Tracking the development of social learning

Jennifer Jane Machin

Thesis submitted for the degree of Doctor of Philosophy

Institute of Neuroscience

Newcastle University

May 2018

Abstract

Background

Developmental shifts in social learning (SL), from selective copying to overimitation, are common in childhood. These differences appear to be both age- and task-dependent. Multiple explanations have been proposed, including participants' understanding of the task goal, affiliation with the demonstrator, and the influence of developing abilities, (memory, theory of mind). Much research focuses on young children or adults, with little understanding into how these behaviours develop during middle childhood and adolescence. Little is known about how participants attend to demonstrations and whether this affects their task performance. This thesis investigated the role of multiple tasks, developing abilities, and attention to demonstrations in children and adults, with the aim of understanding the SL process from beginning to end, and to determine whether a developmental trajectory of copying strategies was present.

Methods

Three studies examined the influence of task types, developing abilities, and attention to a demonstration in typically developing participants aged three to 45. Four tasks were used: a puzzle box, a tool-building task, a puzzle board, and a colouring task. Eye-tracking was used in order to examine the relationship between participants' attention to the demonstration and their task performance. Experiment three investigated participants' understanding of task goals and the demonstrator's intentions.

Results

Minimal differences in eye-tracking patterns between overimitators and selective copiers were observed. Differences in copying strategies were observed between age groups, and appeared to be linked to memory development but not theory of mind. Participants' interpretation of task goals influenced their own task performance. Copying fidelity was task-dependent across all age groups.

Conclusions

Overimitation does not appear to be influenced at the "attention phase", but instead may be driven by participants' memory ability and understanding of task goals. Caution should be used when using one SL task in isolation, as copying behaviours varied across tasks in the same participant groups.

Acknowledgements

To my supervisors, Bess, Vicki, and Sinead: thank you for all your help, support, guidance and wise words for the past four years.

To my family (Mum, Dad, Jess, Nanny J, Grandee F, Nanny W): Thank you for always believing in me, and making me feel like I can achieve anything I set my mind to. I could not have done this without you all. Thanks also to my second family (Lyn, Ian, David and Lesley), who I am very grateful to have.

To Rob: In the words of Daniel Plainview, "I'M FINISHED!" Thank you for your endless patience, for always being there to listen, and for helping me to believe in myself enough that I even attempted this in the first place.

To Alex: What on earth would I have done without you? Maybe now we can finally be breezy (although I'll believe that when I see it)! Thank you for everything. Your turn now!

To The Council (Rachel and Sophie): I am <u>so</u> glad to have you guys in my life. Thank you for your never-ending support, for listening to me rant, and for generally being excellent humans, with excellent taste in music!

To Kat: My mind twin and demonstrating buddy, thank you for being excellent and for making me laugh so hard I've cried on multiple occasions (honk). Hail yourself!

Special thanks also to my non-PhD friends (Fliss, Sarah, John, Helen, and Kieran): Thank you for forcing me to take time out from my PhD when I've needed it, for always taking time to listen to me, and for understanding when I couldn't see you for weeks on end because I had PhD stuff to do.

To A.S: Thank you for all your inspiration and for helping me believe I could really be Dr Jen.

I also wish to thank all of the participants and their families, for giving up their valuable time (often in school holidays!) to take part in my research, and to the Garside Legacy, for providing the funding that allowed me to complete this PhD.

“*That's how you learn. You emulate. You're just a shadow, you're a shell of who you're emulating, and then you fill it with your own personality and ideas, your own ideology, right? It's smart in a way. It's how humans learn.*”
– Ben Kissel

Table of Contents

List of Figures and Tables

Chapter 1: Introduction

Social learning is of vital importance to human culture; it allows us to understand the world around us, as well as objects and individuals in it, to learn new skills, and to build on the work of previous generations. The tendency to learn from one another by copying precise actions, behaviours and gestures appears to be innately human, and has been found cross-culturally, even in remote and isolated tribes (Nielsen & Tomaselli, 2010). Furthermore, the ability to socially learn impacts us all, whether this involves young children learning to write at school, the development of new technology, or learning a new skill or hobby from a more experienced peer. Of particular importance is cumulative culture, which allows an individual or a group of individuals to build on the work of previous generations in order to increase the complexity of an item or an idea. This process is progressive, and develops over multiple iterations, but avoids the necessity to start from scratch each time modifications are made (Shipton & Nielsen, 2015). Cumulative culture has allowed for, amongst other things, the development of complex technology, which we have now come to rely on in our everyday lives, such as computers and mobile phones, and for an increased understanding into complex scientific concepts such as the human genome project, which aimed to fully map the entire DNA sequence of a human being. Neither of these developments could have been conceptualised and carried out by one individual alone. Instead, they were a result of a long process, over years, involving a number of different individuals slowly building on the ideas that had come before them. This process requires the high-fidelity transmission of information, in order that the previous version of the object, idea or piece of technology can be replicated precisely, so that further improvements can be made. For this reason, in recent years, a great deal of attention has been paid to the way in which humans learn precisely from one another, by copying actions, gestures, movements and behaviours from one another. Of particular interest has been the phenomenon of overimitation, in which individuals copy obviously irrelevant actions from others, often resulting in reduced efficiency in task performance.

A vast range of research has been conducted in recent years in order to investigate this phenomenon, alongside other copying behaviours, in children, adults and non-human animals, with the aim of understanding what drives this behaviour, and whether it is limited only to humans. However, due to the vast amount of research conducted, often on very isolated aspects of the behaviour (e.g. restricting studies to particular age groups or tasks), it becomes very difficult to draw conclusions as to what may be driving high fidelity copying in humans across the lifespan, and in different situations. These restrictions have resulted in conflicting findings, and an over-reliance on particular tasks. To date, no single study has aimed to investigate overimitative behaviour across childhood and into adulthood, instead focusing on disparate age groups, meaning that it is difficult to determine how these behaviours develop with age. Additionally, these studies often do not consider the potential contributing factors to these behaviours that participants may face, such as developing memory and theory of mind abilities, and the ability to attend to information provided to them. Furthermore, the majority of studies use only one task, or a variant of the same task, to assess copying behaviours in their participants. This means that even if multiple age groups are tested, it becomes difficult to differentiate between behaviour changes caused due to developmental changes, and those caused due to behavioural shifts that might occur due to the demands of the task. In addition to this, the main focus of many social learning studies has been the behaviours replicated by participants, as opposed to what participants attend to during demonstrations and how that subsequently influences their behaviour. In order to fully understand the process that participants go through when watching demonstrations and then subsequently completing a related task, it would be beneficial to understand whether there are differences in attention to a demonstration between participants who copy faithfully and those who do not. The use of eye tracking during video demonstrations would allow for patterns of attention to subsequently be matched to behaviours displayed by participants when completing the task for themselves.

For these reasons, this thesis aims to use multiple tasks, age groups, and assessments of developing abilities (including memory, theory of mind and attention), to try to paint a broader picture of what exactly may be driving such precise copying behaviours across childhood, and into adulthood, with the aim of explaining the shifts in copying fidelity observed across age groups, and in many cases, from study to study within the literature.

In this introductory chapter, I aim to give an overview of the terms used within the literature to describe the various copying behaviours that may be observed in children and adults, what drives these behaviours, how they may be influenced by developing abilities across childhood, and what methods are appropriate for assessing them. Finally, I will describe the overall aims of this thesis in more detail.

1.1. Distinctions between copying behaviours and definition of terms

When we learn a new skill from someone else, we may choose to copy them precisely, or to use the same broad overall method but with our own adaptations. A number of terms have been identified to distinguish between these high and low fidelity copying behaviours, and these will be outlined further in this section. However, within the social learning literature as a whole, there is often a range of different definitions used to describe the same concept, meaning that a description of imitation in one study might differ from the description of imitation in another study, particularly when switching between the comparative and developmental literature. For this reason, the definitions that I will be using within this thesis are outlined below, in order to avoid confusion. The behaviours discussed in this section are described in the context of an individual observing the behaviour of another (either acting on an object, or performing movements, gestures, or actions that may appear to serve no immediate function). In this section, I will refer to the individual initially displaying this behaviour as the demonstrator, as this is a commonly used term within the literature.

1.1.1. Mimicry

Mimicry has been broadly defined as copying the actions displayed by a demonstrator, without an understanding of the intended goal of the actions (Tomasello, Kruger, & Ratner, 1993), with a further definition describing mimicry as “unconscious or automatic imitation of gestures, behaviours, facial expressions, speech and movements” (van Baaren, Janssen, Chartrand, & Dijksterhuis, 2009). Mimicry is also referred to as “automatic imitation” (Heyes, 2011) within the social learning literature. However, as a precise definition of imitation will be used here, the term “automatic imitation”, which refers to a specific definition as described above, will be avoided in order to reduce confusion.

Although mimicry is not the focus of this thesis, it has been included here in order to fully describe the range of behaviours that may be observed, particularly in younger children. Although mimicry may appear similar to other social learning behaviours, it occurs unconsciously (Chartrand & van Baaren, 2009). It is therefore important to understand how this behaviour can be induced, if individuals are not consciously choosing to mimic others. Wang, Newport, Hamilton, De, & Hamilton (2011) suggested that mimicry can be elicited using ostensive cues, determining that participants were increasingly likely to mimic hand actions performed by a demonstrator if they had made eye contact with that individual. Additionally, participants were more likely to display a motor response, such as wincing, when making eye contact with an individual who had been injured (Eisenberg & Strayer, 1990). This further supports the suggestion that mimicry has a primarily social function: showing affiliation.

As suggested in the definition by van Baaren et al (2009), mimicry can take many forms, with varying complexity. Mimicry of basic facial expressions from adults, such as tongue protrusion and mouth opening, has been observed in newborns under 71 hours old (Meltzoff & Moore, 1989), whilst adults were found to mimic features of angry and happy expressions when shown images and video recordings of angry and happy individuals (Sato & Yoshikawa, 2007). It is suggested that these behaviours occur as a result of the mirror neuron system (Heyes, 2011) and serve a number of social functions, such as establishing rapport between individuals, and understanding the intentions of others (Blairy, Herrera, & Hess, 1999; Hess, Philippot, & Blairy, 1999). This is further supported by the suggestion that unconscious mimicry of facial expressions is impaired in adolescents and adults with Autism Spectrum Conditions, which are characterised by deficits in social behaviour (McIntosh, Reichmann-Decker, Winkielman, & Wilbarger, 2006). There is also some suggestion that motor mimicry in humans, such as wincing when observing another individual's pain, is a primitive form of empathy, thereby further increasing affiliation between the two individuals involved (Eisenberg & Strayer, 1990).

1.1.2. Emulation

Emulation occurs when an individual copies the result or goal state of a demonstration, or learns about the affordances of something within their environment (Carpenter & Call, 2002). Whiten, McGuigan, Marshall-Pescini, & Hopper (2009) suggest that emulation allows an individual to learn about their environment and the results of what others do. One example of emulation refers to what is sometimes called "end-state emulation" or "end-state copying" within the literature, in which an individual produces the same outcome as that achieved by the demonstrator, but without using any of the same actions (Caldwell, Schillinger, Evans, & Hopper, 2012), thereby reproducing the end-state of a task only. An additional type of emulation identified within the literature is object movement re-enactment, in which the actions themselves, rather than any movements or gestures used to perform them, are copied (Custance, Whiten, & Fredman, 1999). Object movement re-enactment is typically investigated using experiments that make use of ghost conditions, in which an object is made to move without the use of a demonstrator, i.e. using fishing wire to move levers rather than showing a demonstrator moving them with their hands (Hopper, Lambeth, Schapiro, & Whiten, 2008; Tennie, Call, & Tomasello, 2006; Custance, Mayer, Kumar, Hill, & Heaton, 2014). This means that participants are only provided with information about object movements/affordances. Replicating object movements is considered to be emulation, whilst replication of actions performed by a demonstrator would be considered to be imitation or overimitation, depending on whether the actions are considered to be relevant or irrelevant for task completion.

This distinction between emulation and imitation is described by Tennie, Call, & Tomasello (2006), who refer to emulation as “results learning” and imitation as “action learning”. One example of emulative behaviour, as described by Jones (2009), might be that a child learns from watching a demonstrator that the door of a dolls’ house can be opened. The focus here is on the fact that the door can be opened, as opposed to the method used to open the door. Therefore, the child may open the door using a method of his or her own (e.g. pushing rather than pulling), as opposed to directly imitating the method used by the demonstrator, therefore achieving the goal state (an open door) via emulation.

1.1.3. Imitation

As previously discussed, imitation refers to the reproduction of movements, gestures or actions from a demonstrator in order to complete a task, using the same method as was demonstrated. It is important to distinguish this from emulation, in which the outcome and/or object affordances (i.e. the inner workings of an object) are copied, rather than the specific actions or gestures used to complete the task. Instead, in imitation, it is the actions used to achieve that goal which are the primary focus (Tennie et al., 2006). However, unlike in mimicry, imitation requires the individual to be consciously aware of the fact that they are copying these actions or movements (Jones, 2009). It has therefore been suggested that imitation broadly serves two purposes: to learn (about the environment or the affordances of objects), and/or to demonstrate affiliation with other individuals (Strouse & Troseth, 2008).

1.1.4. Overimitation

In recent years, a further distinction has been drawn within the social learning literature between imitation and overimitation. Overimitation is an unusual phenomenon in which individuals copy all gestures or actions from a demonstrator, whether relevant to task completion or not, even at the expense of time and efficiency. This is described by Berl et al. (2015) as “the high-fidelity copying of causally irrelevant actions in the presence of clear causal information”. As with imitation, overimitation facilitates the learning of new skills or information, as individuals who overimitate eventually achieve the same outcome as those who imitate, with the difference being that overimitation requires more work to achieve the same goal.

1.2. How are copying behaviours assessed?

1.2.1. Puzzle boxes and artificial fruit tasks

A number of different task types have been used in previous research to investigate emulative, imitative and overimitative behaviour. Perhaps the most commonly used tasks within the literature are puzzle boxes, which usually contain some sort of reward which the participant must attempt to retrieve (Horner & Whiten, 2005; McGuigan, 2013; Nielsen & Hudry, 2010; Simpson & Riggs, 2011b; Watson-Jones, Legare, Whitehouse, & Clegg, 2014). One example of this task type is the artificial fruit task (Custance, Prato-Previde, Spiezio, Rigamonti, & Poli, 2006; Custance et al., 1999), in which two steel bolts must be removed from a small plastic box (referred to as an artificial fruit, as it is intended to mimic foraging problems encountered by primates), in order to allow it to open. The bolts can be removed in a number of ways, such as twisting and poking. In these tasks, a specific sequence of actions is demonstrated to the participants, meaning that the precision of copying can be investigated: in other words, did the participant use the same method to remove the bolts, and in the same sequence? These tasks have been used in adults, children and non-human animals. Additionally, further manipulations may be made, in which participants see either a transparent or an opaque puzzle box, thereby altering the amount of information about the internal mechanisms of the box that is available to participants (Horner & Whiten, 2005).

Puzzle boxes are useful, because they are often novel tasks with which children have no previous experience, meaning that they are often unable to solve them easily on their own, as they have no in-built behavioural repertoires related to these task types (Whiten et al., 2016). In addition, this means that participants may not immediately understand the function or purpose of each of the relevant and irrelevant actions involved, or the inner workings of the box. This makes puzzle box tasks useful for testing hypotheses relating to children's understanding of actions and their subsequent choice to copy them (e.g. the automatic causal encoding hypothesis) (Lyons, Young, & Keil, 2007). However, it can also be suggested that puzzle boxes are very different from more natural social learning situations that children may encounter on a daily basis, due to their artificial nature (Berl et al., 2015; Whiten et al., 2016). An overreliance on puzzle boxes as the only task type used to assess copying fidelity is potentially problematic, as the range of behavioural responses observed might be limited by this task. Additionally, since puzzle boxes are so dissimilar to the types of novel situation encountered by children in their daily lives, the responses observed might simply be specific to puzzle boxes, and in reality, participants may behave very differently if learning to complete a different task. This means that behaviours observed from one singular task type should not be generalised to participants' completion of all task types. However, despite these criticisms of puzzle box tasks, their use is still beneficial, due to their novelty and the manipulations that can be made, such as modifying the amount of information that is available to the participant. When used alongside other more familiar tasks, the concerns relating to ecological relevance can be off-set, and comparisons between task performance can be made to determine how different participants' behaviour is in more natural situations, in comparison to a more artificial puzzle box task.

1.2.2. Tool use and manufacture tasks

An alternative task type that may be used is a tool use task, in which the participant must use a tool in order to retrieve a reward either from inside a box, or inside some other type of container, such as a plastic tube. Puzzle boxes may be used in conjunction with tool use tasks. In this case, they are often composed of a series of defences, which must be removed using a simple tool, for example, using a stick to drag a bolt and slide a door (McGuigan, Whiten, Flynn, & Horner, 2007). Tool use tasks are beneficial, as we encounter tools frequently in day-to-day life, and children learn to use a range of tools (some of which are very complex) rapidly within the first few years of their lives (Cutting, Apperly, Chappell, & Beck, 2014). However, it is also important to consider more complex tool manufacture tasks. One example is a task in which children are required to use pipe cleaners to construct a hook, which can then be used to retrieve a reward from inside a plastic tube (Cutting et al., 2014). These tasks are beneficial as they allow the participant to use a much wider range of actions when constructing the hook tool, in comparison to a puzzle box task, which has a finite number of potential actions. This larger number of potential solutions allows for better investigation into the precision of participants' copying, as well as allowing for a better understanding into how much innovation and creativity, as well as causal understanding, are involved in social learning. For example, they may produce a tool that looks broadly similar (i.e. the same overall shape), but using different materials, or they may solve the task using a completely different tool to the one demonstrated to them. Research suggests that children find it particularly difficult to produce tools on their own, potentially due to the fact that tool manufacture is an ill-structured problem, and therefore these tasks allow us to understand more about which aspects of a demonstration are important to allow participants to succeed (Chappell, Cutting, Apperly & Beck, 2013). Additionally, given the difficulty of tool manufacture tasks, this allows for a more in-depth investigation of overimitative behaviour, since the causal properties of the items used to construct the tools may be more difficult to identify, which is likely to result in increased overimitation in comparison to a causally transparent task (Want & Harris, 2001).

1.2.3. Gesture or movement copying tasks

Tasks with less obvious goal states may also be used to determine what effect they have on participants' behaviour. For example, Dickerson, Gerhardstein, Zack, & Barr (2008), designed a task in which participants use magnetic tangram pieces to produce either a boat or a fish shape on a board. Although this task may not be novel to young children, who are likely to commonly encounter puzzles at home or in school, gestures can be added when completing the puzzle, to determine whether participants overimitate, and copy both gestures and goals, or whether they emulate, and copy the goal state only. The use of gestures, as opposed to potentially functional irrelevant actions, allows for the investigation of overimitation in a different domain to that traditionally investigated in either a puzzle box or tool building task. Additionally, the sequence used by the demonstrator to complete the puzzle is likely to be more arbitrary in comparison to the sequence used to complete a functional task, and so therefore provides a contrast to highly functional overimitation tasks used in most studies. Indeed, some studies rely on gesture alone in order to assess copying behaviours (Bekkering, Wohlschlager, & Gattis, 2000; Huang & Charman, 2005; Mataric & Pomplun, 1998; Paulus, Hunnius, Vissers, & Bekkering, 2011; Wang, Newport, & Hamilton, 2010; Wild, Poliakoff, Jerrison, & Gowen, 2012; Williams, Casey, Braadbaart, Culmer, & Mon-Williams, 2014). In some cases, these movements or gestures are not performed on a target object, in contrast to the puzzle task used by Dickerson et al (2008). This is sometimes referred to within the literature as kinematics. It has been suggested that when a task has no obvious goal, participants will more closely imitate actions or movements, and in fact, the precise copying of these actions becomes the goal of the task (Gowen, 2012). Therefore, it would be expected that these gesture-based tasks with no clear goals would elicit the highest levels of precise copying from participants.

1.2.4. Assessment of different levels of copying fidelity

In many studies examining copying behaviours in children, emulative behaviour without a delay between demonstration and test is induced, using the ghost conditions described previously (Hopper et al., 2008; Thompson & Russell, 2004; Tomasello, Savage-Rumbaugh, & Kruger, 1993). In these conditions, only information about the affordances of an object is available, since any information about the movements of the demonstrator when completing the task has been removed. These tasks are often successful in eliciting emulative behaviours, which could provide a useful contrast against other tasks in which information about object affordances is readily available. These tasks are in direct contrast to the purely gesture-based tasks described previously, which only provide information on movements, and no information on the affordances of objects (since there is no object present in the task).

When investigating imitation in infants and children, two or three step sequences of actions are often used (Barr & Hayne, 1999; Strouse & Troseth, 2008; Taylor & Herbert, 2013). These sequences are useful specifically in the assessment of imitation because they have no clear overall goal, meaning that a clear distinction between goal-state learning and action learning is possible. One example of an action sequence is described by Taylor & Herbert (2013), in which a demonstrator holds a puppet up in front of a child and then proceeds to use the puppet to wave hello, removes a mitten from the puppet and shakes it to ring a bell hidden inside, and then replaces the mitten. In this demonstration, the participant learns that the mitten can be removed, and that shaking it can produce a sound; this information is not obvious when looking at a still image of the puppet. They may then choose to copy these actions (either in or out of sequence), to perform different actions, or to not interact with the puppet at all. These tasks therefore allow for a distinction to be drawn between emulation and imitation. If the participant copies the actions, then they are seen to have imitated. If they do not copy the actions, then they are seen to have emulated, as they have learnt about the affordances of the object but have not used the same strategy as the demonstrator to interact with it.

In addition to the use of gesture-based tasks, overimitation is assessed with the inclusion of irrelevant actions or gestures performed by a demonstrator, alongside the relevant actions or gestures required to complete a task (Lyons, Young, & Keil, 2007; Lyons, Damrosch, Lin, Macris, & Keil, 2011; Nielsen & Tomaselli, 2010). One example of a task involving both relevant and irrelevant actions is a puzzle box task described by Whiten et al. (2016), in which participants were required to use a stick tool to open a puzzle box and retrieve a reward. During the demonstration, participants were shown two irrelevant actions required for task completion: uncovering a hole on the top of the box, and placing the end of the stick tool inside the box. Following this, two relevant actions, moving a defence to uncover a hole on the front of the box, and using the stick to retrieve a reward through this hole, were shown. Participants were described as overimitators if they chose to copy the irrelevant actions alongside the relevant ones.

1.2.5. The use of multiple task types

Although many studies use multiple variants of one task (e.g. multiple puzzle boxes in research by Nielsen & Hudry (2010)), there is little research that examines the differences in task performance by participants completing a number of different task types. As indicated above, the type of task used can greatly influence or limit the range of behaviours seen in participants. For this reason, there is a risk involved when only allowing participants to complete one task, as it is unlikely to allow participants to show a wide range of potential copying strategies. Allowing participants to complete multiple tasks of different types will allow for investigation of a full range of potential copying behaviours within the same participant group.

Additionally, many studies use variations of puzzle boxes, with little exploration of other tasks, such as tool making tasks or those involving puzzle boards. One of the key aims of this thesis is to investigate any changes in participants' copying behaviour that may be task-dependent, by using multiple task types with the same participant population. This is of particular interest given the differences in difficulty between these tasks, which is also likely to influence how faithfully or precisely participants are able to copy aspects of them. Since a hook manufacture task has been identified to be more difficult for younger participants, then it would be expected that if presented with a hook manufacture task and a puzzle box task, these participants may display markedly different behaviour. This would allow for an understanding into how much of their behaviour is driven by the task type itself, rather than their developing abilities, or social influences.

1.2.6. Demonstrations and demonstrator characteristics

Within the social learning literature, both live and video demonstrations have been used, with each having its own set of limitations. Whilst live demonstrations allow for more subtle ostensive cues to be shared between demonstrator and participant, perhaps eliciting a more natural social situation, they do not allow for identical demonstrations between participants in the same way that video demonstrations do (McGuigan, Makinson, & Whiten, 2011). Additionally, although a video deficit effect has been identified in infants, in which they are less able to learn multi-step sequences from a video recording in comparison to a live demonstration, this appears to be reduced by the age of 3, which is the youngest participant age group tested within this thesis (Anderson, 2005a; Barr & Hayne, 1999; Troseth, Saylor, & Archer, 2006).

There is a great deal of research that suggests that the characteristics of the individual demonstrating the task play an important role in influencing copying behaviours; this is known as transmission bias (Wood, Kendal, & Flynn, 2012). Table 1.1. provides an overview of some of the attributes or behaviours of a demonstrator that might encourage an individual to copy more or less faithfully. Since the aim of the experiments within this thesis are not to test demonstrator characteristics, it is important to keep the demonstrator used within each experiment as consistent as possible, both in terms of the individual themselves, and their behaviour when demonstrating the task. Since multiple tasks will be used with the same participant group in this thesis, the demonstrator will be kept consistent across tasks, in order to ensure that any differences seen across tasks are due to the task types themselves, rather than changes in demonstrator behaviour. Furthermore, video demonstrations will be used across all experiments, in order to ensure consistency between demonstrations of the same task across participants. Table 1.1. indicates the importance of maintaining consistency across demonstrations, due to the fact that there are so many potential nuanced influences on participants' behaviour that may be difficult to control for with the use of live demonstrations.

Table 1.1: Demonstrator characteristics influencing that may influence participants' decision to overimitate

Characteristic identified	**Identified by**	**Influence on behaviour**
Age of demonstrator	McGuigan, Makinson, & Whiten (2011); Wood et al. (2012); Zmyj & Seehagen (2013)	Children are more likely to imitate familiar behaviours from peers, but novel behaviours from adults.
Perceived demonstrator success, proficiency or reliability	Wilks, Collier-Baker, & Nielsen (2015); Gergely (2006)	Participants were more likely to copy a demonstrator that was perceived to be successful.
Familiarity with demonstrator	Corriveau & Harris (2009)	Children were more likely to trust, and copy from, a familiar demonstrator.
Presence or absence of an inefficient demonstrator or peer	Nielsen & Blank (2011); McGuigan & Robertson (2015)	Participants imitated an inefficient demonstrator when they were present in the room, but switched to a more efficient strategy when they were no longer present.
Perceived dominance of demonstrator	Hetherington & Mavis (1965)	Participants were more likely to imitate a dominant parent.
Aggressive behaviour by demonstrator	Bandura, Ross, & Ross (1961)	Children displayed more aggressive behaviour when they saw an aggressive demonstrator.
Perceived status of demonstrator	McGuigan (2013)	Children were more likely to copy high status demonstrators.
More or less communicative or friendly demonstrator	Hoehl, Zettersten, Schleihauf, Grätz, & Pauen (2014); Kupán et al. (2017)	Children were more likely to copy friendly demonstrators in comparison to more aloof/unfriendly demonstrators.
Ostensive or pedagogical cues (e.g. eye contact with demonstrator)	Marsh, Ropar, Hamilton, Erdohegyi, & Csibra (2014); Vivanti, Hocking, Fanning, & Dissanayake (2017)	Eye contact with the demonstrator increased copying fidelity.
Number of demonstrators (i.e. group majority/conformity)	McGuigan & Burgess (2017)	When copying from a group of demonstrators, children copied from the group majority as long as they were of a similar age or older than the participant group.
Whether demonstrator belongs to in-group or out-group	Howard, Henderson, Carrazza, & Woodward (2015)	Three year olds were more likely to imitate from in-group demonstrators.

1.3. What drives us to copy faithfully?

Because overimitation initially appears to be so irrational, a number of theories have been identified that aim to explain what may drive individuals to behave in this way, even when it seems maladaptive. These theories are outlined below.

1.3.1. Automatic causal encoding

The automatic causal encoding hypothesis was proposed by Lyons, Young, & Keil (2007) and describes the way in which children have a tendency to interpret all actions as necessary when learning to complete a task for the first time by observing another individual, resulting in automatic overimitation. This theory directly contradicts the suggestion that children imitate for social or affiliative purposes, and also appears to explain why children continue to overimitate even when advised not to. In an experiment by Lyons, Damrosch, Lin, Macris, & Keil (2011), participants aged 4 and 5 were trained not to overimitate prior to being given the opportunity to complete a puzzle box task. During training, participants were praised when they omitted “silly” actions, and guided towards a more efficient solution if they chose to overimitate. Participants were then shown a demonstration of a novel puzzle box, and were given the opportunity to retrieve a reward from it when the demonstrator left the room, in order to reduce any social pressure. Finally, participants were given the opportunity to “race” against a puppet to determine who could retrieve the reward in the quickest time, with the time pressure being intended to reduce overimitation. The majority of children in the experimental condition overimitated, with participants in this condition using a significantly higher number of irrelevant actions in comparison to those in a baseline control condition, who did not view a demonstration. This does appear to suggest that, despite a reduction in social pressure and an increase in time pressure, both of which might be expected to inhibit overimitation to some degree, children were unable to stop themselves from copying. Attempts were also made within this hypothesis to explain situations in which participants did not copy all actions from the demonstrator. Lyons et al. (2007) determined that there are specific circumstances, known as boundary conditions, in which the actions performed go against basic causal principles. The main principle identified here was the contact principle, in which two boxes were used: one which contained a reward, and another which was empty. When these containers were connected by a tube, irrelevant actions performed on the empty box, seemingly in order to retrieve the reward from the second box, were more likely to be imitated by participants. However, when the boxes were not connected, participants tended not to copy actions performed on the empty box, as there was no obvious way that they could affect the box containing the reward, since the two boxes were completely separate objects. However, for participants to determine that there was a contact principle involved in a task, they must be rationally interpreting the actions involved in some way, suggesting that the process is not completely automatic or subconscious.

1.3.2. Rational normative action interpretation hypothesis

Gergely & Csibra (2003) describe the “mentalising stance” used by adults when interpreting actions, in order to infer the beliefs and intentions of the demonstrator. The development of this mentalising stance is reliant on theory of mind, and allows adult participants to determine what actions they believe should be copied following a demonstration, based on their interpretation of the demonstrator’s actions. However, since young children have not yet fully developed theory of mind, it is likely that they use an alternative strategy when determining which actions to copy and which to omit. A teleological stance has been suggested, in which children are able to interpret situational constraints, alongside their own desires, in order to use a rational approach to determine what to copy following a demonstration. This is represented in Figure 1.1 below.

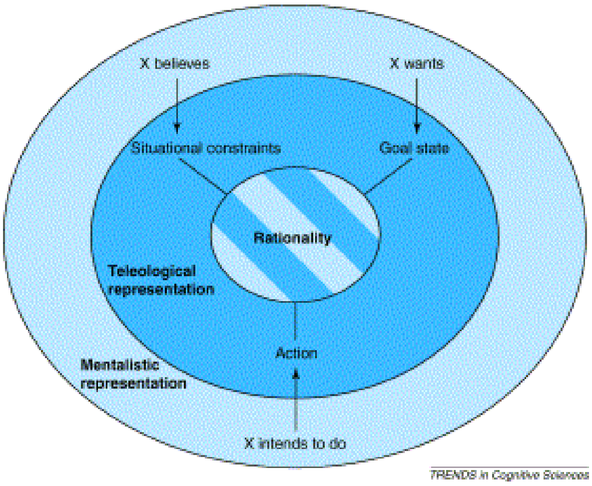


Figure 1.1: Mentalising and teleological stances used when deciding to imitate an action (from Gergely & Csibra, 2003)

This rational interpretation of actions has been further investigated by Keupp, Bancken, Schillmöller, Rakoczy, & Behne (2016), who suggest that children overimitate because they perceive the irrelevant actions to be an important part of a conventional activity. In other words, that these actions are included as part of the social norms of completing this particular task, and so must be repeated in order to complete the task "correctly". In this way, participants are more focused on the ritual of completing a task, as opposed to its instrumental purpose. They therefore interpret the goal of the task to be to imitate all of the actions used, rather than to simply achieve the same outcome as the demonstrator, (e.g. solving a puzzle and retrieving a reward) (Herrmann, Legare, Harris, & Whitehouse, 2013). Additional research determined that when a task was framed in a conventional way, 3- to 6-year-olds were more likely to display increased imitative fidelity in comparison to when a task was framed in an instrumental way (Herrmann et al., 2013). Furthermore, children aged 4 to 6 were also able to determine when a task was intended to be completed in a conventional manner, and when it was intended to be instrumental, using causal information provided about the actions involved in the task. When clear causal information relating to an action was not present, participants interpreted performance of this action as conventional. As before, participants were found to imitate more when the task was intended to be completed conventionally (Legare, Wen, Herrmann, & Whitehouse, 2015). This ability to understand both conventional and instrumental tasks may also partially explain why children are able to display flexible copying behaviours across multiple tasks.

1.3.3. Social or Affiliative theories

In addition to theories investigating participants' understanding of individual actions during a demonstration, social and affiliative theories of overimitation have been discussed, in which participants are aware that an action is irrational or unnecessary, but choose to copy it in order to show affiliation with the demonstrator (Nielsen & Blank, 2011; Wilks et al., 2015). Over & Carpenter (2012) suggest that as well as a child's own learning goals, the most important factors influencing their decision to overimitate are their identification with the model, and any social pressure they experience during either the demonstration or completion of the task. In experiments where the demonstrator appeared aloof and unfriendly, children were found to be less likely to overimitate in comparison to those where the demonstrator made an effort to appear friendly towards participants (Nielsen, 2006), suggesting a strong social influence. Additionally, infants were more likely to overimitate following a live demonstration, with a socially engaged demonstrator that could provide real-time feedback, in contrast to a video demonstration of the same task (Nielsen, Simcock, & Jenkins, 2008). In some situations, it has been suggested that overimitation occurs as a result of conformist behaviour, with research determining that children adapt the strategy used to complete a task, depending on whether they are in the presence of a majority of efficient or inefficient demonstrators (McGuigan, Gladstone, Cook, Macris, & Keil, 2012). It is suggested that this conformist bias is generally beneficial as it would be expected that if a technique is being used by the majority of individuals, then it is likely to be the most successful method of completing that task (Claidière & Whiten, 2012).

Further support for social theories is found in research with children with autism spectrum conditions (ASC), who do not appear to overimitate, potentially due to their lack of motivation to demonstrate affiliation with the demonstrator (Marsh, Pearson, Ropar, & Hamilton, 2013). Marsh et al (2013) used familiar tasks with children with ASC, thereby separating the need to learn information about a novel task from the social component of watching a demonstration. ASC participants were found to demonstrate a reduction in overimitation in the social task, in comparison to performance on a novel task, further supporting the idea that overimitation has an underlying social basis. Additionally, it has been suggested that ASC participants may also overimitate less due to a reduced capability for theory of mind, meaning that they are less able to understand the intentions of the demonstrator (Spengler, Bird, & Brass, 2010).

1.3.4. Unspecified purpose hypothesis

The unspecified purpose hypothesis suggests that individuals imitate actions when their purpose is unclear (Horner & Whiten, 2005). Although this explanation may appear similar to that provided by the automatic causal encoding hypothesis, in the unspecified purpose hypothesis there is some suggestion that participants understand that the action being copied may be unnecessary, but that it is being intentionally performed by the demonstrator for a reason, and therefore merits being copied (Kenward, 2012; Kenward, Karlsson, & Persson, 2011; Marsh, Ropar, Hamilton, Erdohegyi, & Csibra, 2014). In this way, this hypothesis also bears some similarity to the rational normative action interpretation hypothesis, with the difference being that in the unspecified purpose hypothesis, the focus is less on the ritual of performing the task, and more on the fact that the action being copied may actually serve some purpose in completing the task itself.

Although each of these theories may provide a good explanation for overimitative behaviour in different contexts, it is unlikely that any one theory can fully explain a particular copying behaviour across all tasks, but rather that a number of different explanations should be considered in conjunction with one another. This is likely to be dependent on the task and the context being used.

1.4. Benefits of high fidelity copying

At times it can be difficult to understand why overimitation is such a persistent behaviour, since copying irrelevant actions may mean reduced efficiency and an increased time cost when completing a task. However, as previously stated, copying faithfully from another individual can have positive outcomes, both in terms of building on the ideas of previous generations, and in demonstrating affiliative behaviours with others. This section aims to outline some of the beneficial aspects of high fidelity copying.

1.4.1. Cumulative culture

Cumulative culture is the process by which information is transmitted across generations, with each generation making minor improvements to the process, described as the "ratchet effect". As previously described, this process relies strongly on high fidelity copying of the original behaviour (Lewis & Laland, 2012; Tennie, Call, & Tomasello, 2009), meaning that actions, gestures, movements or ideas are copied very precisely, in order to replicate a previous state or outcome, and subsequent improvements are then made. Since overimitation is the most faithful method of copying, in which all actions (both rational and irrational) are copied from another individual, there is some suggestion that this behaviour developed specifically to support cumulative culture (Legare & Nielsen, 2015). Cumulative culture is of particular importance in human society, as it allows for the rapid development of tools, knowledge, and technology from one generation to the next, without each generation needing to begin again from scratch (Dean, Kendal, Schapiro, Thierry, & Laland, 2012). Although a number of attempts have been made to identify cumulative culture in non-human animals, to date no clear evidence of this process, or of overimitation, has been found (Dean et al., 2012; Dean, Vale, Laland, Flynn, & Kendal, 2014; Marshall-Pescini & Whiten, 2008). It is suggested that whilst non-human animals have the ability to innovate, only humans display faithful enough copying behaviours to allow cumulative culture and the ratchet effect to take place (Tomasello, Kruger, et al., 1993). Further research has also suggested that humans' capability for communication through language, as well as prosocial behaviour, may be additional underlying factors that allow cumulative culture to occur (Dean et al., 2012).

1.4.2. Ritual

Ritual can be defined as "culturally standard, repetitive activity, primarily symbolic in character" (Kertzer, 1988). Rituals are an important part of everyday life for infants and young children that allow them to learn daily routines, such as bathtime and bedtime routines, from their caregivers (Nielsen, Kapitány, & Elkins, 2014). Copying from others may be an important part of learning these ritual behaviours, which require that particular steps are undertaken in a certain order. Additionally, research indicates that when children perceive the behaviours shown during a demonstration of a task to be conventional or ritualistic, they are more likely to then display overimitative behaviour when performing the task themselves (Byrne & Russon, 1998; Herrmann et al., 2013; Legare & Nielsen, 2015). Ritualistic behaviour is also seen in adulthood, in a wide range of social contexts, from passing on a family recipe to religious rituals (Gergely, 2006; Laor, 2013). In order to sufficiently transmit ritualistic behaviour, which may at times appear irrational to an outside viewer, high fidelity copying is required (Hoehl, Zettersten, Schleihauf, Grätz, & Pauen, 2014).

1.4.3. Affiliation and conformity

As discussed above, faithful copying may be linked to the performance of rituals, which are a way of developing in-group affiliation (Wen, Herrmann, & Legare, 2016). The influence of conformity on social behaviour in general has been widely studied, with participants tending to agree with the answer provided by the majority of a group, even if that answer appears obviously incorrect (Asch, 1956). The effect of conformity on overimitation in children has been studied, with children being likely to copy irrelevant actions from members of their own assigned group, but not from members of another group (Gruber, Deschenaux, Frick, & Clément, 2017). Yu & Kushnir (2014) suggest that children as young as two are more likely to copy faithfully in a social situation, for example, in a mimicry game involving hand gestures, in comparison with a more instrumental task, involving the opening of a box. Additionally, young children are sensitive to behaviour that may cause them to be ostracised, and therefore copy more closely when ostracism may occur if they do not do so (Watson-Jones, Legare, Whitehouse, & Clegg, 2014). Additionally, children have been found to copy faithfully when irrelevant actions are demonstrated by the majority of their peers (McGuigan & Robertson, 2015), even switching from an efficient strategy to do so, suggesting that they were performing these actions in order to display affiliative behaviour with their peers. Being able to demonstrate this kind of affiliative behaviour is extremely beneficial from a social and affiliative perspective, as research has suggested that individuals are at risk of ostracism if they do not participate in group activities (Whitehouse & Legare, 2013). In the past, maintaining social connections was vitally important, as being part of a group allowed for group hunting and herding, food foraging, and exchanges and bartering, thereby increasing overall fitness and the potential for survival (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012). In modern society, affiliation and being part of a social group prevents isolation and loneliness, which can have a detrimental effect on overall health and wellbeing (Cacioppo & Patrick, 2008).

1.5. How do copying behaviours develop throughout childhood?

1.5.1. Developmental trajectory

As identified in section 1.1., children have been observed to use different copying strategies as they develop. There is a risk in defining participants of a particular age as being always emulative or overimitative, since there is likely to be a range of other influences on their behaviour. Therefore, the aim of this section is not to suggest that children of a particular age will always behave in a particular way, but instead to describe the development in cognitive abilities that might affect a child's ability to display certain copying behaviours throughout ontogeny.. Table 1.2 indicates a potential developmental trajectory of copying behaviours identified in infants and children within the social learning literature. This is not designed to be a comprehensive overview, but is intended to indicate the rough ages at which each of the copying behaviours might arise. Additionally, once an infant or child develops a skill, they may still copy flexibly, meaning that in different situations (for example, in the presence of different demonstrators, as indicated in table 1.1), they may mimic, emulate or imitate, depending on the task and the environment.

Table 1.2: Developmental trajectory of copying behaviours present in infants and children

	Mimicry	Emulation	Imitation	Overimitation
Newborn	Mimicry of facial expressions, e.g. tongue protrusions and mouth opening in newborns up to 72 hours old. (Meltzoff & Moore, 1989)			
6 months	However, Jones (2009) argues that mimicry of actions modelled by parents was not present at 6 months.			
9-12 months	9 month old infants were observed to be capable of mimicry of actions (Longo & Bertenthal, 2006)	Actions copied in terms of outcomes, i.e. emulation, around 12 months. (Malinda Carpenter, Call, & Tomasello, 2005) Infants aged 17 months used a primarily emulative strategy to complete a task (Huang & Charman, 2005)	9-12 month olds begin to understand the relationship between actions and goals (Gergely & Csibra, 2006) Infants become capable of imitating some target actions within a three step sequence at 9-15 months (Elsner, Hauf, & Aschersleben, 2007) Infants begin to imitate "action + result" at 15 months (Esseily, Nadel, & Fagard, 2010)	
18 months			Intentional actions imitated between the age of 14-18 months. (Malinda Carpenter, Akhtar, & Tomasello, 1998) However, this behaviour can be moderated by demonstrator characteristics (Buttelmann, Zmyj, Daum, & Carpenter, 2013; Kupán et al., 2017; Zmyj, Daum, & Prinz, 2012)	
2 years			Children aged 23-30 months performed only the relevant actions when completing a puzzle box task. (McGuigan & Whiten, 2009)	24 month olds copied specific relevant and irrelevant actions from a demonstrator, but again this was modulated by demonstrator characteristics. (Nielsen, 2006)
3-4 years		3 year olds displayed an emulative strategy when available information was degraded (McGuigan, Whiten, Flynn, & Horner, 2007b)	3-4 year olds displayed immediate imitation but delayed emulation (Simpson & Riggs, 2011b)	3, 4, and 5 year olds were found to display overimitative behaviour (Keupp, Behne, & Rakoczy, 2013; Yue Yu & Kushnir, 2014)
5 years				5 year olds were found to overimitate. This behaviour increased further with age (L. E. Marsh et al., 2014)

To date, the majority of research into high and low fidelity copying behaviours has focused on infants, young children, and adults, with little attention being paid to copying strategies used in middle childhood and adolescence. Research into overimitation in adults has produced contradictory results, with some studies suggesting that adult participants display increased levels of precise copying from a demonstrator, including actions irrelevant to task completion, in comparison to children (Custance et al., 2006; McGuigan et al., 2011b; Whiten et al., 2016b). Others suggest that adults display reduced levels of overimitation in comparison to children (Horowitz et al., 2003). This may be task-dependent, or context-dependent, and more research is required to determine at what age copying behaviours begin to become more "mature", as children develop the ability to remember demonstrations, to attend to relevant aspects of a demonstration, and to further understand the intentions of the demonstrator. Still-developing memory abilities may prevent younger children from remembering actions or gestures that they have witnessed, meaning that they are unable to subsequently copy them later. This suggestion is supported by findings by Simpson & Riggs (2011), which appear to suggest that 3 and 4 year old participants imitated immediately following a task demonstration, but emulated after a delay when presented with the same task again. An additional factor affecting overimitation is the development of theory of mind, which allows participants to understand the intentions of others (Meltzoff, 2002). This is particularly important for socially-motivated aspects of copying behaviours, in which participants wish to display affiliative behaviour. In order to fully understand why copying behaviours change across age groups, these underlying developmental factors must also be investigated.

1.5.2. Memory for sequences of actions, and the content of demonstrations

Episodic memory is a memory system that "allows people to consciously re-experience past experiences", and is "about happenings in particular places at particular times, or about "what," "where," and "when"" (Tulving, 2002). It begins to fully develop between the ages of 3 and 5 (Mullally & Maguire, 2014; Scarf, Gross, Colombo, & Hayne, 2013), around the same time as children begin to display overimitative strategies during copying tasks (see Table 1.2). Episodic memory is considered to be an important component of imitative behaviour, with deferred imitation tasks often being used to assess proto-episodic memory ability from a “what-where-when” standpoint. In these tasks, participants are required to remember a sequence of actions (what), and perform them in order (when). The context in which the participant views the demonstration and completes the action sequence may also be altered in order to assess the “where” component (Burns, Russell, & Russell, 2014). In order to assess episodic memory rather than short term working memory, a delay is implemented between the demonstration and participants’ completion of the task (Barr, Dowden, & Hayne, 1996). This delay may be as short as 30 minutes, or as long as 24 hours (Nakano & Kitazawa, 2017; Scarf et al., 2013). These tasks, involving a delay, are known as deferred imitation tasks, These tasks are particularly useful with infants and young children, as they do not rely on verbal cues and therefore fully developed language abilities (Taylor & Herbert, 2014). However, whilst this is a useful method for assessing memory, little attention is paid to the methods used by the child to perform the task. This was highlighted in a criticism of deferred imitation tasks made by Bauer & Kleinknecht (2002), who described how a child using an emulative strategy in a deferred imitation activity would be seen to have “failed” the task, despite achieving the same outcome as the demonstrator. This may lead to the suggestion that the child participating in the task is not capable of forming episodic memories accurately. This is due to the fact that the action used to achieve the outcome are being assessed, ignoring whether the child has achieved the goal state through some other method. The child may have remembered the intended outcome, but chosen to use a different strategy to achieve it. Therefore, in the context of the experiments in this thesis, it would be beneficial to use a separate task to determine participants’ capability for remembering a sequence of actions or steps alongside the tasks assessing copying strategies themselves, rather than combining both into a single task in which one may negate the effects of the other. Since the delay between the task and memory tests used within this thesis is minimal, this cannot truly be referred to as episodic memory, but still relates to the “what” and “when” aspects of memory: what actions were included in the demonstration (content of the task), and when in the sequence did they take place?

1.5.3. Working Memory

In contrast to episodic memory, working memory is defined as “a temporary storage of information while other cognitive tasks are being performed” (Siegel & Ryan, 1989) and is a component of executive function, which is defined as “the ability to execute appropriate actions and to inhibit inappropriate actions for the attainment of a specific goal" (Moriguchi, 2014). Working memory has been identified as a factor that may influence overimitative behaviour, as when the working memory load involved in a task is high (for example, a sequence with a large number of steps involved), participants are more likely to display emulative strategies in order to complete a task (Clay & Tennie, 2017; Subiaul & Schilder, 2014). Previous research has suggested that working memory rapidly develops between the ages of 2 and 5 (Isquith, Gioia, & Espy, 2004), showing improvements around the age of six years old (Gathercole, Pickering, Ambridge, & Wearing, 2004), which also coincides with increases in overimitative behaviour (Moriguchi, 2014). Working memory continues to develop into adolescence and adulthood, allowing for the completion of tasks that require more complex processing (Luciana, Conklin, Hooper, & Yarger, 2005). Working memory is also involved in the reproduction of a sequence of actions. In order to overimitate, individuals must often produce not only the same actions as a demonstrator, but also must reproduce the actions in the same order (Carmo et al., 2017; Subiaul & Schilder, 2014).

1.5.4. Theory of Mind

Theory of mind, or the ability to understand the intentions of others, is thought to develop around the age of 4 (Baillargeon, Scott, & He, 2010). A link has been established between theory of mind and overimitation (Frith & Frith, 2012; Nielsen, Cucchiaro, & Mohamedally, 2012), and imitation of actions in infants has been cited as a precursor to the development of theory of mind (Charman et al., 2000). It may be possible that the development of theory of mind increases overimitative behaviour in social contexts, in which it may be beneficial for individuals to copy a demonstrator in order to display affiliative behaviour (Santiesteban et al., 2012). However, in contrast, participants who have developed theory of mind, and therefore a better understanding of the demonstrator's intended goal, may be less likely to overimitate as they may be more likely to understand that copying the irrelevant actions is not rational, nor is it required for the goal state to be achieved. This is likely to also be related to the development of executive function, which develops around the same time, as discussed above, and allows individuals to inhibit performance of inappropriate actions (Moriguchi, 2014).

1.5.5 Attention

The ability to attend to relevant stimuli increases with age (Hagen & Hale, 1973), which may suggest that the increase in overimitation with age that has been observed in previous studies may simply be due to the fact that older participants are better able to direct their attention towards more relevant aspects of the demonstration. Previous research with infants and ASC children has suggested that there are some differences in gaze patterns between participants who copied more faithfully and those who did not, in which individuals with ASCs attended less to the demonstrator's face and to non-meaningful actions in comparison to typically developing participants (Taylor & Herbert, 2014; Vivanti, Nadig, Ozonoff, & Rogers, 2008). Therefore, as their ability to attend more successfully to important aspects of the demonstration video develops, so too could participants' copying fidelity. One way to assess this is through the use of eye tracking during video demonstrations, which would allow for differences in attention to be examined between participants who copied faithfully and those who did not. Additionally, the use of eye tracking might allow for further understanding into the point at which an individual can be identified as an overimitator or an emulator. Is it the case that overimitators simply look at different aspects of the demonstration than emulators, and subsequently use that information to copy more faithfully, or do both participant groups look at the same areas of interest and then subsequently behave differently due to additional underlying factors? The use of this methodology, along with multiple tasks and assessments of developing abilities such as theory of mind, would allow for a better understanding of the full "end to end" process involved in watching a demonstration and then subsequently completing a task. It may be the case that there is one driving factor alone, or multiple influences on behaviour, but without some measurement of the full process, it becomes very difficult to make that assessment.

Although much of the research into high and low fidelity copying is conducted in young children, during the time at which memory and theory of mind are rapidly developing, measures to assess these abilities are often not included as part of the testing process. In order to fully capture the range of ability in participants, one aim of this thesis is to include assessments of theory of mind, attention and memory alongside tasks which will measure participants' copying fidelity. This will allow for comparisons between participants who have developed these abilities, and those who have not, and should provide further insight into how influential these abilities truly are.

1.6. Aims of this thesis

This thesis contains three experiments designed to add to existing knowledge relating to a range of copying behaviours in children and adults, and what may influence these. In order to achieve this, four aims have been identified. These are outlined in more detail below.

1.6.1. The influence of task type on copying behaviours

As previously mentioned, a variety of task types have been used in previous research into overimitation. However, most studies focus on one task in isolation, and there is little research that combines multiple task types to determine whether participants' behaviour changes depending on the context of the task, and any differences in difficulty or complexity. Additionally, no study to date has combined these multiple task types with a wide age range of participants and an investigation into what participants attend to during the initial demonstration of the task, to provide a full picture of the influences on participants' copying fidelity. Additionally, there is potentially an over-reliance on puzzle box tasks in the social learning literature, and it remains unclear as to whether these are the most appropriate tasks to assess whether participants would be likely to overimitate in their day to day life when learning something new. The frequent use of these tasks means that there is a risk that a given age group, for example, is at risk of being identified as overimitators, when in fact, this may not be the case in every situation. The comparison of a puzzle box task with other task types, including some more arbitrary tasks, will allow for a full range of potential behaviours to be investigated. One purpose of the experiments in chapters two and three is to investigate whether participants are able to display flexible copying behaviours across different task types, and what influences them to do so.

1.6.2. Developmental trajectory of copying behaviours and participants' understanding of task goals

Although the majority of research is conducted with young children (typically aged 2 to 5 years old) and adults, little attention has been paid to copying strategies used during middle-late childhood and early adolescence in typically developing participants. Experiment two aimed to investigate copying behaviours displayed by participants aged 3 to 21, in a variety of tasks, to further add to our knowledge of how social learning behaviours change with age. Additional measures, including memory (including working memory, and memory for sequences of actions and the content of a demonstration) and theory of mind were also taken, to determine what might be influencing these changes in both younger and older participants.

The experiment in chapter four aims to further investigate participants' understanding of the goal of a puzzle box task, as well as the functionality of the actions involved, and how participants' task performance changed across multiple attempts at the same task.

1.6.3. The role of attention during a demonstration, and its influence on task performance

Eye tracking has been used in very few studies investigating copying strategies used by participants, with most eye tracking studies being used either in very young infants or in individuals with Autism Spectrum Conditions. Eye tracking is used in these studies for a number of reasons: to avoid the need for language-based assessments, which are inappropriate for use with infants, and to investigate the differences often seen in ASC children when compared with typically developing children. However, in this case, eye tracking may prove beneficial in understanding the mechanisms behind the different social learning behaviours that are often observed in children and adults. After watching a demonstration, it would be highly unlikely that an individual would be able to describe precisely what they attended to during the demonstration itself. For this reason, this thesis aims to use eye tracking to investigate the role of the demonstration in typically developing populations, as a direct assessment of what participants attend to and how this shapes their behaviour when they were given the opportunity to complete the same task themselves.

In chapters two and three, eye tracking will be used to determine whether attending to particular areas of a demonstration video influenced participants' subsequent task performance, with particular reference to copying fidelity. More specifically, the intention is to determine whether those who copy faithfully (imitators or overimitators) attend more to the movements performed by the demonstrator, and whether those who used their own strategy (emulators, using object movement re-enactment), attend more to the functional aspects of the task (e.g. the way in which the internal mechanisms of a puzzle box operated). Additionally, it may be possible that an assessment of social influences can be made from this data, if participants who copy more faithfully attend more to the demonstrator's face in comparison to those who do not. If this is the case, then it would suggest that it is attention to particular aspects of a demonstration that drives faithful copying behaviours, as opposed to developmental factors (such as the ability to remember and repeat the actions shown) or a conscious decision to copy or omit actions. Although previous research has partially attempted to determine this (Rigamonti, Custance, Previde, & Spiezio, 2005), only the amount of attention paid to the demonstration video as a whole was recorded, meaning that it was unclear as to what participants were actually attending to, and how that subsequently affected their behaviour. Therefore, in experiments one and two, specific areas of interest on the demonstration videos shown to participants were identified, with the aim being to attempt to assess whether the difference in participants who emulate or overimitate begins at the "attention phase", as discussed above. As previously stated, the inclusion of a measurement of attention to the demonstration video alongside other factors will provide a fuller understanding into the influences on copying fidelity throughout the full learning process, from the start (viewing the demonstration), to the end (completing the task).

A further aim is to determine whether attention to the demonstration changed with age, and if so, how this subsequently changed with age. As the ability to attend to relevant stimuli is understood to improve with age (Hagen & Hale, 1973; Plude & Doussard-Roosevelt, 1989), then it would be expected that the length of time attending to the demonstration itself would increase with age, as would attention to relevant aspects (either the task or the movements used to operate it) of the demonstration. However, to date, no study has examined this across a wide age span, using eye tracking technology with typically developing children. For this reason, in chapter 3, participants aged 3 to 21 underwent eye tracking during the demonstration of three different tasks, in order to investigate whether attention to the task changed with age, as well as whether attention to specific areas of interest during a demonstration then resulted in more or less overimitative behaviour. This was investigated in conjunction with additional abilities that develop during early childhood, such as the ability to remember sequences, and the ability to understand the motivations of others, with the aim of providing a full picture into what drives these observed differences in behaviour.

It is hoped that the combination of these aims will allow for a fuller understanding into the process that an individual goes through when they learn to complete a novel task from someone else, and that this will further help us to understand the marked social learning shifts that are seen across the lifespan and between tasks.

Chapter 2. Experiment 1

2.1. Introduction

Overimitation is a widely studied phenomenon, in which individuals copy causally irrelevant actions from a demonstrator when learning to complete a novel task (Lyons et al., 2007; McGuigan et al., 2011b; Whiten et al., 2009). It is suggested that overimitative behaviour is required for the development of cumulative culture in humans, in which one individual can build on the work of others in order to produce a novel item, even though it may not have been possible for them to invent such an item on their own (Tennie, Call, & Tomasello, 2009). It appears that both overimitation and cumulative culture are uniquely human phenomena, which have not been observed in non-human animals to date (Lewis & Laland, 2012; Logan, Breen, Taylor, Gray, & Hoppitt, 2016). The overall aim of this experiment is to understand more about what drives potential overimitative behaviour in both children and adults. It is therefore important to briefly distinguish this overimitative behaviour from a number of additional copying behaviours described within the social learning literature, although these differences are described fully in Chapter 1.

Mimicry can be described as the way in which an individual may copy the movements or actions displayed by someone else, but without an understanding of their purpose (Huang & Charman, 2005). This is in contrast to emulation, in which an individual understands the intended goal of a task, and achieves this goal state, but without using the methods demonstrated to them (Horner & Whiten, 2005). In imitation, individuals copy the relevant actions required for task completion only, omitting any superfluous actions whilst completing the task set for them (Carpenter, Nagell, & Tomasello, 1998). It is important to note that imitation is distinguished from mimicry by an individuals' understanding of the goal of the task or the intentions of the demonstrator: in mimicry, this understanding is not present (Tomasello, Kruger, et al., 1993).

There is some suggestion that mimicry, emulation and the eventual intentional control of imitation form a developmental trajectory, with mimicry of rudimentary gestures even being observed in newborn infants (Meltzoff & Moore, 1989). Emulative behaviour has been observed in 12-month-old infants (Tennie et al., 2006), whilst imitation of simple actions has been observed in infants around the age of 14 months, although this is dependent on the complexity of the task and its intended goal (Sevlever & Gillis, 2010). The introduction of causally irrelevant actions in order to attempt to induce overimitation appears to result in further discrepancies in behaviour in young children and adults. Research by McGuigan, Whiten, Flynn, & Horner (2007) suggests that in particular circumstances, such as when the information provided to participants about a task is degraded, 3 year olds display emulative strategies in order to achieve a solution, whereas 5 year olds employ an overimitative approach, copying causally irrelevant actions. Intriguingly, adults were also found to overimitate, copying causally irrelevant aspects of a demonstration at an even higher frequency than both 3 and 5 year old children, even when time pressure was induced (Flynn & Smith, 2012; McGuigan et al., 2011b). As previously discussed, one potential explanation for this behaviour focuses on social or affiliative factors, in which individuals copy in order to display affiliation with the demonstrator (Nielsen & Blank, 2011; Over & Carpenter, 2012). The changes in observed behaviour may occur as 3 year old children have not fully developed theory of mind, defined as "understanding of persons' mental states" (Wellman & Liu, 2004), which can either refer to desires, beliefs or emotions (Wellman, Cross, & Watson, 2001). This is in contrast to typically developing 5 year old children, who do possess this ability to some degree (Rakoczy & Schmidt, 2013), allowing them to pass false belief tasks. However, theory of mind continues to develop during childhood and into adolescence to allow for a more nuanced understanding of the behaviour of others, particularly in terms of understanding the emotions of others (Vetter, Altgassen, Phillips, Mahy, & Kliegel, 2013). The suggestion that overimitative behaviour is linked to the development of theory of mind is further supported by findings that suggest that children with autism, who often display social deficits, including those relating to theory of mind, do not overimitate (Marsh, Pearson, Ropar, & Hamilton, 2013.; Spengler, Bird, & Brass, 2010); although these findings do not appear to be consistent across tasks (see Nielsen & Hudry (2010)).

Although it is likely that the development of abilities such as theory of mind do play an important role in participants' use of copying strategies, this is further complicated by individuals who display flexible imitation depending on the circumstances involved, indicating that overimitation does not become an automatic process once theory of mind is achieved. Research by Keupp, Behne, Zachow, Kasbohm, & Rakoczy (2015) determined that children's choice to overimitate is also context-sensitive, meaning that participants' understanding of the goal and the actions performed did influence their subsequent behaviour. However, understanding the motivations of very young children when completing these tasks is difficult, due to their still-developing language skills, which may prevent them from articulating why they chose to behave in a particular way. One potential way to investigate this further may be eye tracking, conducted during the demonstration process, to allow further understanding into what exactly participants focus on when they are learning to complete a novel task. Although it is possible to make assumptions regarding what participants attend to during these demonstrations based on participants' behavioural outcomes when completing a task, there has been little investigation into what is actually attended to, either in typically developing children or adults. Eye tracking is suitable for use in both infants and young children, as it is not dependent on language cues, and does not require participants to explain their behaviour (Feng, 2011). Furthermore, information gathered via eye tracking is implicit, and therefore allows researchers to access subconscious processes that participants may not actively be able to articulate, even with a good grasp of spoken language (Odean, Nazareth, & Pruden, 2015; Roderer & Roebers, 2010; Roebers, Schmid, & Roderer, 2010). Of particular interest is participants' tendency to focus on action information during a demonstration, allowing them to further understand how an object works. There is some suggestion that when action information is withheld from participants, they have a tendency to adopt an emulative strategy. In contrast, when action information is freely available, participants are more likely to overimitate (Reindl, Apperly, Beck, & Tennie, 2017). However, this does not account for participants' ability to attend to relevant aspects of a demonstration. Although action information may be freely available during a demonstration, younger participants may struggle to determine what aspects of the demonstration are most relevant (Badger & Shapiro, 2012), especially when faced with a novel task, meaning that they may unintentionally ignore important action information altogether. Furthermore, young children such as those frequently used in overimitation studies are still developing the ability to sustain their attention during demonstrations (McClelland, Acock, Piccinin, Rhea, & Stallings, 2013), meaning that again, although action information may be present, it may not be attended to fully. Collection of eye tracking data during video demonstrations of novel tasks will provide further understanding as to how much of a demonstration a participant attended to, as well as the areas of interest observed.

In addition to more general information on participants' ability to attend to relevant stimuli during a task demonstration, eye tracking may uncover differences in gaze patterns between participants who overimitate and those who do not. Previous research with infants has already used eye tracking during video demonstrations of imitation tasks, albeit in a slightly different context, as deferred imitation is often used as a measure of rudimentary episodic memory (what-where-when) in infants (Taylor & Herbert, 2014). Taylor & Herbert (2014) determined that infants defined as imitators were more likely to attend to relevant areas of the demonstration video, as opposed to emulators, who attended more to irrelevant aspects of the demonstration, such as the background. In contrast, research by Vivanti, Hocking, Fanning, & Dissanayake (2017) determined that typically developing children who overimitated were more likely to look at the demonstrator's face during a demonstration. Using clearly defined areas of interest, it may be possible to further expand upon this research in order to investigate differences in looking patterns at action information as the task is being performed. If a focus on action information is important for overimitation, as described by Reindl et al. (2017), then it may be expected that participants who overimitate will focus more on movements produced by the demonstrator themselves (e.g. the demonstrator's hands operating a task) as opposed to the inner workings of a task itself. In contrast, emulators may show a more generalised looking approach, investigating multiple areas of interest, or focusing specifically on the goal state (e.g. a reward to be obtained upon completion of the task).

As the use of eye tracking in this area of social learning research is so limited, particularly in the age groups used within this field of social learning research, it would be beneficial to conduct further research to investigate these potential differences between participants. However, further considerations must be addressed, namely the suggestion that overimitation is context-sensitive, as previously discussed. Additionally, considerations must be made into how to obtain populations of participants that are likely to overimitate or emulate in order to make a fair comparison between the two. One way to further explore context sensitivity could be to manipulate the task type, using a goal-driven task with a clearly defined number of steps, in comparison with a task which still has an end-stage goal and can be achieved more flexibly, using a greater variety of potential solutions. As previously described, the majority of the social learning literature focuses on 3 and 5 year old children, due to the rapid development of abilities such as episodic memory and theory of mind during this time period, as well as the contrasting copying behaviours displayed by participants of these age groups. For this reason, these age groups will be used to allow for comparison with previous findings. In addition, an adult participant group will also be included, in order to directly contrast adult's task performance with that of children. As previously described in chapter 1, overimitation has been studied in adults, with some studies determining that adults copy even more precisely than young children. The inclusion of adults will also allow for investigation of potential differences in gaze patterns between adults and children, as eye tracking studies investigating copying fidelity have yet to be carried out using adult participants.

Puzzle boxes are one of the most commonly used tasks in studies examining copying behaviours, particularly imitation and overimitation, as they can be adapted for use with children (Lyons et al., 2007; McGuigan et al., 2007; Yu & Kushnir, 2011), adults (Flynn & Smith, 2012; McGuigan et al., 2011) and non-human animals (Horner & Whiten, 2005; Tennie, Call & Tomasello, 2006). Many puzzle box tasks involve a demonstrator acting on the box, using both relevant and irrelevant actions to retrieve some sort of reward, meaning that the task potentially has a clearer goal than the board task. Although puzzle box tasks tend to have a limited number of potential actions that can be performed on them, they can be designed so that these actions can be performed in a number of different sequences. This means that sequence order can be examined as an additional measure of imitation, alongside the performance of relevant and irrelevant actions. Although puzzle boxes are frequently used in social learning studies, little is known about what participants attend to during the demonstration of these tasks. Therefore, as previously stated, eye tracking may help uncover more about the underlying mechanisms of overimitation when using these types of stimuli, particularly when examining whether participants attend to action information or object movement whilst the task itself is being completed. Previous research has suggested that when participants do not have access to action information (for example, when it has been deliberately removed from a demonstration), they are likely to adopt an emulative strategy (Reindl et al., 2017). Therefore, it could be suggested that participants who do not attend to this action information during a demonstration will also be likely to be emulators.

In addition to a puzzle box task, a tangram puzzle task will also be used. Dickerson et al (2013) developed magnetic tangram puzzle boards in order to assess imitative and emulative behaviour in 18 month old infants and 3.5 year old children. One issue with many of the tasks used for this purpose in previous research was that they were designed for use only with young children (e.g. copying actions from a puppet) and so would not be appropriate for use with adults. However, tangram puzzles can be easily adapted to increase or decrease difficulty for a range of age groups and abilities, and are not too “childish” for use with older participants. An additional benefit of these puzzles is that, as demonstrated by Dickerson et al (2013), they can be used to measure high and low fidelity copying behaviours. In order to achieve this, the puzzle is initially completed by a demonstrator, who includes additional actions that are irrelevant to puzzle completion. The actions of the participant can then be observed to determine the copying strategy used.

2.1.1. Aims of the present study

There are three underlying aims of this research, as described below:

To induce overimitative, imitative or emulative behaviours in participants in order to identify potential distinct gaze patterns relating to each behaviour.

As previously described, some research has suggested potential differences in the way that participants who were observed to copy more actions from a demonstrator attended to the demonstration video in comparison to those who chose their own strategy for task completion. Findings by Taylor & Herbert (2014) determined that infants who copied a higher number of target actions were found to have attended significantly more to the task itself during a demonstration video, in comparison to those who did not copy, who attended more to irrelevant background information. Although these findings are limited, making it difficult to make extremely specific predictions, it is expected that participants who achieve the highest quotient scores will attend more to the task, in comparison to those who achieve lower quotient scores, and therefore display lower copying fidelity. Previous research suggests that those who copy less faithfully are less likely to attend to relevant action information (Reindl et al., 2017). In addition, it may also be expected that participants who copy more faithfully may attend more to the demonstrator themselves. Vivanti, Hocking, Fanning, & Dissanayake (2017) determined that typically developing children who looked more at the demonstrator's face were more likely to then copy causally irrelevant actions when completing a task.

The purpose of this initial experiment is therefore to explore the eye tracking patterns of participants across a full range of copying behaviours. As very little previous research has attempted to make these comparisons, we are unsure as to how likely they are to occur naturally in our participant sample, given that the underlying reasons for these behaviours are not fully understood, particularly in typically developing children. Therefore, in order to try to ensure that a range of copying behaviours are captured within the participant sample used, an attempt will be made to induce high and low fidelity copying behaviours, by providing specific instructions to participants before they view the demonstration video. It is hoped that this will create extremes of each of these behaviours in order to determine whether differences in gaze patterns can be observed. Conducting eye tracking during the demonstration itself will provide further insight into the point at which differences between these behaviours can be observed. Do emulators simply look at different things than overimitators or do all participants attend to similar areas of the screen during the demonstration, and then display differences in behaviour due to other factors (such as a conscious choice not to copy what they have seen)?

To determine whether developmental differences in task performance and gaze patterns are observed in 3-year-olds, 5-year-olds and adults.

In multiple previous studies, differences in copying behaviours have been observed between 3-year-olds, 5-year-olds and adults, with overimitation increasing with age (Flynn & Smith, 2012; McGuigan et al., 2011b, 2007b). In fact, a series of experiments conducted by Flynn & Smith (2012) determined that adults continued to overimitate even when time pressure was introduced, and when a monetary reward was offered to the participant who was able to complete a task in the quickest time. If these copying behaviours are pervasive, despite instruction, as observed by Flynn & Smith (2012) then it is expected that they should also be observed in participants within this study. When examining eye tracking data, participants' attention to particular areas of the screen will be investigated. It is expected that overall, adults will be able to attend to relevant aspects of the demonstration with more ease than younger participants, similar to findings by Taylor & Herbert (2014) which determined that infants tended to distribute their gaze more evenly across a demonstration than adults, who attended primarily to the task. However, if the goal to induce copying behaviour is achieved, then we should not see a rise in overimitation with age, and the differences in gaze patterns should be limited to differences between those who achieve high or low quotient scores, as previously discussed. It is additionally possible that differences in behaviour might be due to an interaction between both age differences and the instructions given.

To investigate differences in copying behaviours displayed by participants when completing a more functional task in comparison to a more arbitrary task.
The majority of studies into overimitative behaviour use only one task type, or multiple tasks with similar procedures for completion, which makes it difficult to assess shifts in copying strategies within the same participant group. In order to determine whether task performance is context-dependent, a functional, goal-driven puzzle box task will be used alongside a more arbitrary tangram puzzle task, which does have a clear end state, but this is achievable using a variety of steps that may be completed in any order. As well as the relevant actions required for task completion, irrelevant steps will also be included in the demonstration videos for both the box and tangram tasks, to allow for measurement of overimitation, alongside imitation of relevant actions and emulation (goal state only). If overimitation is primarily driven by social pressure and a desire to show affiliation, then adult participants should display similar levels of overimitation in both the box and tangram puzzle tasks.

2.2. Pilot testing and Task Development

2.2.1. Task Development: Demonstration Videos

There is some argument within the literature as to whether live or video demonstrations are most appropriate in social learning research. This is predominantly due to the fact that a video deficit effect has been identified, in which children appear to learn less successfully from a video demonstration in comparison to a live demonstration (Anderson, 2005; Hayne, Herbert, & Simcock, 2003). However, research has determined that this deficit disappears around the age of 2.5 to 3 (Dickerson et al., 2008), which indicates that it should not affect task performance in participants in this study. One benefit of video demonstrations in comparison with live demonstrations is improved consistency, as it is unlikely that a live demonstration would be identical for each participant. As previously discussed, subtle ostensive cues appear to affect performance on a task (Gergely, 2006), and these can be kept consistent across participants if video demonstrations are used. Additionally, video demonstrations allow for the inclusion of a number of manipulations, for examples, changes in instructions given to participants, whilst keeping all aspects of the demonstration itself consistent. Furthermore, video demonstrations are required in order for eye tracking to be conducted whilst participants watch the task being demonstrated. For these reasons, video demonstrations were used in this research. In order to reduce bias caused by demonstrator characteristics as much as possible, initially it was decided that the demonstrator themselves would not be visible during the completion of the puzzle; only their hand would be seen. However, following pilot testing it was determined that this was too artificial. Therefore the demonstrator was visible during all demonstration videos, and made eye contact with the camera at the end of the task. However, the demonstrator did not provide any spoken cues during the videos, in order to reduce the amount of language used as much as possible. This was particularly important for the youngest participants.

2.2.2. Task Development: Box Puzzle

The puzzle box developed for this experiment consisted of a clear Perspex box with three levers – two functional, and one non-functional. A plastic egg is placed on the top lever. The aim of the task is to manipulate the levers in order to cause the plastic egg to fall out of a chute at the bottom of the box. As the box is transparent, the movements of the levers both inside and outside of the box are visible to participants during the demonstration and whilst using the box.

The box task comprised of a clear Perspex box, measuring 20cm tall, 30cm wide and 16cm deep (see figure 2.1). The box contained levers on the left and right hand sides which could be operated in any order to release a reward. Both the internal and external mechanisms of the task were visible to participants throughout the demonstration, and whilst completing the task. On the left hand side, a wheel could be rotated either forwards or backwards to turn the platform containing the reward, in order to tip it onto the lever below. However, only turning the lever forwards would result in a successful outcome. On the lid of the box, a distracter lever could be moved up or down but this had no effect on the outcome of the task. On the right hand side of the box, a lever could be moved up or down, which moved the attached platform up and down. If the reward was on the lever at the time that it was tipped downwards, then it would fall through the chute at the bottom of the box. This chute was designed to be small enough that participants could not use it to retrieve the reward without using the levers. If participants did attempt to reach into the box through the chute, they were discouraged from doing so. The reward consisted of a small plastic egg which could be swapped with the experimenter for a sticker. The lid of the box was held closed using a strong magnet, and a key was required to release the lid to reset the task, which was not accessible by participants. This prevented participants from simply opening the lid of the box to retrieve the reward.

Figure 2.1: Box task used in demonstration and testing

The demonstration video for the box task can be viewed by clicking on the following link (password: demo). Both relevant and irrelevant actions were included as part of the demonstration. The demonstrator first tips the lever on the right hand side of the box to move the lower platform into a downwards position. She then knocks on the lid of the box – an unnecessary action for task completion. A second unnecessary action involves the demonstrator pulling the lever attached to the lid of the box. Finally, the demonstrator rotates the wheel on the left hand side of the box, tipping the platform that releases the reward, which falls down the chute. The demonstrator retrieves the reward, and then faces the camera (see figure 2.2). Although it may have been preferable to split the irrelevant actions up further, the fact that only two relevant actions were used, and that the left lever action had to be operated last, made it difficult to do so.

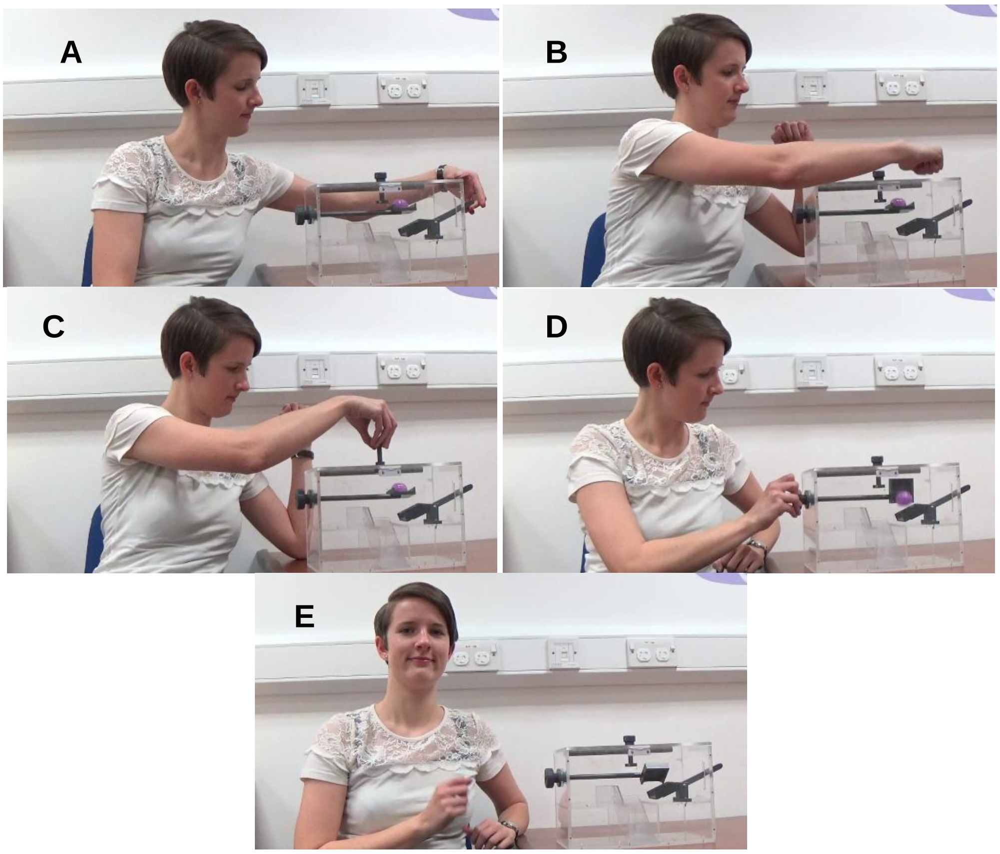


Figure 2.2: All actions involved in the box demonstration video. A: The demonstrator tips the right lever. B: She knocks on the lid of the box. C: She lifts the irrelevant top lever and lets it go so that it falls back into its initial position. D: She turns the left lever to release the ball, which falls onto the right lever and down the chute. E: The demonstrator picks up the ball, faces the camera and smiles.

Participants viewed each demonstration video three times, in the sequence displayed in Figure 2.3. A yellow smiley face indicated the end of the demonstrations.

Figure 2.3: Trial sequence shown to participants.

Pilot Testing

A control group was used to determine whether the box task was suitable for use with participants. Participants were also observed to determine whether they spontaneously produced any of the irrelevant actions that were to be included in the demonstration video.

18 participants (8 female, 10 male), aged 3 (N=2), 4 (N=7), and 5 (N=9), mean age 4.39, took part. Participants did not view any video demonstration prior to completing the task, but were asked if they would like to try a puzzle. Participants were then presented with the box puzzle, and were given time to explore it and complete it, although they were given no further instructions. Actions matching those in the demonstration video were recorded, and no time limit was given.

All participants were able to solve the puzzle. Of the 18 participants, 12 (67%) spontaneously used the irrelevant lever when completing the task. No participants used the knock action, and therefore none were able to complete all of the actions in order. No changes were made to the box task or actions used before the remainder of testing was completed.

2.2.3. Task Development: Board Puzzle

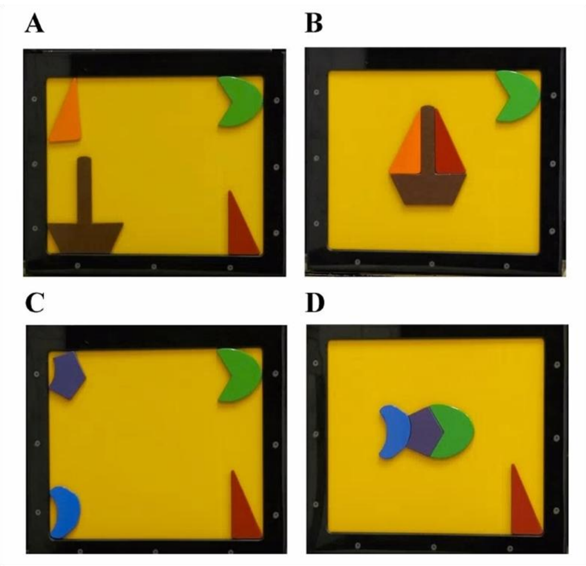


Figure 2.4: Magnetic board tangram puzzle used by Dickerson et al (2008)

The puzzle was adapted from a magnetic tangram puzzle, initially used by (Dickerson et al., 2008), as shown in figure 2.4. It consisted of a painted steel sheet inside a 30x40cm plastic frame, producing a magnetic board which could be placed into a stand to be displayed upright. Coloured perspex magnets of varying sizes were produced to allow for creation of a simple fish shaped puzzle (see figure 2.5).

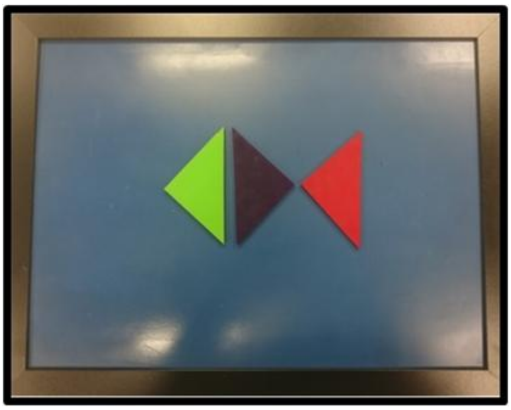


Figure 2.5: Initial "fish" magnetic tangram puzzle used in pilot testing

2.2.4. Pilot Testing: Board Puzzle

Initial pilot testing of the board task took place at local museums, with the aim of this stage of testing being to determine whether the difficulty level of the puzzle was appropriate for both 3 and 5 year old participants. 25 children, aged 3 (N=9), 4 (N=9) and 5 (N=7), mean age 3.92, took part. The puzzle (see figure 2.5) was placed vertically on a table, with the three triangular puzzle pieces, designed to form a fish shape placed on the table in front of it, ensuring that the child could easily reach both the board and the pieces. Prior to taking part, parents or guardians of the children provided written consent. The child was then shown a photograph of the completed fish puzzle, and was asked what shape they could see in the picture. The participant was then asked to use the pieces in front of them to produce the same shape as the one in the photograph, which was visible to the child throughout. If participants failed to produce the correct colour order or shape, they were prompted with the question “does it match the picture?” Completion success and time taken was recorded.

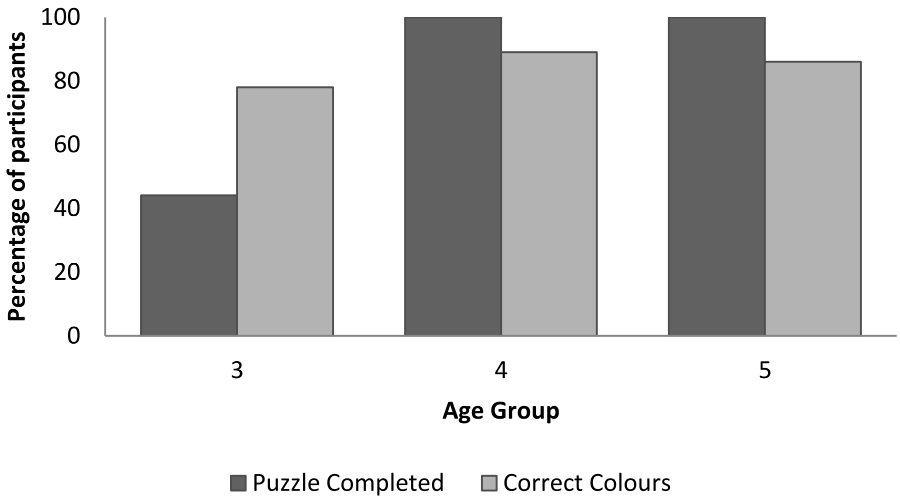


Figure 2.6: Percentage of three, four and five year old participants that successfully completed the puzzle and used the correct coloured pieces during pilot testing.

Figure 2.6 indicates results obtained from pilot testing. Although all 4 and 5 year old participants were capable of copying the puzzle shape from a photograph, the majority of 3 year olds were not, potentially indicating that the task remains too difficult. One 4 year old and one 5 year old failed to produce the same colour order of the puzzle pieces, despite producing the correct shape overall.

As indicated by the low completion rates found in the pilot data in figure 6, 3 year olds appear to find the puzzle task difficult, with more than half of 3 year old participants failing to achieve the correct puzzle shape. This may have occurred for a number of reasons. Testing took place on a table at the edge of a soft play area where other children were still playing. The environment was quite noisy, and other children continued to play whilst the participant was being tested. Since working memory and inhibitory control is still developing in 3 year olds (Wiebe et al, 2011), it could be the case that younger participants found it more difficult to “switch off” from the distractions around them and focus on the task. To ameliorate this, future participants were tested in a less distracting environment.

In each age group, at least one participant failed to use the correct colour order when completing the puzzle. Participants in this stage of testing were not tested for colour blindness, so it is possible that this could have been an issue in these cases. It may also be the case that children were copying the puzzle using higher level hierarchical information (the overall shape) rather than copying specific details (correct colour order). Flynn & Whiten (2008) suggested that young children were more likely to copy hierarchical information as opposed to specific action details when completing a task, so this may also have affected performance.

In sorting tasks, 3 year old children experience difficulties when switching from sorting by colour to sorting by shape since their understanding of higher order rules is limited (Perner & Lang, 2002). It may be the case that participants do not understand consider both shape and colour order to be important. Additionally, using clearer instructions prior to the task may result in more children being able to complete the puzzle. Finally, it may simply be possible that the puzzle is too difficult for younger participants. For these reasons, the puzzle was refined to make the pieces more distinct from one another as in the Dickerson et al (2008) version (see Figure 2.7).

Figure 2.7: Final magnetic "fish" tangram puzzle design

Task

The board consisted of a wooden frame and a painted blue steel sheet and measured 30cm tall and 40cm wide. The fish puzzle was made up of three coloured pieces, as seen in Figure 2.7. Each piece had a magnetic backing, which allowed pieces to be affixed to the board and easily moved or removed by participants.

Demonstration Video

As well as placing the three pieces onto the board to form the fish shape, three additional unnecessary actions were present in the demonstration video for this task, and were identified as “swirl”, “slide” and “wave”. In the swirl action, the demonstrator picked up the orange puzzle piece from the table and moved it in a circular clockwise motion before placing it onto the board. The demonstrator then placed the green puzzle piece normally. In the wave action, the demonstrator waved her hand over the puzzle board from left to right. In the slide action, the demonstrator picked up the purple puzzle piece and placed it at the right hand edge of the board, before sliding it into the correct position next to the green piece. Finally, the demonstrator faced the camera at the end of the task (figure 2.8).

Figure 2.8: Actions performed by the demonstrator during the board demonstration video. A: The demonstrator picks up the orange puzzle piece and moves it in a clockwise circular motion above the board, before placing it (swirl action). B: The demonstrator moves her hand in a wave-like motion across the board (wave action). C: The demonstrator places the purple tail piece at the right hand edge of the board, and slides it into place (slide action). D: The demonstrator faces the camera and smiles.

Participants viewed each demonstration video three times, in the sequence displayed in Figure 2.9. A yellow smiley face indicated the end of the demonstrations.

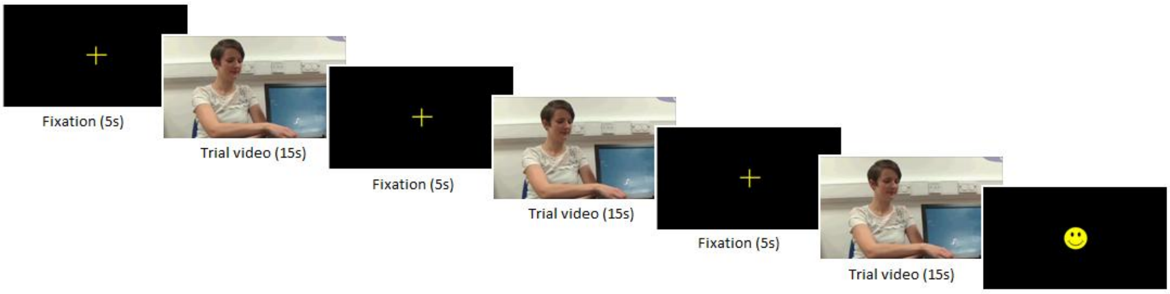


Figure 2.9: Board trial sequence shown to participants

2.3. Method

2.3.1. Participants

98 typically developing children (51 male, 47 female), aged from 3 to 5 years (mean age 54 months) were recruited from local primary schools to participate in this study. Additionally, 49 adults (aged 18-21) were recruited from the undergraduate Psychology programme at Newcastle University to take part. Adults received course credit in return for their participation. Prior to completion of the research, written consent was provided by parents and verbal consent was given by the participants themselves in the case of children. Adult participants also provided written consent. Eye tracking data was collected from 44 children in total; the remaining 54 were excluded due to failed calibration. Two additional children were excluded from the analysis due to special educational needs (SEN) status. Eye tracking data was collected for 36 participants, with the remaining 13 participants also being excluded due to failed calibration.

2.3.2. Apparatus

An Eye Tribe eye tracker was used to collect gaze data. This is a small, portable desktop tracker that allows for binocular gaze data to be recorded at a sampling rate of 60Hz. Two tasks were used within this study; a board task and a box task. These are described below. The tasks were designed to be appropriate for use by both child and adult participants. The same female demonstrator appeared in both the board and box demonstration videos, and both videos were filmed in the same location.

2.3.3. Design

Children were randomly assigned to one of three conditions: copying, goal-focused and video control. Each participant completed both the board and box tasks. Task order was counterbalanced across participants, although participants always viewed all three video trials of one task before moving onto the next. Since no changes were made to the task following pilot testing, data from pilot testing was determined to be appropriate for use as control data. Note here that there is a difference between video control participants, who viewed the demonstration video but did not receive instructions on how to perform the task, and “true” control participants, who did not view a demonstration video. “True” controls were included to ensure that participants did not spontaneously produce the irrelevant actions, which would have suggested that they were not truly irrelevant, and were instead perhaps a natural response to the task.

2.3.4. Procedure

In schools, testing took place either in a quiet area of the child's classroom, or in a separate room within the school. Children were tested one at a time and were encouraged not to tell their classmates how to solve the puzzles, which were presented as games. Additional testing of both adults and children also took place in the child development lab at Newcastle University. Parents completed a written consent form before children were allowed to participate, and additional verbal consent from the participant was sought before beginning each task. Participants were free to withdraw at any time. In all conditions other than the baseline condition, participants were seated at a table in front of a laptop computer with a 15" monitor. An Eye Tribe desktop eye tracker was placed on the table in front of the laptop. The tracker was placed below the monitor so that participants could clearly see the entire screen. Each participant completed the standard Eye Tribe 9-point calibration, and participants were instructed to keep as still as possible whilst they viewed the videos. Demonstration videos were played using open-source OGAMA software (Voßkühler, 2008), which also simultaneously records eye tracking data, including fixations and saccades. For each task, participants viewed the appropriate demonstration video followed by a 5 second fixation cross three times. Before viewing the demonstrations, participants were provided with specific instructions depending on the condition they had been assigned to, as discussed below. In all eye tracking conditions, participants were prompted ("you're doing really well, keep going" or "try to keep looking at the screen") if necessary.

Video Control Group

Participants in the video control group were shown the demonstration videos and given the following instruction: "I'm going to show you a video of someone doing a puzzle. You will see the video three times, and afterwards it'll be your turn to have a go at the puzzle".

Copying Group

Participants in the copying group were shown the demonstration videos and given the following instruction: "I'm going to show you a video of someone doing a puzzle. You will see the video three times, and afterwards I'd like you to do the puzzle exactly the same as the person on the video”.

Goal-focused Group

Participants in the goal-focused group were shown the demonstration videos and given the following instruction: “I’m going to show you a video of someone doing a puzzle. You will see the video three times, and afterwards I’d like you to do the puzzle as quickly as you can”.

Following the video demonstration, participants were immediately provided with the puzzle that corresponded to the demonstration video that they had just viewed. The instruction that had been given prior to the video demonstration was repeated (e.g. “Now I’d like you to do the puzzle as quickly as you can”). Participants were given as long as necessary to complete the puzzle, but children were prompted with “well done, keep going” if they strayed off task. Following completion of both tasks, children were praised and given a sticker, and the number of relevant and irrelevant actions performed during task completion were recorded by the experimenter.

2.3.5. Data Analysis

Participants who failed the calibration procedure were excluded from analysis of eye tracking data. The demonstration videos were split into individual frames, at a rate of 25 frames per second, and fixation data was extracted from OGAMA. A Matlab script was designed to display video frames one by one, with any corresponding fixations overlaid onto them. As each fixation had a corresponding start and end time, the duration of fixations in each relevant area of interest was calculated per trial for each task. Four areas of interest were identified: the demonstrator’s face, the demonstrator’s hands, the levers, and the box itself. Examples of fixations coded at each of these areas of interest can be seen in Appendix A. A fixation was coded as belonging to the levers AOI if it appeared on any part of the lever, either inside or outside the box. Fixations coded as belonging to the box were those focused on the “blank space” inside the box, rather than any of its internal mechanisms.

Participant data was scored depending on the number of irrelevant actions produced and the order in which the relevant actions were copied. Within this task, there were two possible irrelevant actions and two potential measures of the order of actions completed: the order in which participants completed the relevant actions (left lever, right lever) only, and the order in which they completed all actions. Participants were awarded a score of 1 for each of these elements that they completed, meaning that the maximum possible score for this task was 4. A quotient score was then calculated for each participant, by taking their actual score on the task and dividing it by the maximum possible score. A quotient score closer to 1 would indicate higher fidelity copying (overimitation), whereas a quotient score closer to 0 would indicate lower fidelity copying (emulation). It is acknowledged that other methods for task scoring might also be appropriate for directly defining an individual as a high or low fidelity copier (e.g. grouping participants based on whether they performed specific aspects of a task or not – see Carr, Kendal, & Flynn (2015)). However, the use of a quotient score in all tasks across all experiments within this thesis allows for direct comparisons to be made between each of these tasks, which is one of the overall aims of this thesis. In order to fully investigate participants' behaviour, performance of individual aspects of the task, such as the irrelevant actions used, was also examined in addition to quotient score, in order to gain a better understanding of more nuanced behaviour during task completion.

2.4. Results

2.4.1. Box Task: Behavioural Data

Children

Table 2.1 indicates the mean quotient score for child participants in each condition. As expected, those in the copying group achieved the highest mean quotient score.

Table 2.1: Mean box quotient scores for child participants in each of the assigned conditions

Condition	N	Mean	Std. Deviation
Copying	26	0.67	0.337
Goal Focused	26	0.56	0.303
Video Control	28	0.65	0.307
Control	18	0.17	0.121
Total	98	0.54	0.341

To further understand the variability in task performance between participants in each of the assigned conditions, the percentage of participants achieving each of the possible quotient scores in each category was examined. These results can be seen in figure 2.10. Using an ANOVA to compare the mean quotient scores between each of the assigned conditions, a significant difference was observed between groups: $F(3,97)=13.15$, $p<.001$. Bonferroni post-hoc tests indicated a significant difference in quotient score between those in the control group (M=.017) and those in the copying group (M=0.67, $p<.001$), the goal focused group (M=0.56, $p<.001$) and those in the video control group (M=0.65, $p<.001$). The purpose of including the control group in this initial analysis is to indicate that control participants did not spontaneously perform the actions in the demonstration video, nor did they spontaneously use the same order as the demonstrator, as their quotient scores were found to be very low. Most of the control participants who scored .25 did so because they spontaneously operated the irrelevant lever on the lid of the box. However, there is no indication that this lever is not functional until it is operated, so this spontaneous use of it by the majority of participants was perhaps to be expected. Additionally, control participants may also have achieved a score of .25 if they unknowingly operated the right and left levers of the box in the same order as used in the demonstration video.

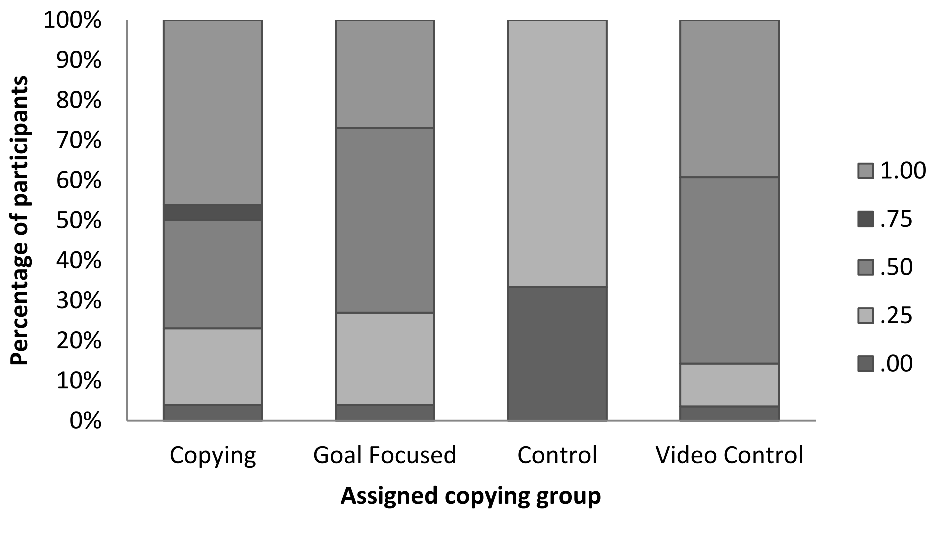


Figure 2.10: Percentage of child participants in each assigned condition that achieved each of the possible quotient scores (0 (N=9), .25 (N=26), .50 (N=32), .75 (N=1), 1.00 (N=30))

Comparisons were then made between the three experimental groups (copying, goal focused, video control), excluding the control group. No significant difference in mean quotient score was observed between groups: $F(2,79)=.989$, $p=.377$, although participants in the goal focused group did have a slightly lower score than those in the copying and video control groups. The similarity in scores between participants in the copying (M=.67) and video control (M=.65) groups may indicate that children of this age group are naturally more inclined to copy from a demonstrator, since participants in the video control condition were given no instructions on how to complete the task.

To further understand any potential differences between groups, the order in which participants completed the relevant actions was examined (figure 2.11). Using a chi-squared test, no significant difference was observed between groups: $x^2(2)=3.09$, $p=.214$. This is to be expected, given the results indicated in figure 15, which show that the majority of participants in each of the three assigned conditions completed the relevant actions in the same order as the demonstrator.

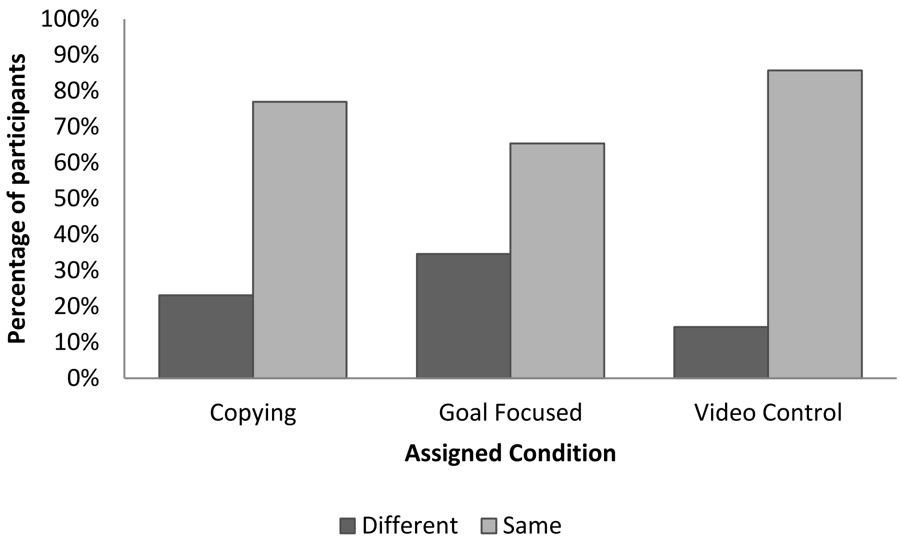


Figure 2.11: Percentage of child participants in each of the assigned conditions that copied the order of relevant actions from the demonstrator

Performance of irrelevant actions by participants in each of the assigned conditions was then investigated. The percentage of participants that performed the knock action is displayed in figure 2.12. Equal numbers of participants in the copying and video control groups knocked on the lid of the box. However, in the goal focused group, fewer than half of all participants used the action, suggesting that the instruction provided to these participants did affect their behaviour slightly. As would be expected from this pattern of results, there was no significant difference observed between groups: $x^2(2)=.416$, $p=.812$

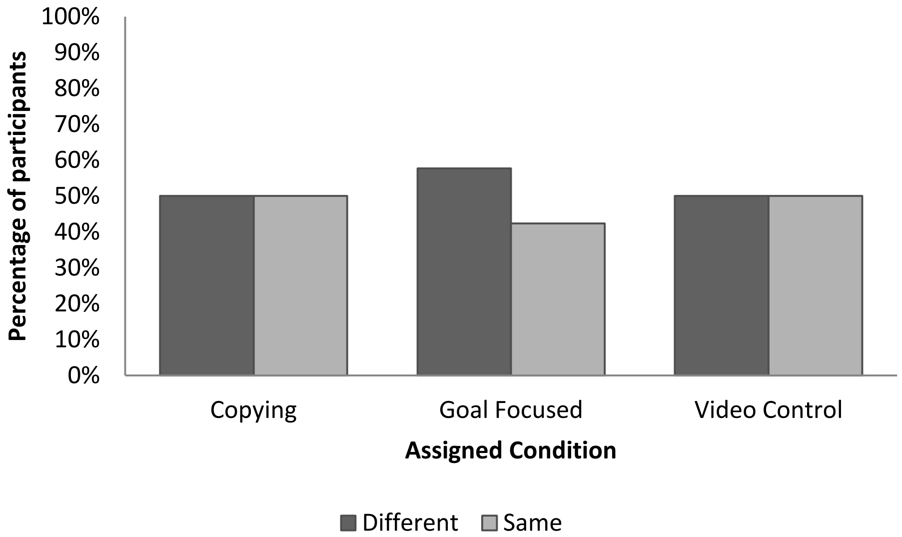


Figure 2.12: Percentage of child participants that performed the irrelevant knock action.

The percentage of participants that used the irrelevant lever is displayed in figure 2.13.

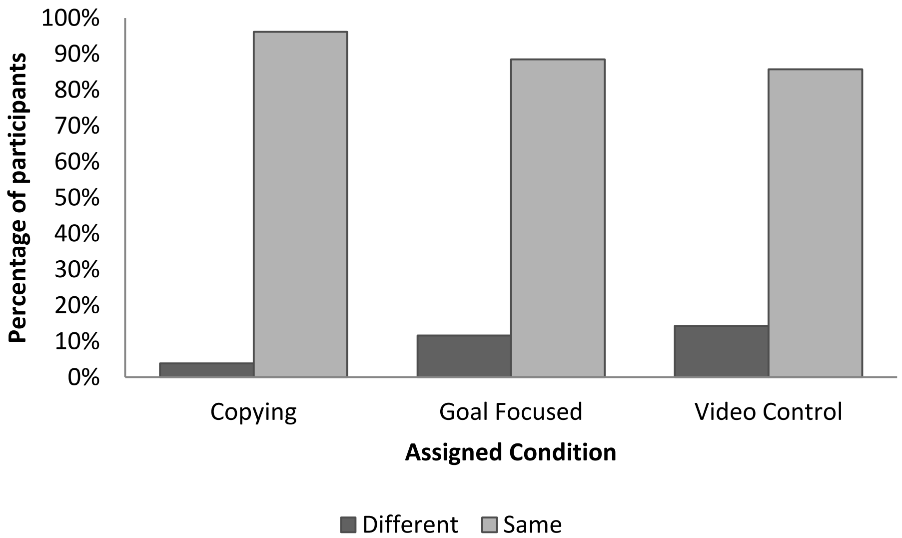


Figure 2.13: Percentage of child participants that completed the irrelevant lever action

Unlike the knock action, the majority of participants in all groups imitated the irrelevant lever action. Again, as would be expected from the results displayed in figure 17, no significant difference was observed between groups: $x^2(2)=1.73$, $p=.420$. These results potentially suggest that children interpreted the lever action to be functional, and therefore necessary for task completion.

Adults

Table 2.2 indicates the mean quotient score for adult participants in each condition. Adults in the copying condition achieved the highest mean quotient score, as expected. However, adults in the video control condition achieved a lower score than those in the goal focused condition, suggesting that they copied less precisely than participants who were encouraged to complete the task as quickly as possible. This suggests that the instruction given to participants in the goal focused condition did alter their behaviour, but perhaps not in the way that was intended.

Table 2.2: Mean box quotient scores for adult participants in each of the assigned conditions

Condition	N	Mean	Std. Deviation
Copying	16	0.891	0.241
Goal Focused	16	0.641	0.302
Video Control	14	0.589	0.348
Total	46	0.712	0.320

A significant difference between groups was observed using an ANOVA, $F(2,45)=4.52$, $p=.017$. Bonferroni post-hoc tests indicated a significant difference in mean quotient score between participants in the copying ($M=.891$) and those in the video control group ($M=.589$, $p=.025$). The difference between participants in the copying and goal focused ($M=.641$) groups was also approaching significance ($p=.067$).

To further understand the variability in task performance between participants in each of the assigned conditions, the percentage of participants achieving each of the possible quotient scores in each category was examined. These results can be seen in figure 2.14. The majority of participants in the copying group achieved the maximum possible score of 1, indicating high fidelity copying. However, performance in the goal focused and video control conditions was more variable.

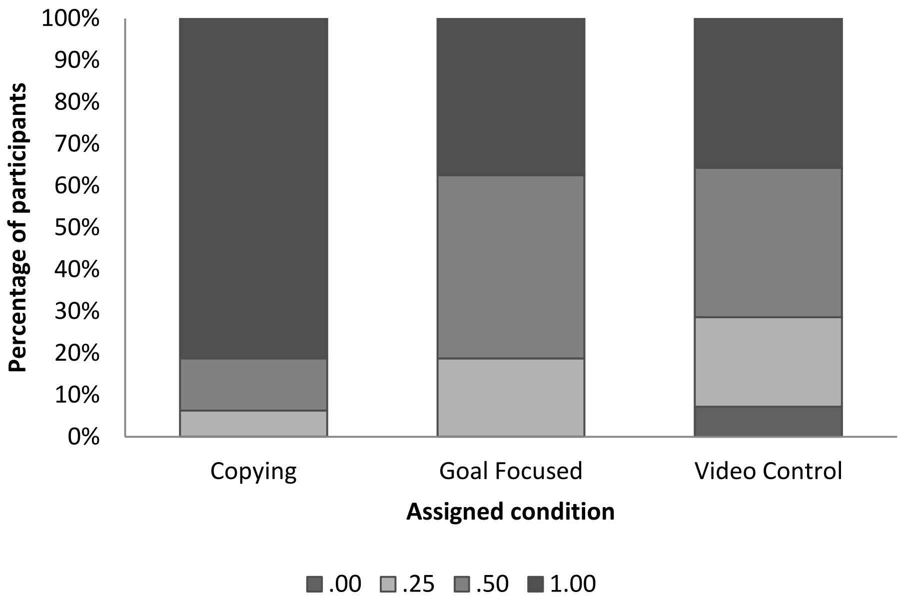


Figure 2.14: Percentage of adult participants in each of the assigned conditions that achieved each of the possible quotient scores (0 (N=1), .25 (N=7), .50 (N=14), 1.00 (N=23)).

The order in which participants completed the relevant actions (right lever, left lever) was examined. Results are shown in figure 2.15. The majority of participants in all groups performed the relevant actions in the same order as the demonstrator. No significant difference between conditions was observed, using a chi-squared test, $x^2(2)=1.13$, $p=.569$.

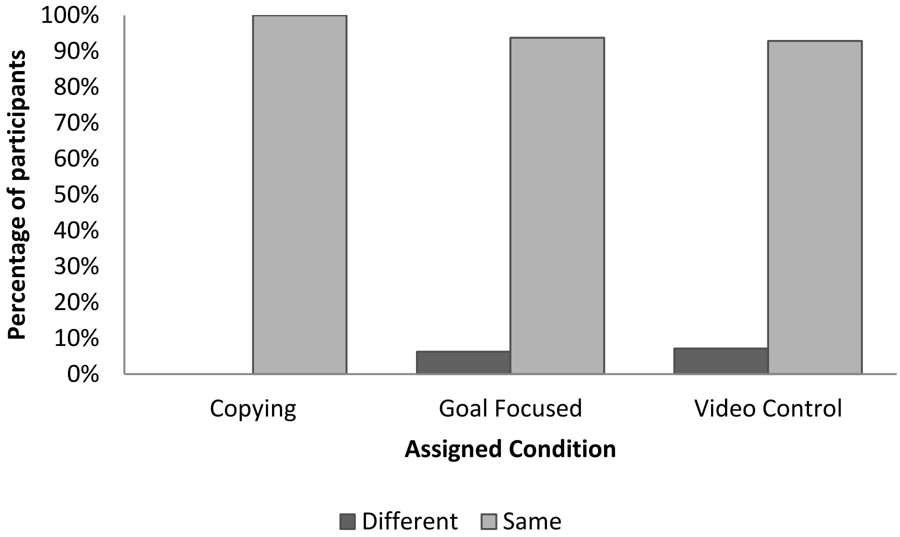


Figure 2.15: Percentage of adult participants in each of the assigned conditions that copied the order of relevant actions from the demonstrator

Performance of irrelevant actions by participants in each of the assigned conditions was then investigated. The percentage of participants that performed the knock action is displayed in figure 2.16. The majority of participants in the copying group did perform the knock action, whereas the majority of those in the video control and goal focused conditions did not. This suggests that the instructions provided to the copying group were successful. Using a chi-squared test, a significant difference was observed between groups: $x^2(2)=7.35$, $p=.025$.

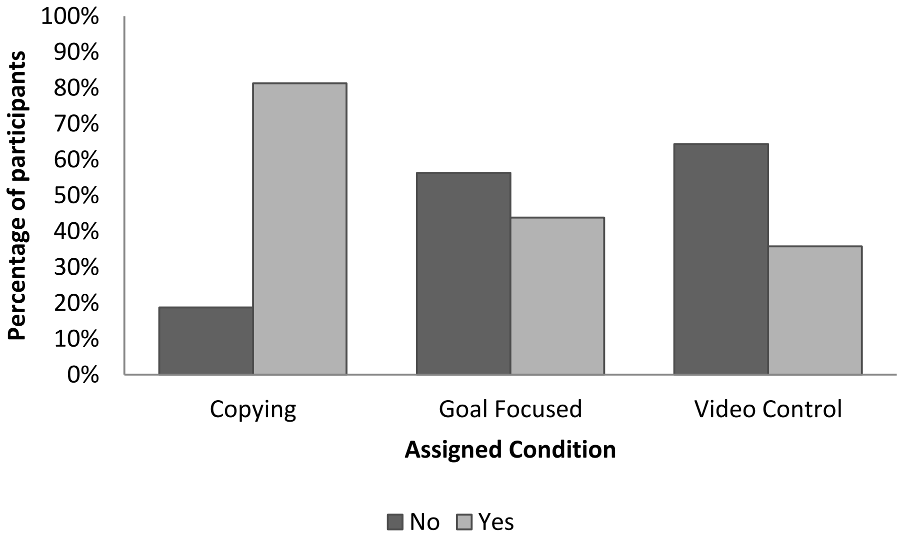


Figure 2.16: Percentage of adult participants in each of the assigned conditions that imitated the irrelevant knock action

The percentage of participants in each of the assigned conditions that performed the irrelevant lever action is indicated in figure 2.17. Unlike the knock action, the majority of participants in all conditions used the irrelevant lever when completing the task. Again, this may suggest that the lever actually appeared to be functional. No significant difference in the performance of the lever action was observed between conditions, using a chi-squared test: $x^2(2)=2.62$, $p=.270$.

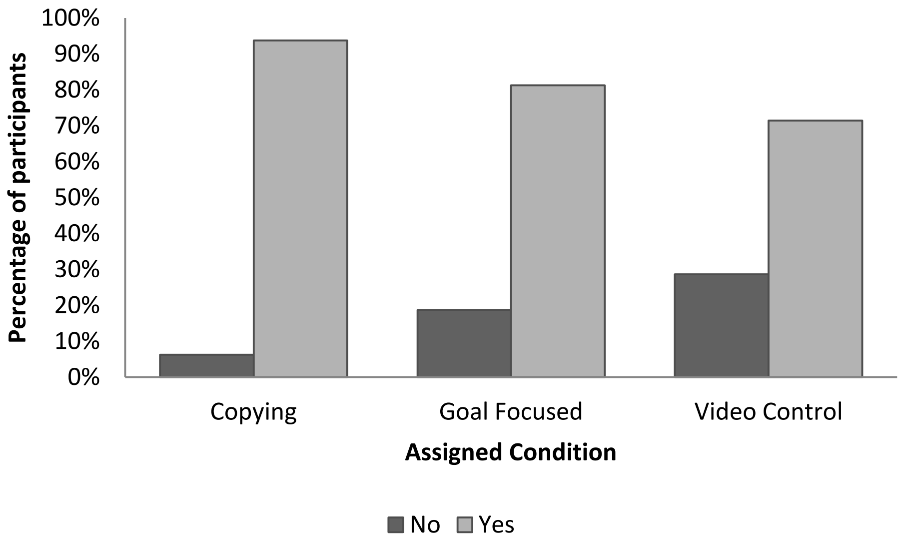


Figure 2.17: Percentage of adult participants in each of the assigned conditions that imitated the irrelevant lever action

Age Differences

Due to issues with recruitment, direct comparisons between 3 and 5 year old participants were not possible due to low participant numbers. However, behavioural differences between the ten youngest (mean age 45.8 months) and ten oldest child participants (mean age 60.4 months) were also examined to determine if there was a difference in copying fidelity, as previous literature suggests that a shift in copying behaviours should be observed as children age. The mean quotient scores for the ten youngest and oldest participants can be seen in table 2.3.

Table 2.3: Mean quotient scores on the box task for the ten youngest and ten oldest child participants

Age Group	N	Mean	Std. Deviation

Youngest	10	0.5250	0.18447
Oldest	10	0.5000	0.37268
Total	20	0.5125	0.28648

The ten youngest participants were found to have a higher mean quotient score (M=.525) than the ten oldest (M=.500), but this difference was not found to be significant: $F(1,19)=.036$, $p=.851$.

Comparisons were then made between child and adult participants, using an ANCOVA with assigned category as a covariate. There was no significant difference observed for participant age (adult or child): $F(1,124)=.147$, $p=.226$. However, the difference between participants' assigned categories was approaching significance, as would be expected from the individual age group results: $F(1,124)=3.55$, $p=.062$. The mean quotient scores for adults and children are displayed in table 2.4.

Table 2.04: Mean box quotient scores for adult and child participants (all conditions)

Age Group	N	Mean	Std. Deviation
Adults	45	0.71	0.321
Children	80	0.63	0.316
Total	125	0.66	0.319

2.4.2. Box Task: Eye Tracking Data

Eye tracking data was collected for 35 children and 31 adults. Heatmaps using the mean fixation positions for each group were calculated in order to make an initial assessment of any potential gaze patterns (see figure 2.18). The heatmaps appear to suggest that adult participants in all groups were better able to direct their attention towards the task than child participants.

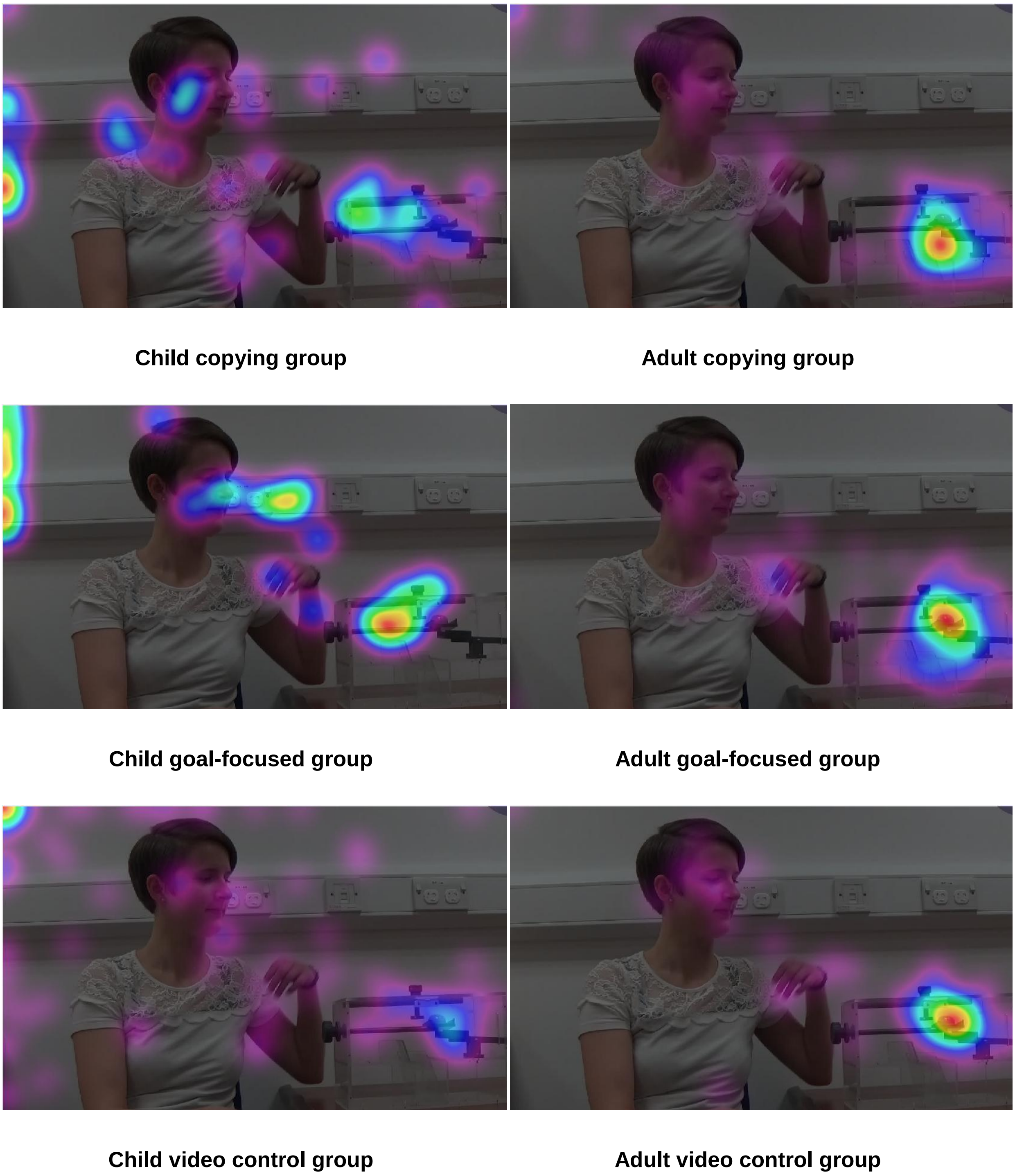


Figure 2.18: Heatmaps of mean fixation locations in each group

In order to investigate these differences further, the proportion of time spent looking at areas of interest (demonstrator's face, demonstrator's hands, the levers attached to the box, and the box itself) was calculated. Examples of fixations in each of these areas of interest can be found in Appendix A. Looking time at the background of the demonstration video was analysed separately, in order to measure participants' ability to attend to relevant stimuli. This is discussed further in the age comparisons section below. A significant difference was found between trials in terms of mean proportion of looking time at the box AOI in child participants: F(2,91)=4.588. p=.013, and therefore the mean proportion of looking time for each area of interest was averaged across the three trials for the purposes of analysis. Although no significant differences were found between trials for adult participants for any of the areas of interest, averaged trial data was also used in order to allow for appropriate comparison with the data obtained from child participants. With regards to both child and adult eye tracking data across trials, the time spent looking at relevant areas of the demonstration dropped across trials as would be expected.

2.4.3. Overall looking time at areas of interest

Children

Figure 2.19 indicates the mean time spent looking at each of the AOIs for each condition.

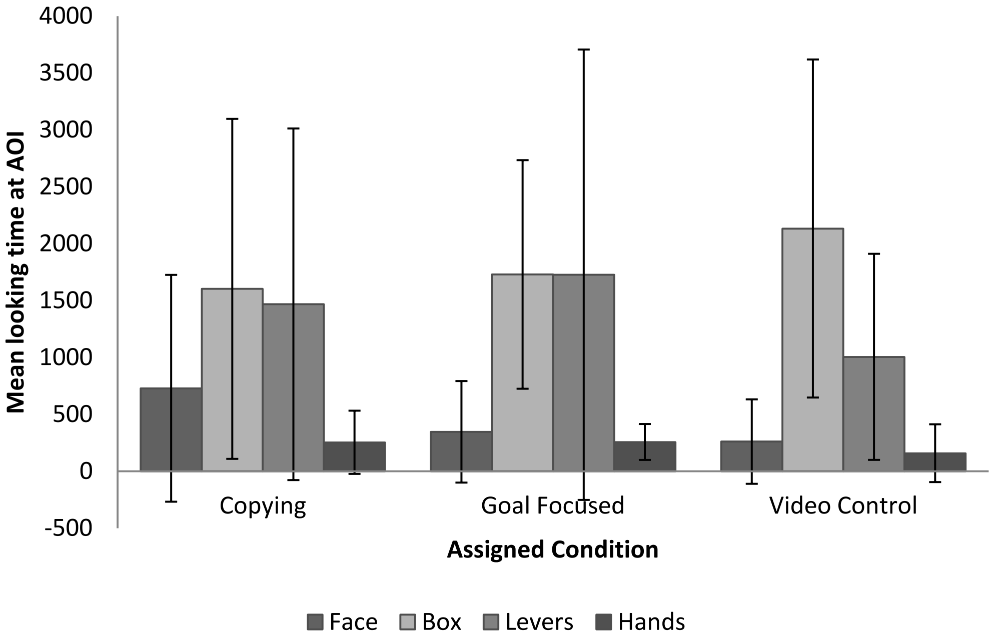


Figure 2.19: Mean looking time at each of the areas of interest (face, box, levers, hands) for child participants in each of the assigned conditions (copying, goal focused, video control). Error bars represent standard deviation.

There were no significant differences observed between participants in each of the assigned conditions at the face AOI: F(2,34)=.700, p=.504, the box AOI:

F(2,34)=.222, p=.802, the levers AOI: F(2,34)=.994, p=.381, or the hands AOI: F(2,34)=1.71, p=.197. However, participants in the video control condition spent the majority of their time looking at the box, rather than at its internal mechanisms or the movements used to operate it. This potentially suggests that participants in the copying and goal focused groups tended to focus their attention on more relevant areas of the video (e.g. the levers) in comparison to those who were provided with no specific instructions.

The difference in mean looking time at each of the areas of interest was then investigated for child participants achieving each of the possible quotient scores. Participants who achieved a score of 0 did not copy any aspects of the demonstration as scored, i.e. they did not copy any of the irrelevant actions, the order of the relevant actions, or the order of all actions. Instead, they used their own method (completing the relevant actions in the opposite order) to complete the task. These results can be seen in figure 2.20.

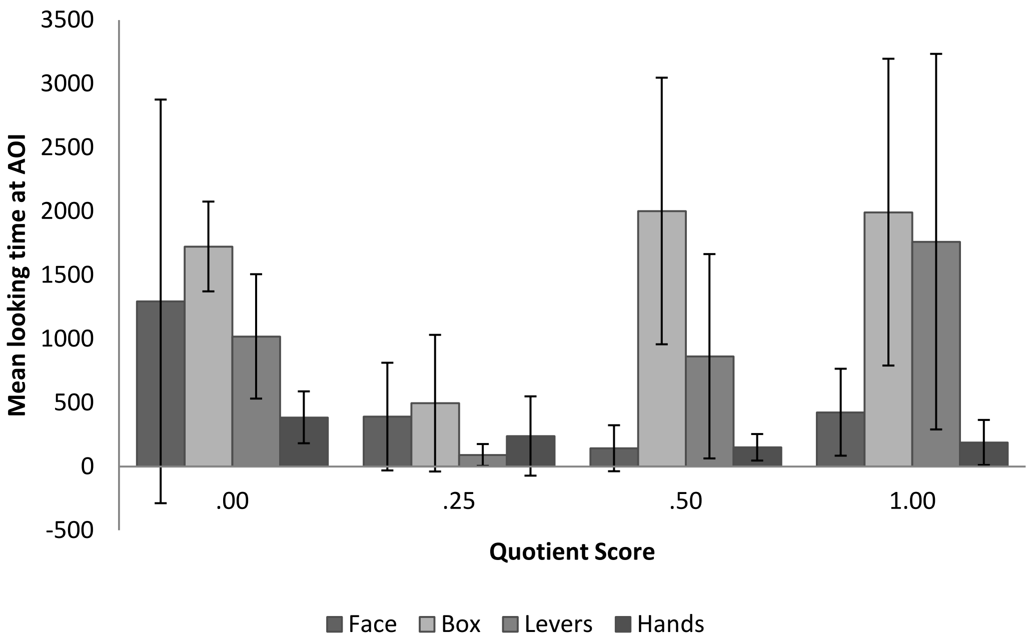


Figure 2.20: Mean looking time at each of the areas of interest (face, box, levers, hands) for child participants that achieved each of the potential quotient scores (0, .25, .50, 1.00). Error bars represent standard deviation.

An ANCOVA was used for comparison, using participants' assigned condition as a covariate. A significant difference in looking time at the face AOI was observed based on quotient score: F(3,34)=.4.38, p=.011. Bonferroni post-hoc tests indicated a significant difference between participants who scored 0 (M=1293.17) and those who scored .50 (M=142.05, p=.007), as well as those who achieved a score of 1 (M=424.39, p=.050). However, there was no significant effect of assigned condition: F(1,34)=.208, p=.652. These results suggest that participants who scored 0 spent much more time attending to the demonstrator's face than participants who achieved higher scores. It may simply be the case that since these participants were looking at the demonstrator, they missed other, more relevant aspects of the demonstration and so were not able to imitate them.

No significant effect of assigned condition: F(1,34)=.150, p=.702, or of quotient score: F(3,34)=2.22, p=.107 was found for the box AOI. No significant effect of assigned condition: F(1,34)=.095, p=.760 or of quotient score: F(3,34)=2.58, p=.072 was found for the levers AOI. Finally, no significant effect of assigned condition: F(1,34)=1.54 or of quotient score: F(3,34)=1.05, p=.384 was found for the hands AOI.

Adults

Figure 2.21 indicates the mean time spent looking at each of the AOIs for each assigned condition.

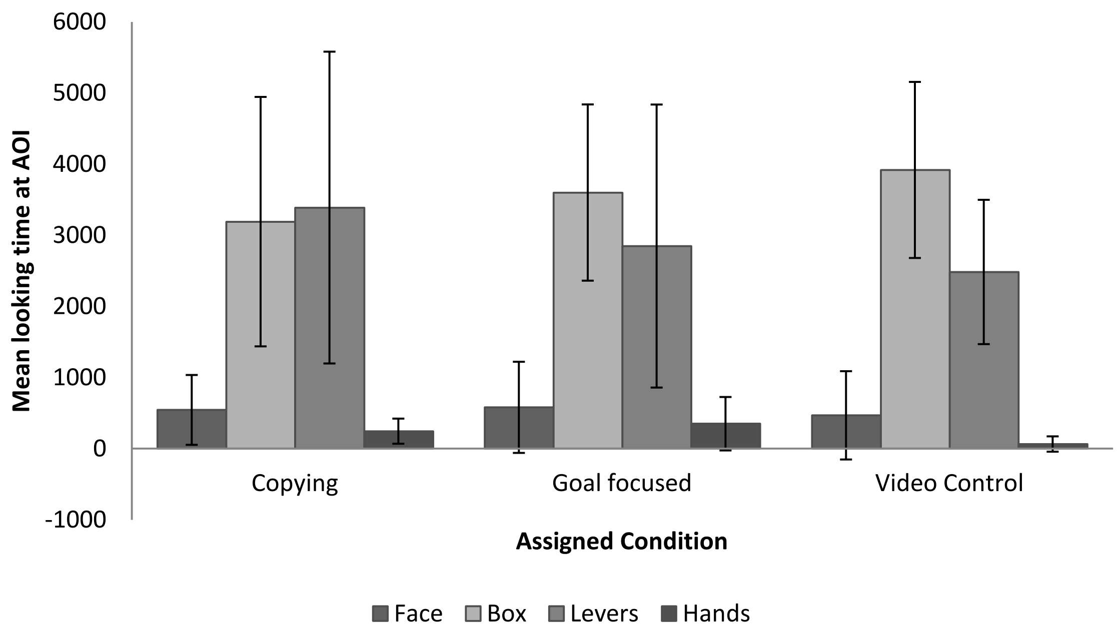


Figure 2.21: Mean looking time at each of the areas of interest (face, box, levers, hands) for adult participants in each of the assigned conditions. Error bars represent standard deviation.

The difference in mean looking time between participants in each of the assigned conditions for the hands AOI was approaching significance: F(2,30)=3.13, p=.059, with participants in the video control condition looking less at the hands than participants in the copying and goal focused conditions. No significant difference between assigned conditions was found for the box AOI: F(2,30)=.639, p=.536, the levers AOI: F(2,30)=.611, p=.550, or the face AOI: F(2,30)=.093, p=.911.

Figure 2.22 indicates the mean time spent looking at each of the AOIs for each quotient score.

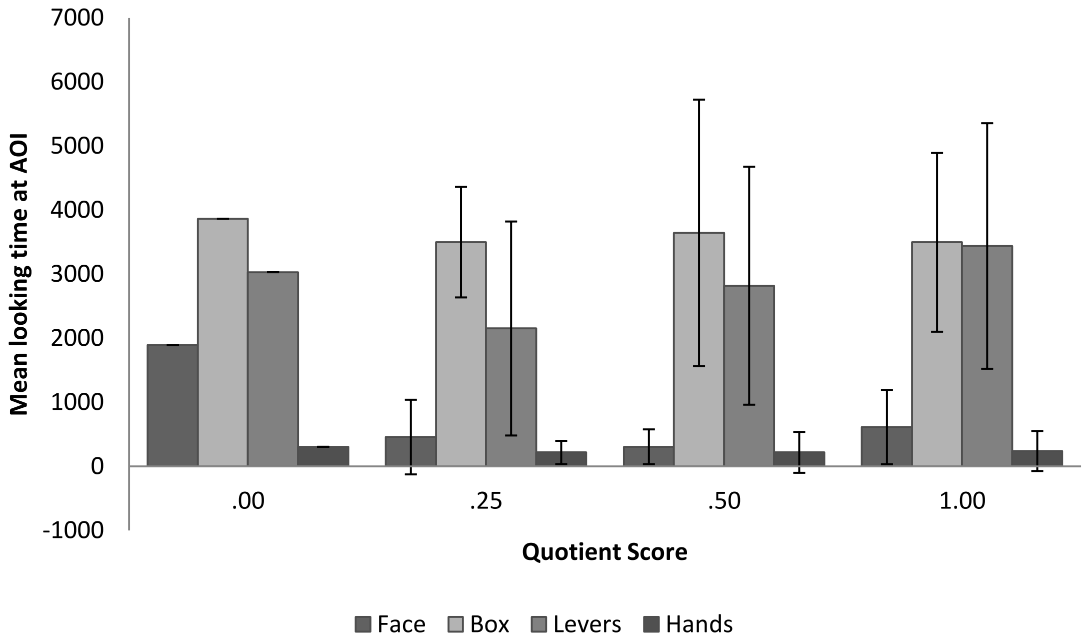


Figure 2.22: Mean looking time at each of the areas of interest (face, box, levers, hands) for adult participants that achieved each of the possible quotient scores (0, .25, .50, 1.00). Error bars represent standard deviation.

For the face AOI, a significant effect of quotient score was found: F(3,30)=3.01, p=.048. Bonferroni post-hoc tests indicated a significant difference between participants who scored 0 (M=1891) and those who scored .50 (M=304.75, p=.043). However, these results should be interpreted with caution as only one adult participant scored 0. No significant effect of assigned condition was found for this AOI: F(1,30)=.397, p=.534. Since only one adult participant scored 0, the ANOVA was run a second time, but excluding this participant. No significant effect of assigned condition: F(1,30)=1.16, p=.292, or of quotient score: F(3,30)=.026, p=.994 was found for the box AOI. No significant effect of assigned condition: F(1,30).562, p=.460 or of quotient score: F(3,30)=.620, p=.609 was found for the levers AOI. Finally, no significant effect of assigned condition: F(1,30)=2.06, p=.164 or of quotient score: F(3,30)=.142, p=.932 was found for the hands AOI.

Age comparisons

Differences in the mean looking time at each of the four areas of interest between adults and child participants were investigated.

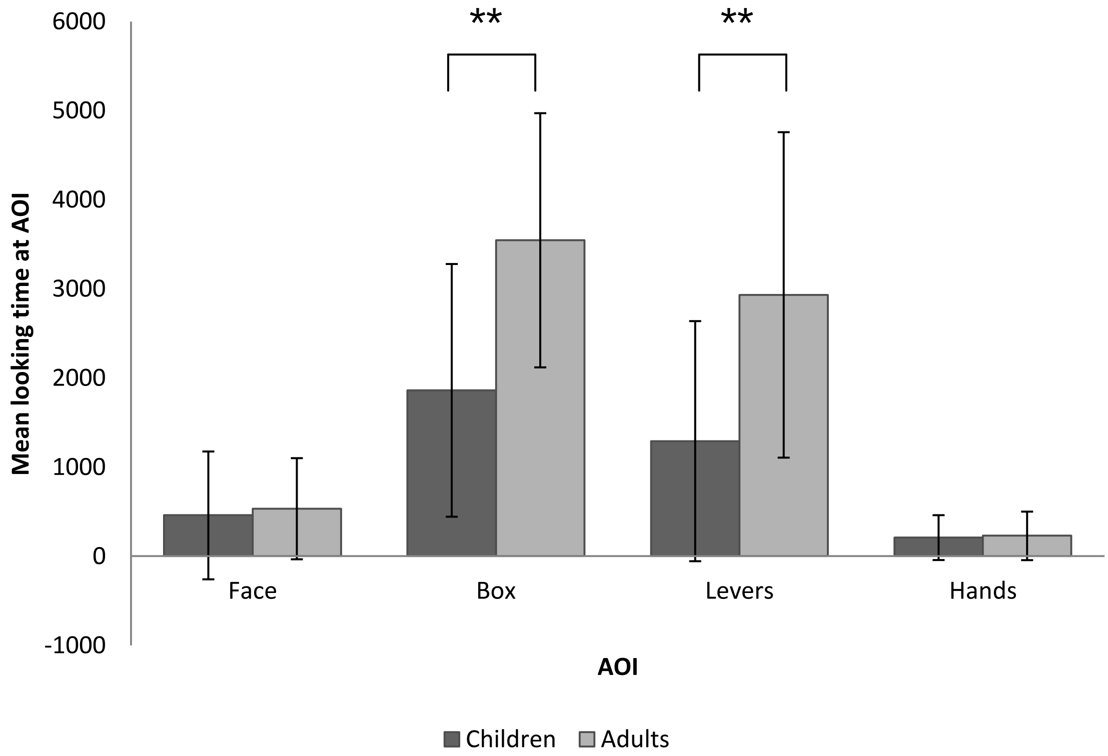


Figure 2.23: Mean looking time at each of the areas of interest (face, box, levers, hands) for child and adult participants. Error bars represent standard deviation. ** indicates significance at the .01 level.

A significant difference in the mean looking time at the box AOI was observed between children and adults: F(1,64)=23.07, p<.001, with adults having a higher mean looking time as indicated in figure 2.23. Similarly, a significant difference in the mean looking time at the levers AOI was observed between children and adults: F(1,64)=17.52, p<.001, again with adults attending more to the levers than children. No significant difference between children and adults was observed for the face AOI: F(1,64)=.214, p=.643, or for the hands AOI: F(1,64)=.089, p=.766. These results suggest that adults were better able to determine and subsequently attend to relevant aspects of the demonstration video more successfully than children.

2.4.4. Board Task: Behavioural Data

Participant data was scored depending on the number of irrelevant actions produced and the order in which the relevant actions were copied. Within this task, there were three possible irrelevant actions and two potential measures of the order of actions completed: the order in which participants completed the relevant actions (adding the pieces to the board) only, and the order in which they completed all actions. Participants were awarded a score of 1 for each of these elements that they completed, meaning that the maximum possible score for this task was 5. A quotient score was then calculated for each participant, by taking their actual score on the task and dividing it by the maximum possible score. A quotient score closer to 1 would indicate higher fidelity copying (overimitation), whereas a quotient score closer to 0 would indicate lower fidelity copying (emulation). As well as quotient score, participants' performance of individual aspects of the task, such as the irrelevant actions used, was also examined, in order to gain a better understanding of more nuanced behaviour during task completion.

Children

Table 2.5 indicates the mean quotient score for participants in each of the assigned conditions.

Table 2.5: Mean board quotient score for child participants in each of the assigned conditions

Assigned Condition	N	Mean	Std. Deviation
Copying	26	0.20	0.150
Goal Focused	25	0.20	0.216
Video Control	28	0.16	0.167
Control	17	0.02	0.066
Total	96	0.16	0.175

A significant difference was found between groups, using an ANOVA: $F(3,95)=4.87$, $p=.003$. Bonferroni post-hoc tests indicated a significant difference in mean quotient score between participants in the copying (M=.20) and control conditions (M=.002, $p=.005$), as well as between the goal focused (M=.20) and control conditions (M=.002, $p=.006$). There was no significant difference observed between participants in the video control and control conditions. As indicated in table 2.5, the mean quotient scores for all four groups were quite low, suggesting that either the instructions provided to participants in the copying group were not successful in influencing their behaviour when completing the task, or that participants were just less likely to copy faithfully when completing this task in comparison to the box task. Comparisons between participant performance on the box and board tasks are discussed in section 2.4.7.

To further understand the variability in task performance between participants in each of the assigned conditions, the percentage of participants achieving each of the possible quotient scores in each category was examined. These results can be seen in figure 2.24. A significant difference between groups was observed: $x^2(12)=31.18$, $p=.002$. The majority of participants in the control group achieved a score of 0, indicating that they were unlikely to spontaneously produce any of the irrelevant actions.

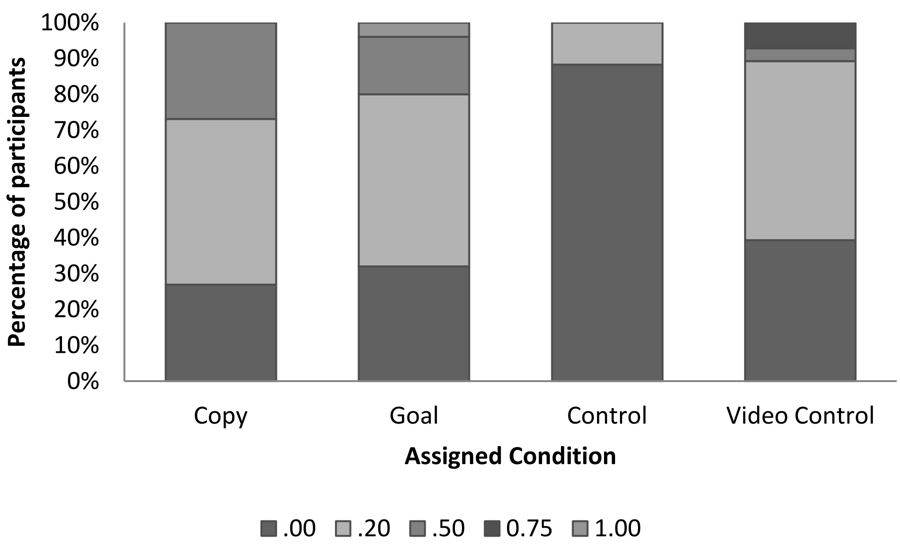


Figure 2.24: Percentage of child participants in each of the assigned conditions that achieved each of the possible quotient scores (0, .20, .50, .75, 1.00). No child participants achieved a score of .75.

The order in which participants completed the relevant actions (adding the pieces to the board) was examined. Results are shown in figure 2.25. There was no significant difference between participants in each of the conditions: $x^2(2)=.856$, $p=.652$, as the majority of participants in all conditions completed the relevant actions in the same order as the demonstrator.

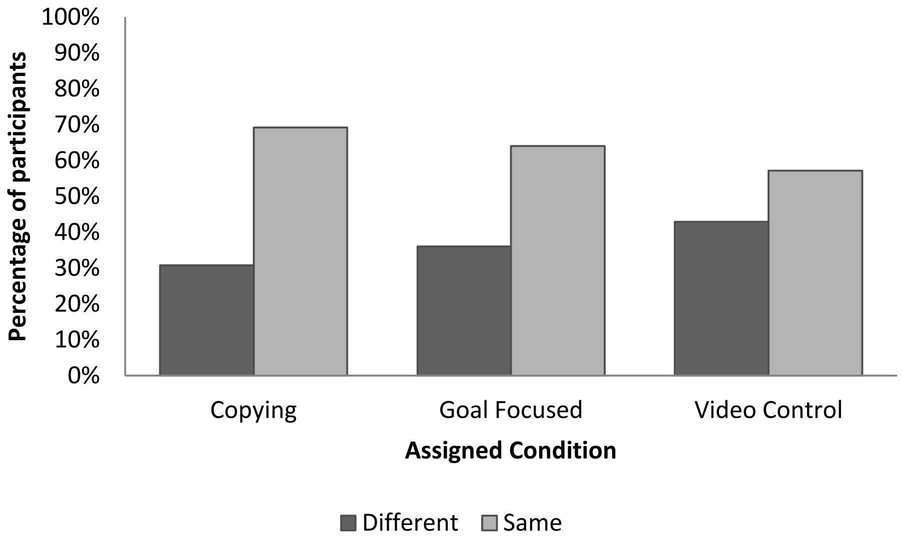


Figure 2.25: Percentage of child participants that completed the relevant actions in the same order as the demonstrator

The percentage of participants in each of the assigned conditions that completed each of the irrelevant actions was then investigated. Figure 2.26 indicates the percentage of participants in each group that completed the swirl action.

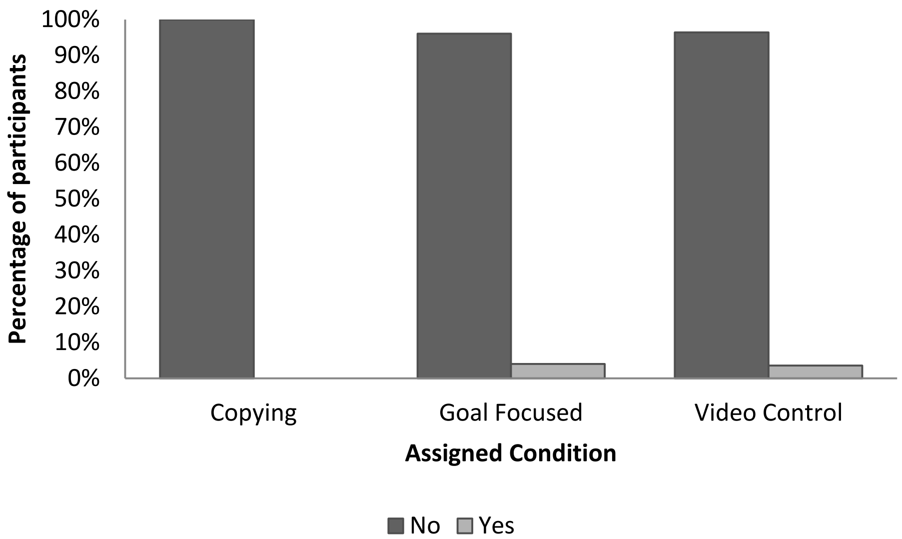


Figure 2.26: Percentage of child participants that completed the irrelevant swirl action

No significant differences between groups were found: $x^2(2)=1.02$, $p=.602$, since the majority of participants in all three of the assigned condition did not complete this action. Next, the wave action was examined. Results are shown in figure 2.27.

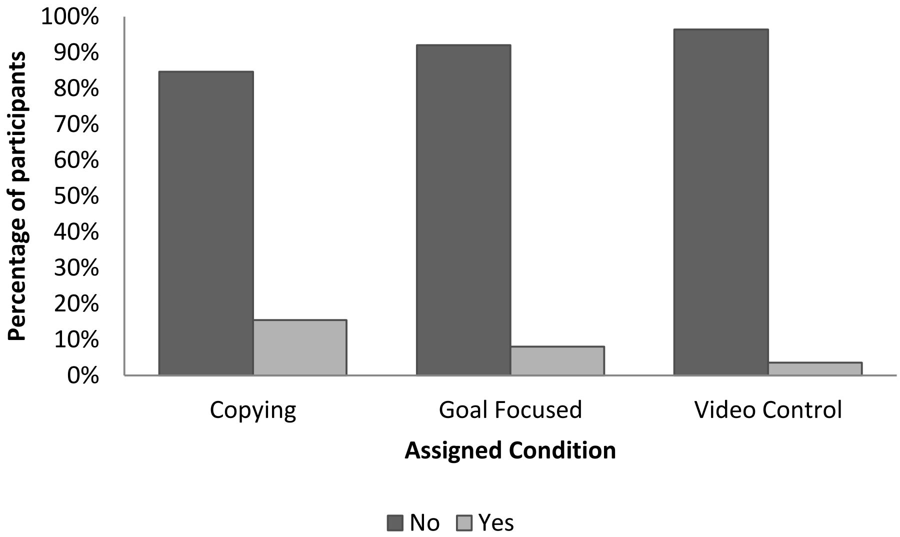


Figure 2.27: Percentage of child participants that completed the irrelevant wave action

As with the previous action, no significant differences between conditions were found: $x^2(2)=2.36$, $p=.307$. Again, the majority of participants in all groups did not complete the wave action. Finally, the slide action was examined. Results are shown in figure 2.28.

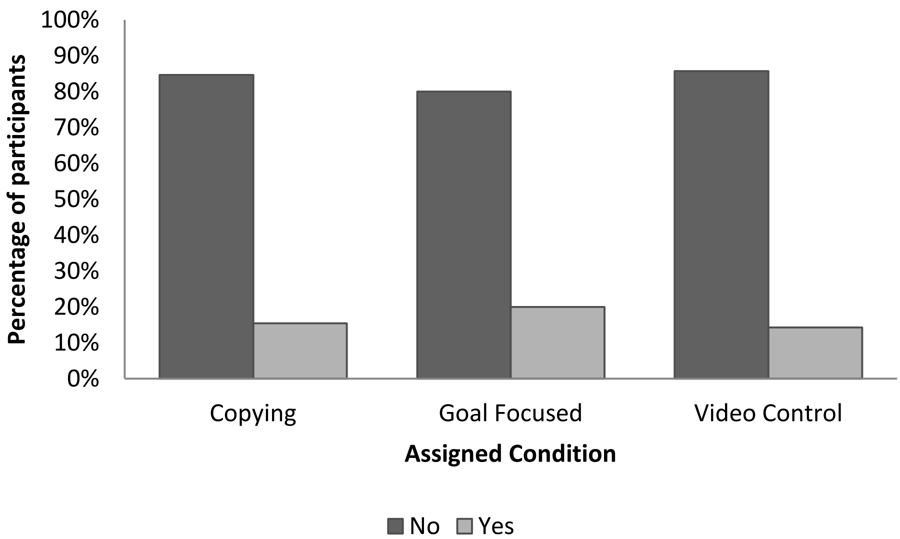


Figure 2.28: Percentage of child participants that completed the irrelevant slide action

Again, no significant differences were observed between groups: $x^2(2)=.346$, p=.841. As with the swirl and wave actions, the majority of participants in all groups did not complete this action. These results suggest that the instructions provided to participants in the copying group did not influence their subsequent behaviour when completing the task. However, it is also the case that the irrelevant actions in the board task were arguably far more obviously irrelevant than those in the box task, and so participants may have been more likely to omit them through choice if they believed that they genuinely were not necessary for task completion.

Adults

Table 2.6 indicates the mean quotient score for adult participants in each of the assigned conditions.

Table 2.6: Mean board quotient scores for adult participants in each of the assigned conditions

Assigned Condition	N	Mean	Std. Deviation
Copying	16	0.75	0.390
Goal Focused	16	0.35	0.278
Video Control	15	0.24	0.112
Total	47	0.45	0.357

Using an ANOVA, a significant difference in mean quotient score between assigned conditions was found: F(2,46)=13.76, p<.001. Bonferroni post-hoc tests indicated a significant difference in mean quotient score between participants in the copying (M=.75) and the video control (M=.24, p<.001) groups, as well as between the copying and goal focused (M=.35, p=.001) groups. As indicated in table 2.6, adults in the copying group achieved the highest mean quotient score, followed by participants in the goal focused group, and then by participants in the video control group. This suggests that the instructions provided to adults in the copying group successfully influenced participants to imitate actions shown in the demonstration video.

To further understand the variability in task performance between participants in each of the assigned conditions, the percentage of participants achieving each of the possible quotient scores in each category was examined. These results can be seen in figure 2.29.

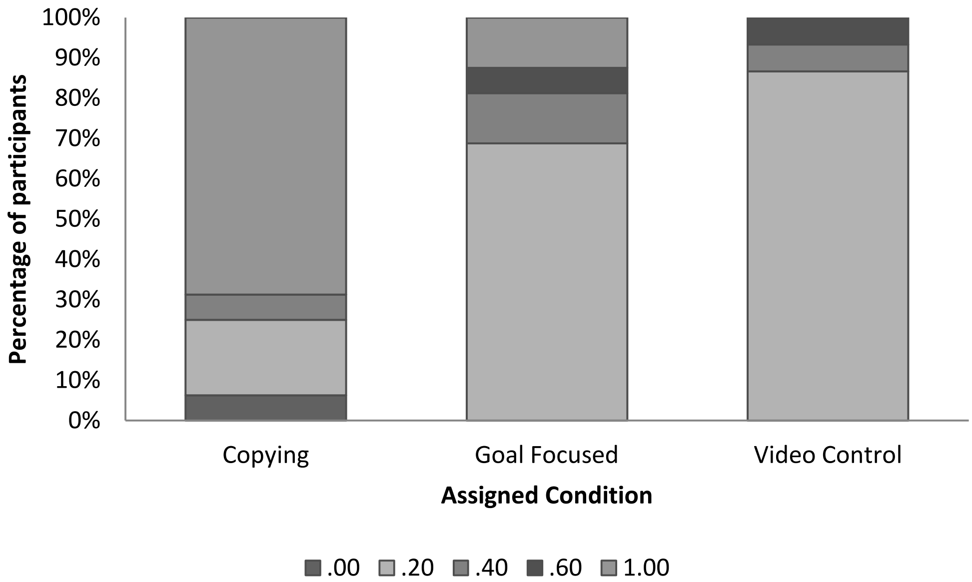


Figure 2.29: Percentage of adult participants in each of the assigned conditions that achieved each of the possible board quotient scores (.00, .20, .40, .60, 1.00). No adult participants achieved a score of .80.

A significant difference in quotient scores between participants in the assigned conditions was found: $x^2(8)=25.43$, p=.001. As indicated in figure 2.29, the majority of participants in the copying condition imitated precisely from the demonstrator and achieved the maximum score of 1. However, the majority of participants in both the goal focused and video control conditions achieved a score of .20. This further supports the suggestion that the instructions given to adult participants in the copying group were successful in influencing their behaviour. However, rather than adults in the video control group being “naturally” overimitative as expected, their performance was much closer to participants in the goal-focused group.

The percentage of participants in each of the assigned conditions that completed each of the irrelevant actions was then investigated. Figure 2.30 indicates the percentage of participants in each group that completed the swirl action.

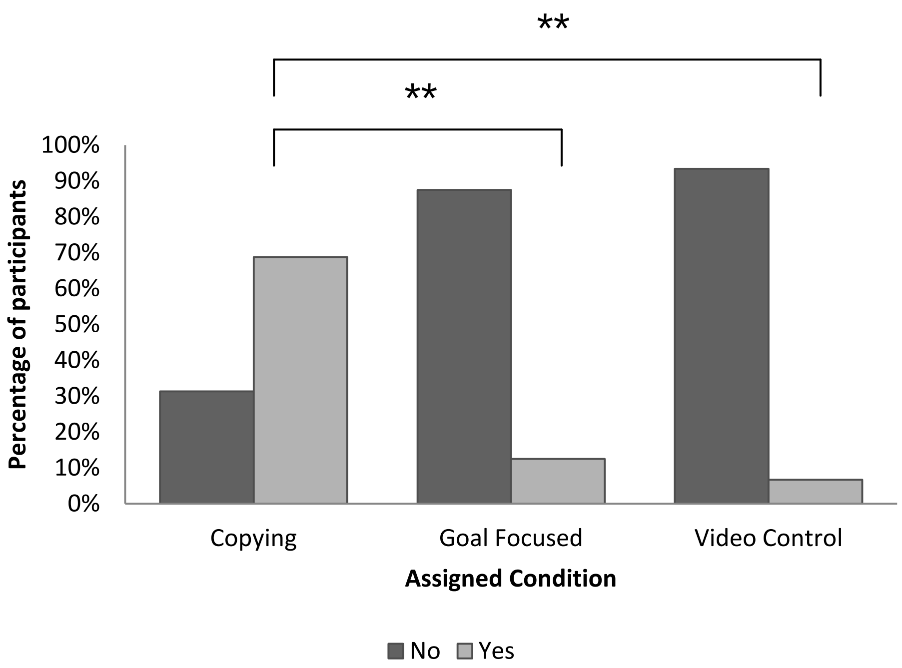


Figure 2.30: Percentage of adult participants that copied the irrelevant swirl action. ** indicates significance at the .01 level.

A significant difference between conditions was observed: $x^2(2)=17.73$, $p<.001$, with the majority of participants in the copying condition imitating this action. However, the majority of participants in the goal focused and video control conditions omitted the swirl action. Post-hoc comparisons of the proportions of participants in each condition that copied the action revealed a significant difference between participants in the copying and goal focused ($p=.004$) and the copying and video control ($p=.001$) groups. Next, the wave action was examined. Results are shown in figure 2.31.

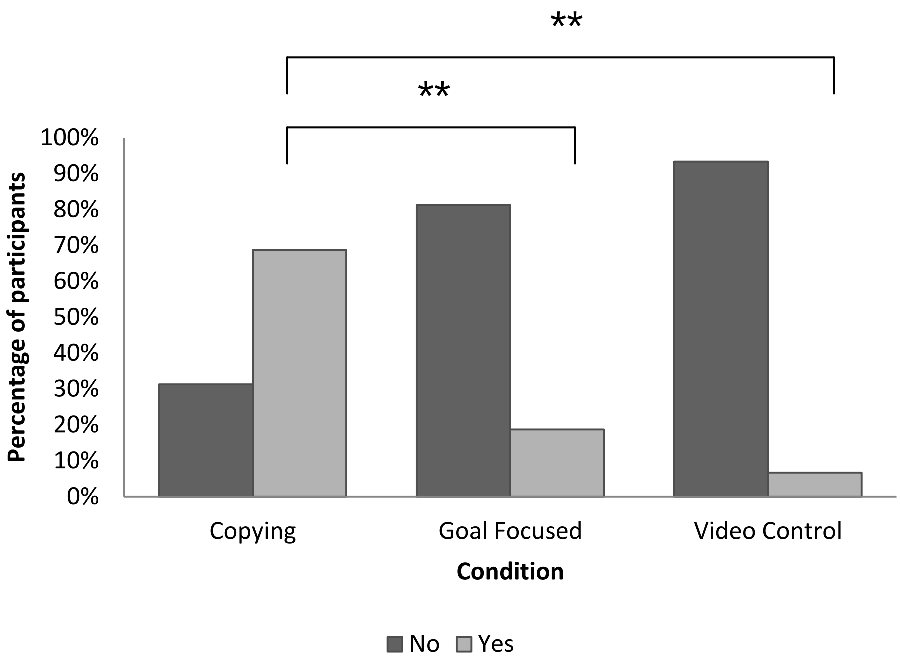


Figure 2.31: Percentage of adult participants that completed the irrelevant wave action. ** indicates significance at the .01 level.

As with the previous action, a significant difference was found between conditions: $x^2(2)=15.67$, $p<.001$. As before, the majority of participants in the copying group imitated the wave action, whereas the majority of participants in the goal focused and video control conditions did not. As before, a post-hoc test was used to compare the proportion of participants in each condition that copied the action. A significant difference was observed between participants in the copying and goal focused groups ($p=.013$) and participants in the copying and video control groups ($p=.001$), with participants in the copying group being significantly more likely to copy the action than those in the goal focused or video control groups. Finally, the slide action was examined. Results are shown in figure 2.32.

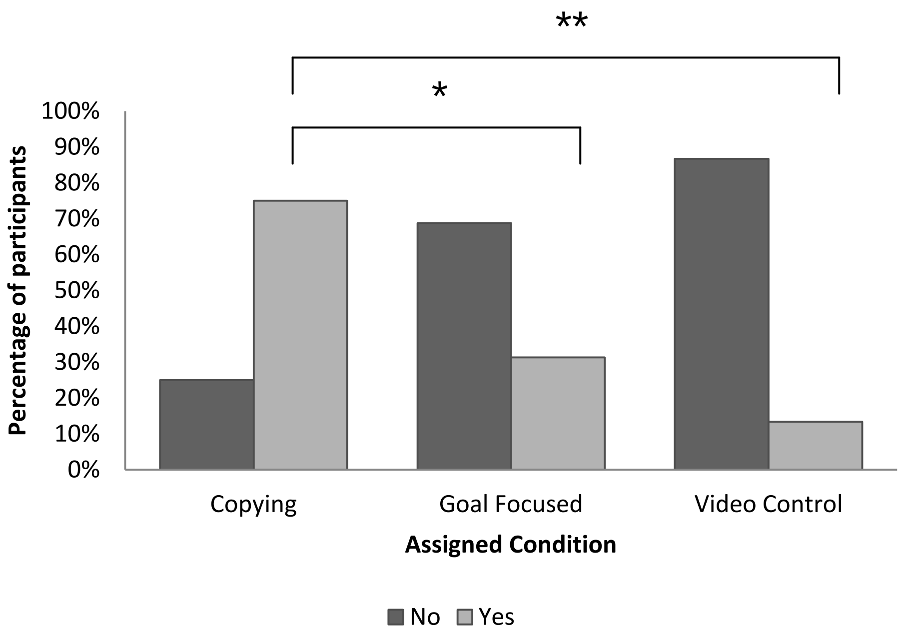


Figure 2.32: Percentage of adult participants that imitated the irrelevant slide action. * indicates significance at the .05 level, and ** indicates significance at the .01 level.

As with the previous two actions, a significant difference between conditions was found: x2(2)=13.07, p=.001, with the majority of participants in the copying group imitating the slide action, and the majority of participants in the goal focused and video control conditions omitting it. Post-hoc tests were used to compare the proportion of participants in each condition that completed the action. As before, a significant difference was found between participants in the copying and goal focused groups (p=.039) and the copying and video control groups (p=.002), with participants in the copying group being significantly more likely to imitate the action than those in the goal focused and video control groups.

No statistical comparisons were made between adult participants in each of the three conditions (copying, goal focused, video control) based on the order in which they copied the relevant actions from the board task, since only one participant in the copying group and one participant in the video control group used a different copying order than that demonstrated in the video. All participants in the goal focused group used the same order as the demonstrator.

Age Differences

As in the box task, behavioural differences between the ten youngest and ten oldest child participants were also examined to determine if there was a difference in copying fidelity. The difference between the ten youngest (mean age 45.8 months) and ten oldest (mean age 60.4 months) participants is indicated in table 2.7.

Table 2.7: Mean board quotient scores for the ten youngest and ten oldest child participants

Age Group	N	Mean	Std. Deviation
Youngest	10	0.20	0.163
Oldest	10	0.18	0.175
Total	20	0.19	0.165

No significant difference in mean quotient score was found between the ten youngest and the ten oldest participants: $F(1,19)=.070$, $p=.795$.

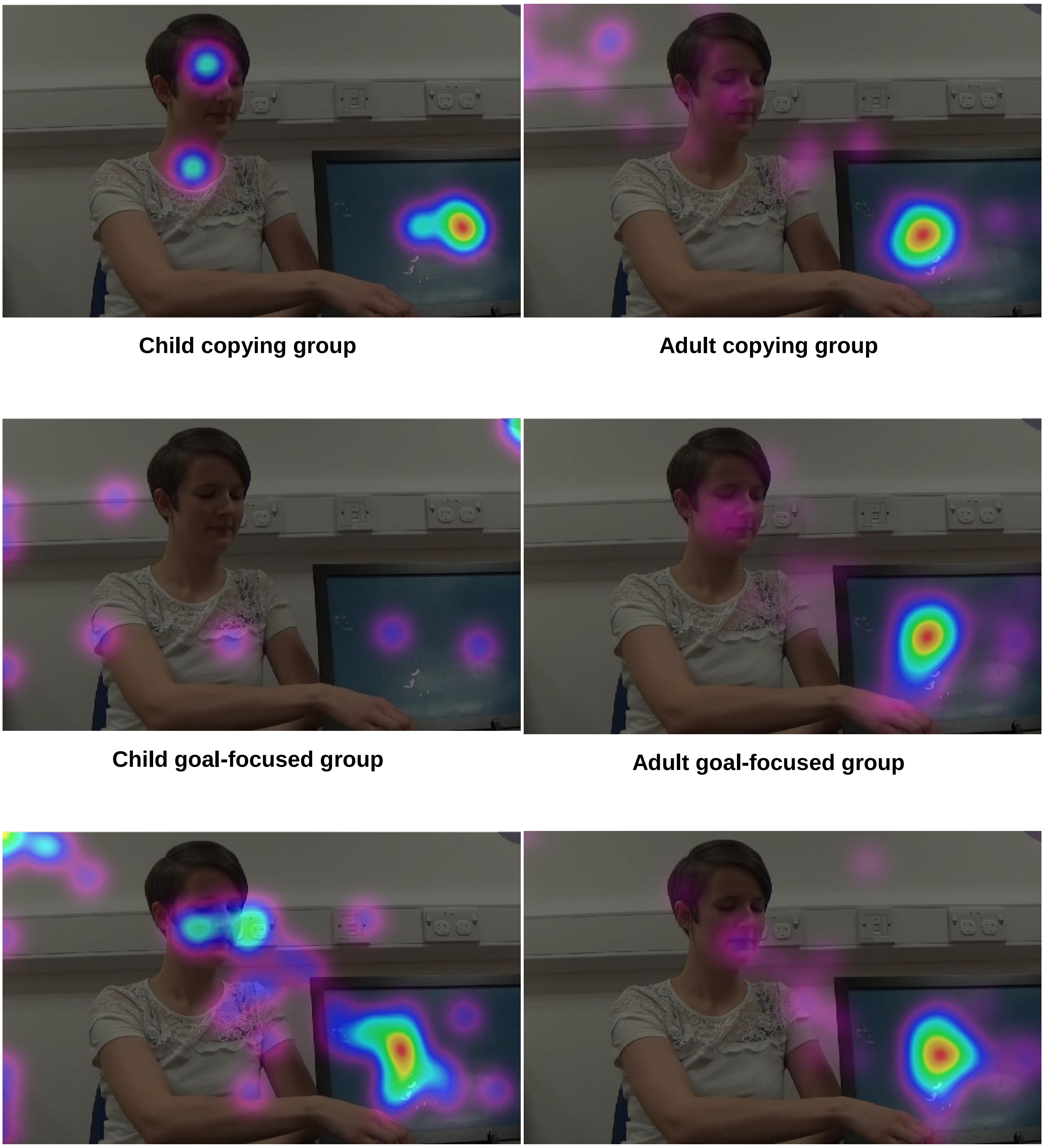
Comparisons were then made between child and adult participants, using an ANCOVA with assigned category as a covariate. There was a significant difference between participants in different age groups: $F(1,125)=32.95$, $p<.001$, and a significant difference between categories: $F(1,125)=14.13$, $p<.001$. As indicated in table 2.8, adult participants achieved a higher mean quotient score overall (M=.45) in comparison to child participants (M=.18). This may be due the fact that child participants may have struggled to remember the actions used by the demonstrator, and therefore used their own strategy instead.

Table 2.8: Mean board quotient scores for adult and child participants (all conditions)

Age Group	N	Mean	Std. Deviation
Children	79	0.18	0.178
Adults	47	0.45	0.357
Total	20	0.28	0.289

2.4.5. Board Task: Eye Tracking Data

Eye tracking data was collected for 38 children and 45 adults. Heatmaps using the mean fixation positions for each group were calculated (see figure 2.33), which appear to suggest that adult participants in all groups were better able to direct their attention towards the task than child participants, as with the box task.



Child copying group | **Adult copying group**

Child goal-focused group | **Adult goal-focused group**

Child video control group | **Adult video control group**

Figure 2.33: Heatmaps of mean fixation locations in each group

In order to investigate these differences further, the proportion of time spent looking at each of the previously identified areas of interest (demonstrator's face, the magnetic board, the puzzle pieces, and the demonstrator's hands) was calculated for the first video trial. Information from the first trial was used as no significant differences between looking times at the four areas of interest were found between trials. Examples of each of the four areas of interest can be found in Appendix B. Looking time at the background of the demonstration video was analysed separately, in order to measure participants' ability to attend to relevant stimuli. This is discussed further in the age comparisons section below.

2.4.6. Overall looking time at areas of interest

Children

Figure 2.34 indicates the mean time spent looking at each of the AOIs for each condition with child participants (copying N=14, goal-focused N=11, video control N=13).

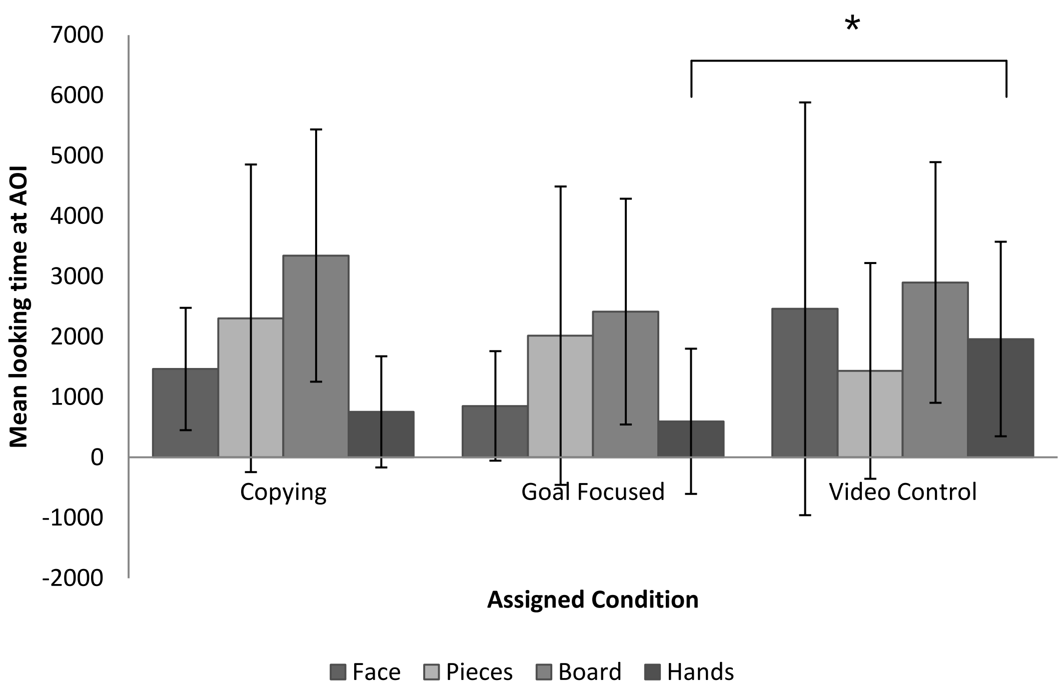


Figure 2.34: Mean looking time at each of the areas of interest (face, board, pieces, hands) for children in each of the assigned conditions. Error bars indicate standard deviation. * indicates significance at the .05 level.

There was a significant difference observed between participants in the three conditions for the hands AOI: F(2,36)=4.49, p=.018. Bonferroni post-hoc tests indicated a significant difference between participants in the goal focused group (M=598.34) and those in the video control group (M=1961.63, p=.043). Additionally, the difference between participants in the copying group (M=754.92) and the video control group was approaching significance (p=.051). Participants in the video control group spent more than twice as much time attending to the demonstrator's hands in comparison to those in the copying and goal focused groups.

No significant difference between groups was observed for participants' looking time at the pieces AOI: F(2,36)=.557, p=.578, the board AOI: F(2,36)=.633, p=.537, and the face AOI: F(2,36)=1.63, p=.211.

The difference in mean looking time at each of the areas of interest was then investigated for child participants achieving each of the possible quotient scores. These results can be seen in figure 2.35.

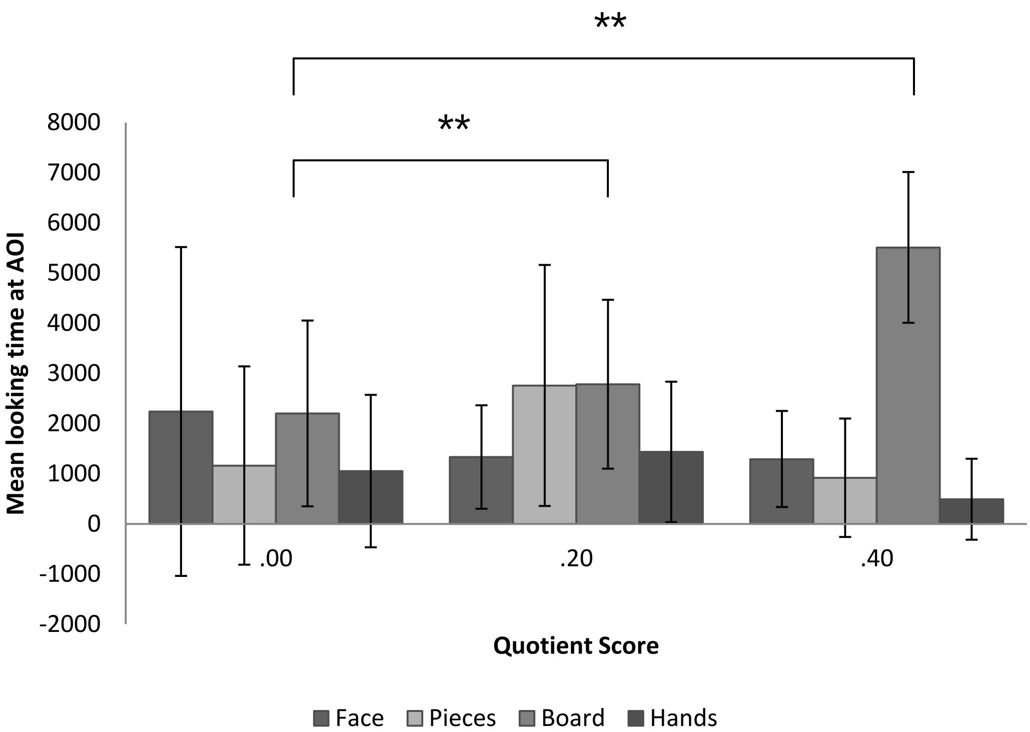


Figure 2.35: Mean looking time at each of the areas of interest (face, board, pieces, hands) for child participants achieving each of the possible quotient scores (.00, .20, .40). No child participants achieved a score of .60, .80 or 1.00. Error bars represent standard deviation. ** indicates significance at the .01 level.

Comparisons between participants who achieved each of the quotient scores were made, using an ANCOVA, with assigned condition as a covariate. It is important to note that no child participants scored higher than .40 in this task. There was a significant difference in the mean looking time at the board AOI based on quotient score: $F(1,38)=7.48$, $p=.002$, with participants who achieved a score of .40 spending much longer looking at the board (M=5508) in comparison to those who scored 0 (M=2199) or .20 (M=2782). Bonferroni post-hoc tests indicated that the difference between those who scored 0 and .40 was significant ($p=.002$), as was the difference between those who scored .20 and .40 ($p=.011$). However, there was no significant effect of assigned condition: $F(1,38)=1.91$, $p=.176$.

The difference between participants in the quotient score groups was approaching significance for the pieces AOI: $F(1,38)=2.95$, $p=.066$, but again, there was no significant effect of assigned condition: $F(1,38)=1.55$, $p=.222$. For the hands AOI, there was a significant difference between participants in each of the assigned conditions, as expected: $F(1,38)=6.14$, $p=.018$, but no significant difference between quotient score groups: $F(1,38)=1.02$, $p=.372$. For the face AOI, there was no significant difference for either assigned condition: $F(1,38)=.662$, $p=.471$ or quotient score: $F(1,38)=.346$, $p=.710$.

Adults

Figure 2.36 indicates the mean time spent looking at each of the AOIs for each condition with adult participants (copying N=16, goal-focused N=15, video control N=14).

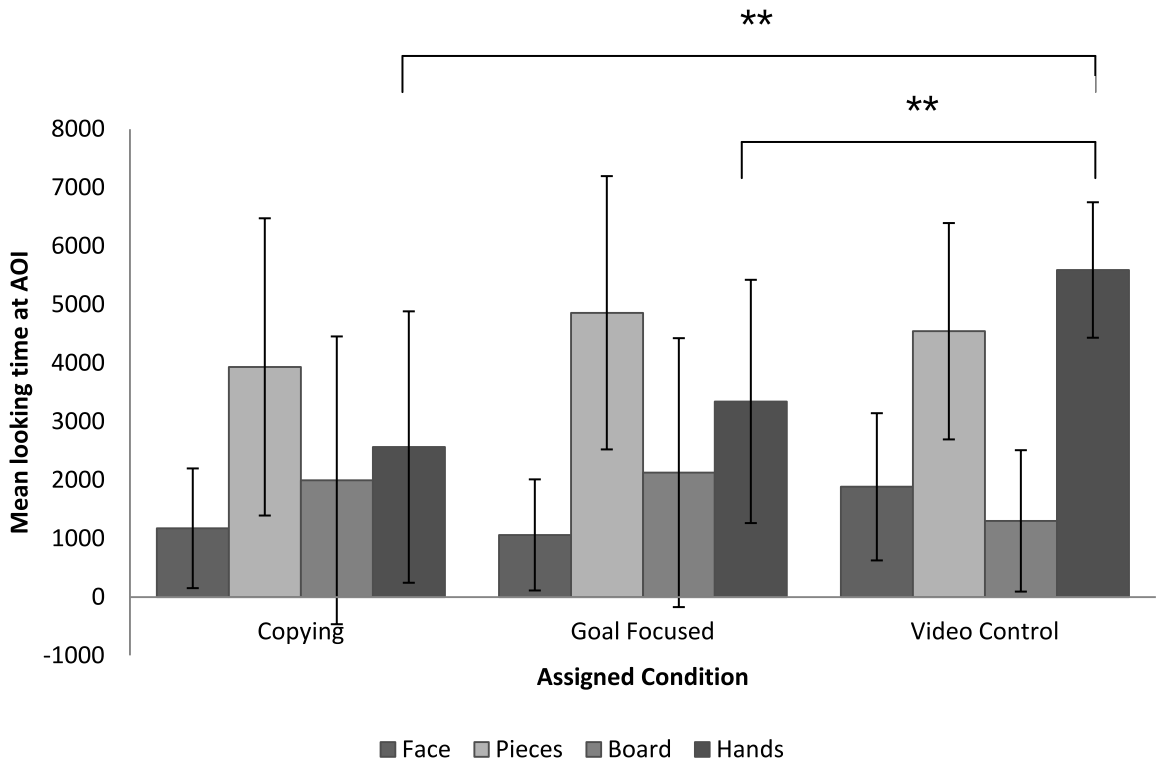


Figure 2.36: Mean looking time at each of the areas of interest (face, pieces, board, hands) for adult participants in each of the assigned conditions. Error bars represent standard deviation. ** represents significance at the .01 level.

An ANOVA indicated significant differences between conditions for the hands AOI: $F(2,44)=9.58$, $p<.001$. Bonferroni post-hoc tests indicated a significant difference between participants in the copying (M=2563) and video control groups (M=5590, $p<.001$), as well as between participants in the goal focused (M=3343) and video control groups ($p=.010$). Participants in the video control group attended to the demonstrator's hands for a significantly longer period of time in comparison to those in the other two conditions. No significant difference was observed between conditions for the face AOI: $F(2,44)=2.46$, $p=.098$, the pieces AOI: $F(2,44)=.664$, $p=.520$, or the board AOI: $F(2,44)=.649$, $p=.529$.

Comparisons between participants who achieved each of the quotient scores were made, using an ANCOVA, with assigned condition as a covariate. Results are displayed in figure 2.37.

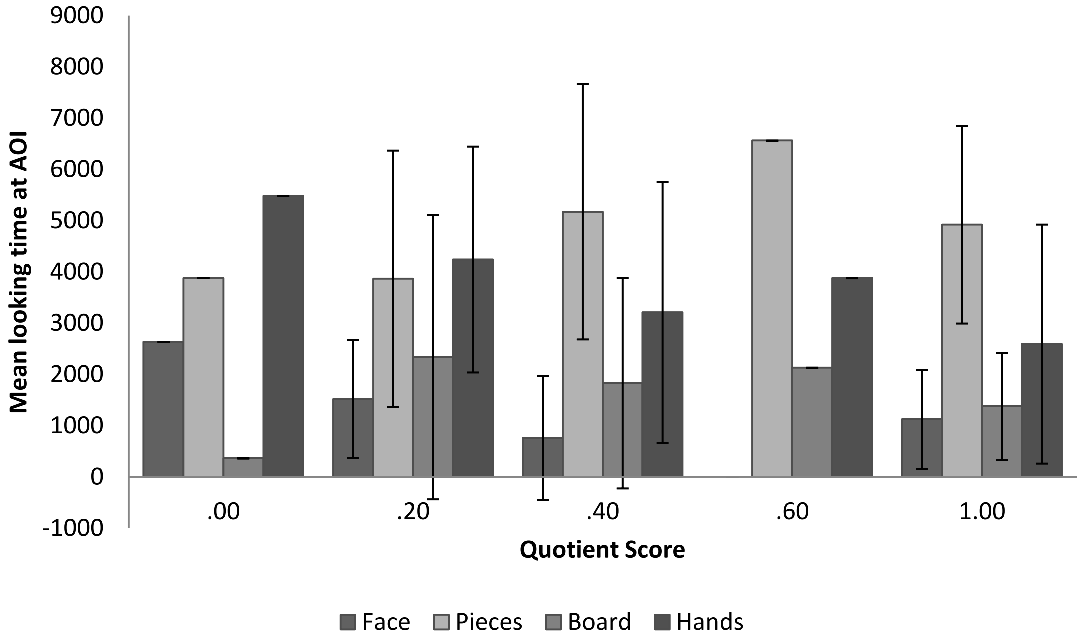


Figure 2.37: Mean looking time at each of the areas of interest (face, pieces, board, hands) for adult participants achieving each of the possible quotient scores (.00, .20, .40, .60, 1.00). No adult participants achieved a score of .80. Error bars represent standard deviation.

For the hands AOI, there was a significant difference between participants in the assigned conditions, as expected, F(1,44)=13.20, p=.001. However, there was no significant difference between quotient scores: F(4,44)=.980, p=.430. Similarly, for the board AOI, there was a significant difference for participants in the assigned conditions: F(1,44)=5.25, p=.027, but no significant difference between quotient scores: F(4,44)=1.51, p=.218. For the pieces AOI, there was no significant difference between participants in the assigned conditions: F(1,44)=3.74, p=.060, or for the different quotient scores: F(4,44)=1.53, p=.214. Finally, for the face AOI, there was no significant difference between participants in the assigned conditions: F(1,44)=3.16, p=.083, or for the different quotient scores: F(4,44)=1.52, p=.216.

Age comparisons

The mean looking time at each of the areas of interest was compared for adult and child participants (see figure 2.38).

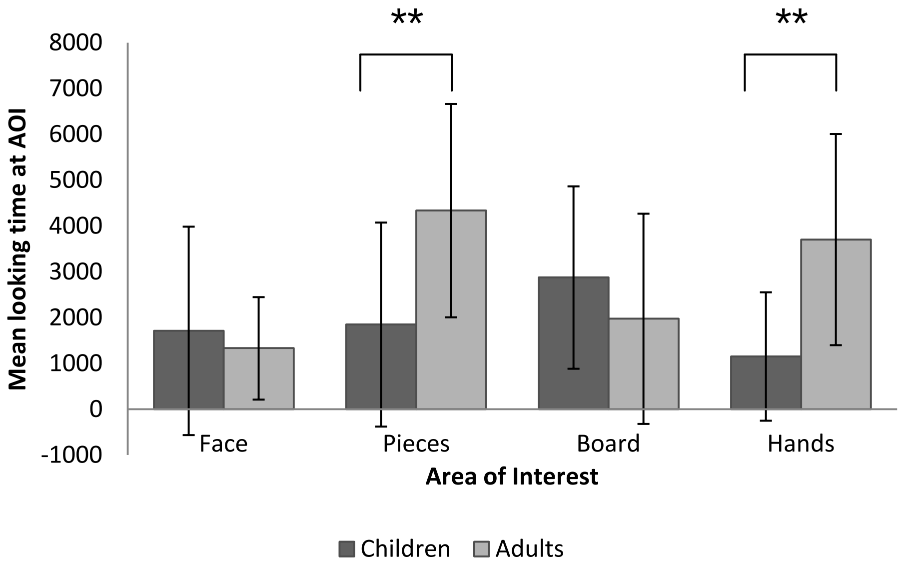


Figure 2.38: Mean looking times at each of the areas of interest (face, board, pieces, hands) for child and adult participants. Error bars represent standard deviation. ** indicates significance at the .01 level.

A significant difference between age groups was observed for the pieces AOI: F(1,85)=25.43, p<.001, with adults having a significantly higher mean looking time (M=4336) at this AOI than children (M=1849). Similarly, a significant difference between age groups was observed for the hands AOI: F(1,85)=37.08, p<.001, with adults attending significantly more to the demonstrator's hands (M=3704) than children (M=1152). The difference between age groups for the board AOI was approaching significance: F(1,85)=3.73, p=.057, with children (M=2875) attending more to the board than adults (M=1974). However, no significant difference was observed for the face AOI: F(1,85)=1.01, p=.317.

2.4.7. Comparisons between tasks

Differences between tasks and conditions

Since all participants completed both the box and board tasks, comparisons could be made between tasks to determine whether particpants' performance remained consistent, or was task-dependent.

Children

The mean quotient score obtained for child participants in each condition (copying, goal focused, video control) in both tasks can be seen in table 2.9. A Spearman's rho correlation was used to compare the mean quotient scores for participants in each group across the board and box tasks. A strong, positive, significant correlation was found for board and box quotient scores for participants in the copying group: r_s=.987, p<.001. For the goal condition, a strong, positive, significant correlation was found between board and box scores: r_s=.792, p<.001, suggesting that participant performance remained consistent across tasks. Finally, for the video control condition, a very weak negative correlation was found between task scores. This was not found to be significant: r_s=-.210, p=.282. Although this result was not significant, the negative correlation between scores suggests that participants' score on one task was likely to increase as their score on the other task decreased. This can be seen in the difference in participants' mean quotient score for the board task (M=.16, SD=.167) and the box task (M=.65, SD=.306), which are markedly different in comparison to other groups. For this reason, a repeated measures ANOVA was used to compare mean scores between the board and box tasks. There was a significant difference between board and box scores for participants in the copying group: F(1,25) = 42.94, p<.001, the goal focused group: F(1,24)=8.84, p=.007 and the video control group: F(1,27)=48.34, p<.001, with participants in all groups achieving higher mean scores in the box task, as indicated in table 2.9.

Table 2.9: Mean board and box quotient scores for child participants in each assigned condition

	N	Board Quotient Score	Std. Deviation	Box Quotient Score	Std. Deviation
Copying	26	0.20	0.14967	0.48	0.38443
Goal	25	0.20	0.21602	0.37	0.32759
Video Control	28	0.16	0.16651	0.65	0.30686
Total	79	0.18	0.17765	0.51	0.35549

Adults

The mean quotient score obtained for adult participants in each condition (copying, goal focused, video control) in both tasks can be seen in table 2.10. As with child participants, Spearman's rho correlations were used to compare participant performance across both tasks. For the copying group, a significant positive correlation was found between board and box scores: r_s=.718, p=.002, indicating that participants' behaviour remained similar across both of the tasks. For the goal focused group, there was a weak positive correlation between board and box scores, but it was not found to be significant: r_s=.409, p=.116. Finally, for the video control group, there was a weak positive correlation between board and box scores, but again, this was not found to be significant: r_s=.326, p=.256.

Table 2.10: Mean board and box quotient scores for adult participants in each of the assigned conditions

	N	Board Quotient Score	Std. Deviation	Box Quotient Score	Std. Deviation
Copying	16	0.75	0.38987	0.89	0.24098
Goal	16	0.35	0.27809	0.64	0.30233
Video Control	14	0.21	0.05345	0.59	0.34817
Total	46	0.45	0.36070	0.71	0.32045

In order to further investigate where differences lay, the differences between mean quotient scores on the box and the board tasks for each condition in adults were examined using a repeated measures ANOVA. There was no significant difference found between board and box scores for participants in the copying group: F(1,15)=4.22, p=.058. However, the difference was significant for participants in the goal focused condition: F(1,15)=14.42, p=.002 and the video control condition: F(1,12)=17.67, p=.001.

Whilst the differences in score between conditions are very marked in the board task, the differences in scores between conditions in the box task are much lower. This suggests that all participants were more likely to overimitate in the box task, regardless of the instructions that they had been given.

These results indicate that performance was different between tasks for both adults and children in all assigned conditions, suggesting that participants were able to use flexible copying strategies where necessary.

2.5. Discussion

The purpose of this experiment was to investigate copying behaviours in 3- to 5-year-old participants and adults. Three broad aims were identified: to induce particular patterns of copying behaviour in both child and adult participants with the aim of establishing templates of looking patterns with reference to overimitators, imitators and emulators; to determine whether developmental differences in task performance and looking behaviours are observed in children and adults, and to investigate differences between functional and arbitrary task types. An additional overall aim was to investigate the viability of eye tracking as a means to determine what information participants attend to during a demonstration of a novel task, with particular reference to how this then affects their subsequent behaviour when completing the task themselves.

2.5.1. Summary of findings

- No significant differences in quotient scores were found for child participants in each of the assigned conditions (copying, goal focused, video control) in either the box or board tasks.
- Differences in task performance were observed for child and adult participants when completing the box and board tasks: mean quotient scores in the board task were lower than those in the box task.
- In both the board and box tasks, adults in the copying condition had higher quotient scores than those in the video control and goal focused groups.
- No gaze patterns that might be used to distinguish a high fidelity copier from a low fidelity copier were observed in either children or adults.

2.5.2. Attempts to induce emulation and overimitation

In the box task, the attempt to induce particular copying behaviours appears to have been unsuccessful, as there was no significant difference in the mean quotient score for each of the assigned categories (copying, goal focused, video control). The instructions given to adults in the copying group do appear to have been successful in both the board and box tasks, as adults in this group were found to achieve a significantly higher quotient score in both tasks. However, the scores for adults in the video control and goal focused groups were very similar in both tasks, indicating that the instruction to complete the tasks as quickly as possible did not encourage participants to emulate in the way that was hoped. This is consistent with previous findings by Flynn & Smith (2012), who suggested that adults continued to overimitate, even when time pressure and the possibility of losing a monetary reward were involved. One explanation for overimitation suggests that it is primarily a social behaviour, which allows people to show affiliation with an individual or group (Over & Carpenter, 2012). This may be reliant on theory of mind (Seyfarth & Cheney, 2013), which develops around the age of 5 and allows children to have an awareness of the motivations of others (Rakoczy & Schmidt, 2013). As only 26% of children tested in this study were 5 years old, it may be that this has affected the results, as no data on participants' theory of mind ability was collected. It would be beneficial to include a measure of participants' theory of mind ability in future studies, in order to determine whether this is a crucial factor in overimitation. This would also help to explain why adults were found to display more flexible copying across tasks, as they may be able to better interpret the demonstrator's intentions. However, if this is the case, it may still be possible that adults overimitated in order to display affiliative behaviour.

In the board task, the majority of children in all three conditions were found to imitate the demonstrator, but not overimitate, meaning that they copied the order in which the demonstrator placed the pieces onto the board, but omitted the irrelevant actions. No significant differences in behaviour were observed between the copying, goal-focused and video control groups, or in the specific irrelevant actions performed. Again, this suggests that the attempt to induce particular copying behaviours in these participants failed, and that eye tracking patterns relating to copying behaviours should be interpreted with caution. These differences may have occurred because children did not understand the instructions given, or because the cognitive load involved in remembering the actions themselves, as well as the task goal, was too high, resulting in children prioritising task completion over following specific instructions to achieve this goal (Subiaul & Schilder, 2014). For this reason, including a measure of participants' memory ability in future studies may be beneficial in determining whether participants adopt emulative strategies due to the fact that they have forgotten what took place in the demonstration video. Furthermore, the irrelevant actions within the board task were gesture-based. Although children have been found to overimitate gestures as well as functional actions (Subiaul, Winters, Krumpak, & Core, 2016), it has been suggested that they are less likely to copy gestures that are not linked to an object or tool (Taniguchi & Sanefuji, 2017). The wave action in particular was not at all linked to any of the puzzle pieces, which may explain why this action was so rarely copied.

2.5.3. Comparisons between board and box task performance

Comparisons between the board and box task were made, which indicated that performance across tasks was consistent for adult participants in the copying group, and for child participants in the goal-focused group, but not for participants in any of the other groups. Of particular interest was the marked difference in board and box quotient scores for both child and adult video control participants. This potentially suggests that participants were simply more likely to overimitate on the box task than the board task, perhaps because the irrelevant actions used in the board task were arguably more obviously irrelevant and therefore unnecessary for task completion. The board puzzle was also more likely to be familiar to participants, and therefore they may have used existing strategies and knowledge about those types of puzzle to complete it, even when asked to copy. However, these results do suggest that both adults and children were capable of flexible copying. One potential explanation for this finding in children is that there were four actions involved in completing the box task, and five in the board task. It may be the case that younger participants found it too difficult to remember all five actions, and therefore employed a more emulative strategy. However, this does not explain the differences in task performance for adults, in which the vast majority of participants who had been identified as imitators in the board task were then found to overimitate in the box task. These results may suggest that overimitation is context- and task-dependent, with higher levels of overimitation being observed in the more novel box task, even with adult participants who were likely to have been able to determine how the box functioned from the demonstration video. One explanation may be that children have been observed to display selective copying of actions that appear causally relevant to task completion (Keupp, Behne & Rakoczy, 2013). It may be the case that the actions used in the board task obviously did not appear to be relevant or necessary, and therefore participants omitted them. Additionally, as discussed, it is likely that the board task was more familiar to participants, who may have had experience with similar objects such as jigsaw puzzles. Therefore, participants may have chosen to omit the irrelevant actions as they understood that they were not necessary for task completion; these results are consistent with those found by Keupp et al., (2015), who found that children are able to determine whether actions are rational or not, and therefore choose whether it is necessary to copy them. This is in direct opposition to the suggestion that overimitation is an automatic process.

Two tasks were used in order to investigate the differences in task performance during both a functional, goal-driven task (box task), and a more arbitrary task (puzzle board task). The board task was adapted from Dickerson et al (2013), and involved completing a puzzle using an arbitrary sequence of actions, meaning that participants were not restricted to a smaller range of possible methods that they could use to achieve the goal state. This was intended to fully explore a range of overimitative versus emulative behaviours. However, it is likely that puzzle completion was familiar to the vast majority of participants, whereas the box task was much more novel, and it is possible that this had an effect on the behaviour of participants as they may have been more aware that the irrelevant actions were not necessary for puzzle completion, and therefore omitted them. As previously discussed, research has suggested that overmitation occurs in order to allow children to better learn and understand causally opaque tasks (Whiten et al., 2009). If the board task was truly very familiar to the participants, then perhaps this explains the lack of overimitation observed. Similarly, the puzzle board was less goal-driven in that there was no observable “reward” for completing the task, unlike the box task, which involved the retrieval of a reward-like object (the plastic egg). This may also have reduced overimitation in children. In fact, Loucks & Meltzoff (2013) suggest that young children are more likely to prioritise, and therefore remember, information related to goals as opposed to arbitrary sequences. However, this does not explain the increased levels of overimitation seen in some adult participant groups. Adults appear to have followed the instructions provided more carefully than children, with adults in the copying group being much more likely to copy irrelevant actions in comparison to those in the goal focused and video control groups. In fact, adults in the video control group for both the box and board tasks showed barely any overimitative behaviour, suggesting that the previous high levels of overimitation occurred due to the instructions provided prior to task completion.

2.5.4. Eye tracking patterns

An additional aim was to determine whether there were differences in gaze patterns based on task performance. However, calibration success rates were very poor with child participants. Future research should investigate ways to reduce the number of datasets lost through failed calibration.

No significant differences in looking patterns were observed for child participants between assigned conditions in the box task. As previously stated, since the attempts to induce behaviour were not successful, this may be expected, and if any significant differences had been observed then these gaze patterns may not have been truly representative of overimitators or emulators, since those behaviours were not observed during task performance. A significant difference in the looking time at the face AOI was observed for both child and adult participants achieving lower quotient scores. It may be the case that since these individuals chose to focus more on the demonstrator's face, they missed the chance to attend to the actions as they were being performed, and so could not subsequently copy them, even if they wished to. Further research should investigate what drove these participants to attend to the demonstrator's face, as opposed to the task, and whether individual differences play a role in this. When comparing the mean looking time at each AOI between child and adult participants, it was determined that adults looked significantly more at the levers and the box than children did, suggesting that adults are better able to focus their attention towards relevant stimuli than children are (Lane & Pearson, 1982; McClelland et al., 2013). This may mean that they may better understand the task, but also that they are more likely to observe irrelevant actions taking place. It may also be the case that adults are better able to direct their attention towards relevant stimuli because they may have more of an understanding as to how the tasks could potentially work, and so therefore focus their attention on areas of the screen in which they expect the most important information to be. This would be consistent with findings that suggest that the ability to attend to relevant stimuli increases with age (Hagen & Hale, 1973)

Unlike the box task, a significant difference between the mean looking time at the hands AOI was observed between conditions for child participants, with participants in the video control group attending to the demonstrator's hands for significantly longer than participants in the copying or goal focused groups. Additionally, a significant difference in the mean looking time at the board was seen for child participants who achieved the highest quotient score. Children who achieved a higher quotient score were found to have attended to the board for significantly longer than children who achieved lower scores. Again, it may be the case that their ability to attend to this relevant area of interest simply meant that they had observed the actions taking place and were therefore better placed to copy them if they chose to do so. As with child participants, adults in the video control group were also found to attend to the demonstrator's hands for significantly longer than participants in the copying and goal focused groups. It is particularly interesting that this pattern is consistent across age groups, and that increased attention to the demonstrator's hands does not appear to have resulted in higher fidelity copying. Finally, when comparisons were made between adults and children, adults were found to look significantly more at the puzzle pieces and the demonstrator's hands, again suggesting that they were better able to attend to more relevant stimuli as expected.

Overall, differences between age groups and low and high fidelity copiers were identified in experiment 1, as well as differences in gaze patterns between groups, but some inconsistencies remain. Further steps are required in order to investigate why these differences in copying behaviours occur.

2.5.5. Future Research

Instructions provided to participants

The attempt to induce particular copying behaviours was unsuccessful with the majority of child participants, although adult task performance appeared to match more closely with the instructions provided to them. With relation to child participants, it should be noted that the inclusion of specific instructions are likely to have affected behaviour somehow, however, these differences were not as obviously apparent in the results obtained as previously hoped, which makes it difficult to determine how effective the instructions actually were. It may be the case that the attempts to artificially increase rates of specific copying behaviours in participants were in fact overridden by participants' natural copying tendencies, which would agree with previous research that suggests that overimitation is a highly pervasive, and possibly automatic, behaviour, as suggested by Lyons, Damrosch, Lin, Macris, & Keil (2011). The inclusion of specific instructions in this study was designed to determine whether particular gaze patterns could be identified for those who copied faithfully, and those who copied selectively. However, even though the instructions did not always result in the expected behaviours, the comparison of gaze patterns between participants was still possible, based on their task scores as opposed to the groups they were assigned to. Further experiments will therefore not include focused instructions to participants on how to complete the tasks, in order to investigate more natural behaviour patterns and methods of task completion.

Task issues

Although the board puzzle was included in order to understand differences between functional and arbitrary task types, it may be that the board was not arbitrary enough as there was still a fairly limited potential sequence of actions, and a clear end goal. It was also difficult to incorporate irrelevant actions into this task that were not either obviously irrelevant, or linked to one of the relevant actions. Therefore, a more open-ended task, with a larger range of potential responses, should be used instead.

Although puzzle boxes are commonly used in imitation research, the results seen in adults were not consistent with previous findings. The inclusion of only two irrelevant actions, and the ease of interpretation of the task affordances (particularly for adult participants) may not have been enough to elicit overimitation. This is of particular importance if the function of overimitative behaviour is to allow humans to better understand a novel task or object (McGuigan & Whiten, 2009). Therefore, an extra action should be included in experiment 2. Another way to increase task difficulty further may be to use a different task altogether as a direct comparison to the puzzle box, which was potentially too easy for participants, particularly adults. Tasks involving tools are also frequently used in overimitation research (DiYanni & Kelemen, 2008; McGuigan et al., 2007b; Whiten et al., 2009). However, less research has involved tool building tasks, which involve the creation of novel, functional objects (Hernik & Csibra, 2009), and previous research has determined that young children are especially poor at tool innovation (Nielsen, Kapitány, et al., 2014). There is also a suggestion that tool manufacture tasks can further investigate the underlying mechanisms behind cumulative culture (Legare & Nielsen, 2015a). In order to fully understand the range of overimitative and emulative behaviours displayed by children, it would be beneficial to also include a tool building task to allow for comparison between other commonly used tasks.

Developmental factors

An additional potential explanation for the differences in behaviour seen in child and adult participants within both tasks, as well as the eye tracking data, may be underlying developmental factors which have the potential to impact social learning, such as memory and theory of mind. Deferred imitation is commonly used in children as an assessment of their episodic memory ability (Bauer & Kleinknecht, 2002), suggesting a link between the ability to imitate or overimitate and the capability for episodic memory. One example deferred imitation task requires infants to perform a sequence of actions on a puppet that is presented to them following a demonstration (Taylor & Herbert, 2014). Although 3-year-olds show some rudimentary capability for episodic memory (Mullally & Maguire, 2014) research suggests that it does not fully develop until around age 5 or 6 (Perner & Ruffman, 1995; Scarf et al., 2013; Terrace & Metcalfe, 2005). It would be beneficial to investigate the link between developing memory and overimitation, to determine whether reduced overimitation may have been seen in the board task as children were unable to remember the full sequence of actions. Additionally, as previously identified, theory of mind may be a factor in overimitation (Nielsen, Moore, & Mohamedally, 2012), particularly if overimitation is socially motivated (Over & Carpenter, 2012). In order to investigate this relationship further, a theory of mind measure should be used with participants.

Investigating a wider age range

The expected differences between adults and children were observed; adults appeared to be better at focusing their attention towards relevant stimuli. However, there is little research investigating when these changes occur, and when both gaze patterns and imitative behaviour become more adult-like. It would therefore be beneficial to investigate copying behaviours throughout early and middle childhood, as well as in adolescence and adulthood. Since the heatmaps from both the board and box tasks appear to show quite a drastic change in gaze patterns between child and adult participants, it would also be beneficial to use eye tracking in future studies to investigate this further.

Individual differences

Not all adult participants in the video control group behaved in the same way; some chose to overimitate, whilst others did not. The video control participants were not given any instructions other than “now it’s your turn”, and so should not have been influenced to behave in a particular way, unlike those in the copying and goal-focused groups. It is therefore unclear as to what drove adult participants to behave in a particular way when completing the tasks. To date, there has been little research investigating the role of personality and individual differences in overimitative behaviour, and therefore including a personality measure may provide additional information about why some people choose to copy where others do not.

Chapter 3: Experiment 2

3.1. Introduction

The purpose of experiment one was to investigate copying behaviours in 3-5-year-old children and adults, with the primary aim being to induce particular copying behaviours in participants, whilst conducting eye tracking during task demonstrations, in order to determine whether gaze patterns displayed by high and low fidelity copiers could be differentiated in any way.

In addition to this, two tasks were used in order to determine whether differences in copying behaviours were context sensitive. Participants' behaviour on both of these tasks was assessed using a quotient score, which indicated how faithfully they had copied the demonstrator. A score closer to 0 indicated a lower fidelity, more emulative copying style, whereas a score closer to 1 indicated a higher fidelity, more overimitative copying style.

Attempts to induce variations in these copying behaviours were unsuccessful across both tasks for child participants, and were only partially successful in adult participants for the board task, in which the majority of adults instructed to copy the demonstrator were found to achieve a higher quotient score in comparison to adult participants in the remaining two conditions. In the video control condition, in which participants viewed the demonstration but were given no instructions, participants' behaviour was observed to be similar to those in the goal focused group. Adult behaviour did not indicate the high levels of overimitation that have been observed in previous studies (Flynn & Smith, 2012; McGuigan et al., 2011b). Interestingly, all participants were able to successfully complete the tasks, potentially suggesting that they were too easy. This is particularly the case with the board task, as it is possible to determine how the pieces fit together without viewing the demonstration video first.

Results obtained from the eye tracking data did identify clear differences between children and adults, particularly in their ability to attend to relevant stimuli. In both box and board demonstration videos, children spent the majority of the time attending to the background (i.e. irrelevant areas of the video), whereas adults attended primarily to the task. These results were in line with previous findings, which suggest that attention and the ability to identify relevant stimuli increase with age (Courage, Reynolds, & Richards, 2006; Hagen & Hale, 1973; Plude & Doussard-Roosevelt, 1989). Interestingly, although clear instructions were provided to participants in the copying and goal-focused groups, no clear differences between gaze patterns obtained from participants in these conditions were observed, potentially suggesting that there are additional factors at work, which drive participants' copying behaviours. Some of these potential developmental factors are discussed in greater detail below, and warrant further investigation.

A number of limitations relating to experiment one were identified:

The number of irrelevant actions included in the puzzle box task was too low to fully assess overimitative behaviour, and the puzzle board task, which was intended to have a less clearly defined goal, was too limited in terms of the possible number/sequence of actions involved.

Calibration levels for younger participants were low, which led to large numbers of participants being removed from analysis and reduced the chance for comparison between 3- and 5–year-old participants. Additionally, the eye tracker used lacked precision in some cases, and so finer differences between participant groups may have been missed as a result of this.

No measures of developmental ability (memory, theory of mind) were collected for participants, and it is suggested that these may influence copying ability. The development of these abilities around the age of 4 or 5 (Perner & Ruffman, 1995; Perner & Lang, 1999) may help to explain the increase in overimitative behaviour observed in the literature in participants around this age.

Little is known about changes in copying behaviours during middle childhood and adolescence, and further research is required with participants in these age groups to determine whether a developmental trajectory is present.

Individual differences such as temperament were also not accounted for, and may influence participants' use of particular copying strategies, with research indicating that extroverts may be more likely to copy actions from a demonstrator in comparison to introverts (Hilbrink, Sakkalou, Ellis-Davies, Fowler, & Gattis, 2013).

The purpose of experiment 2 is therefore to address these limitations, and to expand on experiment one by again using eye tracking during demonstration videos, but with additional tasks and a wider age range, in order to see how copying behaviours and potential gaze patterns change with age and whether these behaviours are context-sensitive and task-dependent.

Within the social learning literature, there are a number of differing explanations for shifts in copying fidelity, particularly overimitative behaviour, in both children and adults. One potential issue is that these theories are often discussed in isolation and are rarely investigated simultaneously. However, it is likely that overimitation does not occur due to one singular underlying influence, but instead as a result of multiple processes or abilities. For this reason, one purpose of this study is to attempt to investigate a number of these underlying differences concurrently, in order to further understand their relationship to high and low copying fidelity within the same participant sample.

3.1.1. Developmental differences

Learning and memory are often broadly discussed as two related concepts in the literature, wherein memory is a requirement for successful learning to take place (Fogarty, Rendell, & Laland, 2012). Despite this, there remains a paucity of research with reference to the relationship between social learning strategies, specifically copying behaviours, and episodic memory. Consistent with the idea that memory is a prerequisite for learning, it can be suggested that the ability to learn from others must require the ability to store this information and subsequently retrieve it at a later date. This is of particular relevance in the developmental literature, as between the ages of 3 and 5 children demonstrate a remarkable change in their ability to perform particular tasks.

Deferred imitation is a commonly used measure of episodic memory in infants and young children, in which the participant is shown a sequence of actions and is required to imitate them following a delay (Meltzoff, 1985). 18 and 24 month old infants have been shown to be capable of replicating actions shown to them 24 hours previously, suggesting that they are capable of episodic memory in some form (Barr et al., 1996). However, a distinction has been made between proto-episodic memory, in which an infant or young child is capable of “what-where-when” memory, such as that required to imitate a sequence of actions, in comparison to autonoesis, in which children can remember having actually experienced previous events (Burns et al., 2014). It has been suggested that children are not capable of autonoetic memory, and therefore “true” episodic memory, until around the age of 4 or 5 (Perner & Ruffman, 1995). One issue identified with this focus on deferred imitation as a measurement of memory, however, is that it does not allow a full investigation of emulative behaviour, as emulation is seen as failure to complete the task, and for this reason, deferred imitation studies are designed in order to avoid emulative behaviour, rather than to measure it (Bauer & Kleinknecht, 2002).

If episodic memory is truly required for overimitation, then this may explain why differences in copying behaviour are observed between 3 and 5 year old children, where 3 year olds have been found to emulate more than 5 year olds (Dickerson et al., 2008; Flynn & Whiten, 2008; McGuigan et al., 2007b). It could therefore be suggested that 3 year olds use an emulative strategy not by choice, but simply because they cannot remember everything that has been shown to them during the demonstration. It is therefore expected that participants above the age of 5 will perform better on a memory task than those aged 3 and 4, and a higher memory score will also be related to increased imitation of actions. However, the methods used to assess memory within this study cannot truly be described as a measure of episodic memory, due to the length of time between the demonstration videos and the memory tests being administered, which was anywhere from 10 to 40 minutes, depending on task order and memory test order. Traditionally a longer delay, such as 24 hours, would be used in an episodic memory assessment, but this was not possible without asking for participants to return the following day. Nonetheless, the tasks used were designed to assess two important components of participants' memory of the task itself (content), and the order in which the actions occurred (sequence), in the hope that even a basic understanding of participants' ability to remember actions observed and the order in which they occurred would provide more information regarding their task performance.

Theory of mind, or being able to understand the motivations of others (Frith & Frith, 2005), has also been identified as important in the literature discussing overimitative behaviour. In particular, a disparity has been observed when examining overimitation in typically developing children and those with Autism Spectrum Conditions (ASC), in which typically developing children overimitate significantly more than ASC children (Marsh, Pearson, Ropar, & Hamilton, 2013). A underlying factor here may be theory of mind; ASC children are often found to perform poorly in theory of mind tests (Baron-Cohen, Leslie, & Frith, 1985). It may be the case that ASC children are therefore be less likely to imitate for social reasons such as affiliation with the demonstrator (Spengler et al., 2010). As theory of mind does not develop until around the age of 4 or 5 (Perner & Lang, 1999), when children begin to be able to pass false belief tasks, then it would be expected that if overimitation serves a social purpose, 3 year olds should overimitate less as they lack understanding of the motivations of the demonstrator. Indeed, theory of mind continues to develop through childhood and into adolescence, (Wellman, 2002). Additionally, it has been suggested that autonoetic memory relies on the development of theory of mind (Perner & Ruffman, 1995), providing an additional link between these abilities, and further supporting the suggestion that younger children do not consciously choose to emulate.

Since these developmental changes occur in young children, the majority of the literature discussing overimitation tends to focus on the differences immediately before and after these abilities develop, as well as making comparisons with adults. However, there is little focus on overimitative behaviour in late childhood and adolescence in typically developing children. This is particularly important, since underlying developmental features such as improvements in episodic memory continue to increase throughout childhood due to the ongoing development of the hippocampus and the prefrontal cortex during middle childhood and into adolescence (Ghetti & Bunge, 2012; Vetter, Leipold, Kliegel, Phillips, & Altgassen, 2013). In fact, children only begin to display adult-like patterns of activation in the hippocampus during an episodic memory task around the age of 14 (Ghetti, DeMaster, Yonelinas, & Bunge, 2010), and Vetter et al (2013) determined that social cognition continues to develop throughout adolescence and into early adulthood.

Whilst social cognition has been widely investigated in teenagers in a broader context, it is often related to the development of antisocial behaviours, with the suggestion that teenagers are susceptible to social pressure from others, particularly peers (Akers & Lee, 1996; Krohn, Skinner, Massey, & Akers, 1985). If overimitation is a purely social process relating to a need to display affiliation with a demonstrator, or a chance for participants to show the experimenter that they are capable of performing the task “correctly” (in other words, in the same way as the demonstrator), then it would be expected that teenage participants may demonstrate more overimitative behaviour if they are more susceptible to social pressure from others. It would therefore be interesting to investigate the difference between teenage participants, who have been shown to display increased conformity in comparison to other age groups (Lashbrook, 2000). If this is the case, then we would expect teenagers to overimitate more than younger participants, and more than adults, as although younger children do display conformity (Haun & Tomasello, 2011; Morgan, Laland, & Harris, 2015), it has been suggested to peak in adolescence (Coates, Petersen, & Perry, 1982; Costanzo & Shaw, 1966). As previously stated, social cognition continues to develop during adolescence (Vetter et al., 2013), and so it is expected that differences will also be observed between adolescents and adults, with adults potentially being less likely to be susceptible to the need to conform to a particular behaviour or method of task completion.

3.1.2. Task differences and features of the demonstration

Task Type

The nature of the task itself has also been shown to affect copying behaviours. Flynn & Smith (2012) determined that when 5 year old children viewed a demonstration involving a transparent puzzle box, they were less likely to overimitate in comparison to when an opaque box was used, possibly because participants were able to interpret the causality of the actions performed. However, Horner & Whiten (2005) observed that 3- to 5-year-old children overimitated actions in both opaque and transparent conditions, although it was suggested that this may have been due to the complexity of the task.

As discussed in chapter 1, a number of different tasks are used in studies investigating overimitation. However, whilst there is research involving simple tools (for example, using a stick to retrieve an object from inside a puzzle box (Horner & Whiten, 2005)), less attention has been paid to how much children or adults overimitate during tool innovation tasks, which tend to be far more complex in nature. Tools are of particular importance in human culture, and the development of complex tools is reliant on cumulative culture, in which individuals are able to build on the work of others that have gone before them. For this reason, high-fidelity transmission of information, such as overimitation, is vital in the development of new tools and technology, in order that the previous iteration of a tool can be replicated and subsequently improved (Lewis & Laland, 2012).

In tool building tasks, participants are provided with components and must use them to produce a tool, usually in order to retrieve a reward (Beck, Apperly, Chappell, Guthrie, & Cutting, 2011). Tool innovation tasks are often described as “ill-structured” due to the wide variety of potential actions that can be completed, unlike in a puzzle box task, which has a clearly limited number of actions that depend on its function (Cutting et al., 2014). Therefore, in these tasks, overimitating serves to benefit participants (Beck et al., 2011), who will be able to complete the task much more quickly and efficiently by copying someone else, even if this means including some irrelevant actions. Additionally, participants have many more potential solutions available to them, as they can use the pieces provided in multiple different ways in order to construct a hook that will allow them to retrieve the reward. Unlike the box task, which has a very limited number of potential actions necessary to achieve the goal, the hook task has a much wider variety of potential strategies that can be used to achieve the solution. It would therefore be expected that higher levels of overimitation might be observed in this task in particular, as it will allow participants to complete the task more efficiently than trying to develop their own strategy. This is particularly true for young children, who are generally poor at tool innovation (Beck et al., 2011; Cutting, Apperly, & Beck, 2011; Sheridan, Konopasky, Kirkwood, & Defeyter, 2016). This finding has also been found across a number of cultures, suggesting that it is robust (Nielsen, Tomaselli, Mushin, & Whiten, 2014).
Determining what children and adults focus on when viewing a demonstration of a tool manufacture task may provide a better understanding as to what adults consider to be the most important properties of tools, why children find these tasks so difficult, and what can be done to better scaffold their learning during a demonstration or explanation. Additionally, a further benefit of using tool manufacture tasks may be the increased difficulty itself – the board and box tasks used in experiment one were found to be too easy, particularly for adult participants. The use of a more complex task may reduce the ceiling effect observed previously, and should allow for a much wider range of responses in order to fully investigate a variety of copying behaviours in all participants. Little is known about the development of tool making abilities, and the use of the same task across multiple age groups should provide further understanding into how this skill develops throughout childhood and into early adulthood.

In this task, copying behaviours can be investigated in a number of ways. Two irrelevant actions have been included in the demonstration video, but the pieces used by participants, the overall shape of the hook tool produced, and the order in which the hook is built can also be used as an assessment of copying fidelity. Hernik & Csibra (2009) determined that even very young children pay attention to the features of a tool that may contribute to a goal, so it may be more likely for children in the youngest participant groups to copy the shape of the hook in comparison to the order in which it is built. Previous research has determined that younger children are more likely to copy the hierarchical structure of a task, such as the general shape and the order in which it is built, rather than attending to specific action detail, whilst older children are likely to copy more intricate details (Flynn & Whiten, 2008b). Although this study involved a multi-step puzzle box, it may be possible that similar results to these will be observed in this experiment, and therefore precision of imitation will increase with age. One of the irrelevant actions (handle) has been designed to appear functional, whereas the other irrelevant action (tap) is more obviously unnecessary for task completion. It could therefore be expected that the handle action will be more frequently imitated than the tap action, and the tap action may be most commonly imitated by the youngest participants who do not fully understand how the task functions. Analysis of eye tracking data is more difficult in this task than in the puzzle box or colouring tasks, as the pieces required to build the hook are very small and so broad areas of interest cannot be drawn. For this reason, fixations at the plastic pieces, the tube and the hook itself will all be coded as belonging to the task area of interest.

In experiment one, the board task was intended to act as a more arbitrary task in comparison to the more functional, goal-driven box task. As discussed in chapter 2, a number of limitations were identified in relation to the board task that indicated that an alternative task was required in order to fully assess participants' copying behaviours when completing a more abstract activity with a less clearly defined goal. The main limitation identified was that the board task still had a clear end goal and a very limited number of potential actions. A colouring task, using shapes as opposed to more meaningful images, was therefore identified as the replacement for the board task from experiment one. An "irrelevant" action in the form of drawing a stick figure was also included, although this was far less obviously irrelevant than those included in the previous tasks. As with the hook and box puzzles, this new task had to be appropriate for use with all participants, taking complexity and fine motor skills into consideration. For this reason, the shapes and the irrelevant drawing were kept as simple as possible, and thick marker pens were provided to all participants as these were the easiest for use with young children who may have lacked the fine motor skills required to hold smaller pencils or crayons. The task involved colouring in three shapes: a triangle, a rectangle, and an oval, which were pre-printed onto an A4 sheet of paper, and drawing the stick figure as described previously. The rectangle was placed inside the oval, whilst the triangle was separate. In the demonstration video, the paper was attached to an upright board using two triangular magnets (see figure 16), with one in each of the bottom corners. These magnets were not present during the testing phase. The upright board was used in order to ensure that the participant could see all areas of the colouring sheet easily during the demonstration. However, when participants completed the task, the colouring sheet was placed flat on the table in front of them, without the use of the board, as it was felt that colouring on an upright board may be too difficult for younger participants, who were still developing their fine motor skills. As in the previous tasks, participants were not provided with any specific instructions on how to complete the task, but instead were told “now it’s your turn”.

In contrast to puzzles used in previous chapters, this task has no obvious reward (e.g. the plastic egg in the box task and the sticker in the hook task), therefore meaning that it also has no well-defined goal. As previously mentioned, research by Carpenter, Call, & Tomasello (2005) indicates that when there is no clear goal in a task, copying the actions themselves becomes the goal, meaning that overimitation actually increases. However, colouring in is likely to be a familiar activity for even the youngest participants, and it has been suggested that familiarity with a task reduces overimitation of redundant actions as participants have more of an awareness of how the task normally functions (Gergely, 2006). Although the stick figure is intended to be irrelevant, however, it is not obviously unnecessary, and so may just be interpreted as part of the task by participants. Although this makes it difficult to predict participants’ behaviour, this task may provide more insight into the underlying social motivations for copying, due to the fact that participants do not risk losing out on a reward by using their own method to complete the task. Therefore, it is likely that if participants in the 7-15 and adult age groups copy, they are doing so for social reasons rather than to use the best method to achieve a set goal. It is also possible to assess this by investigating both the amount of time spent looking at the demonstrator's face during the demonstration video, as well as whether participants look at the demonstrator's face prior to the beginning of the task. It is expected that participants who look more at the demonstrator's face are more likely to feel affiliation with the demonstrator, and are potentially therefore more likely to overimitate in order to appear similar to them. There is some support for this in research by Vivanti, Hocking, Fanning, & Dissanayake (2017), in which typically developing participants who had looked more at the demonstrator's face were subsequently more likely to copy irrelevant actions. As previously discussed, affiliation is considered to be a highly motivating factor in overimitation, with Watson-Jones, Legare, Whitehouse, & Clegg (2014) finding that 3-6 year olds were more likely to overimitate when primed with ostracism, meaning that they wished to display more affiliative behaviour in order to avoid being ostracised from a social group.

One potential issue with this task is that the demonstration video is almost two minutes long, meaning that younger participants may struggle to sustain their attention for this length of time (Hagen & Hale, 1973). An additional difference between this task and the box and hook tasks is that the results of the actions performed by the demonstrator (i.e. the coloured shapes) remain on screen once she has completed them. In order to prevent this, there were two possible options; once a shape had been coloured, the camera could have zoomed in to focus on the next shape to be coloured, therefore only showing one shape at any one time, or the colour information could have been removed from shapes that had already been coloured, (they would appear "empty" once again). Neither of these options were considered to be appropriate, as they may have been too confusing for the youngest participants. Zooming in onto one shape would have made the demonstration drastically different from the box and hook demonstrations, in which the tasks (puzzle box and plastic tube with bucket) are visible at all times. Additionally, zooming in on one particular aspect of the board would mean that the demonstrator was no longer visible, therefore limiting information available to the participant, and potentially skewing the results, as participants would have less potential options in terms of the aspects of the demonstration they were able to see. Removing the colour from the shapes after they had been coloured may have been extremely confusing to younger participants, and may have changed their understanding of the overall goal of the puzzle.

The fact that participants are able to view the intended colour of some of the shapes for an extended period of time may mean that younger participants may be more likely to match the end result in this task than in the other two, since they are not required to remember information after only having seen it for a very brief period of time as in the box and hook tasks. The fact that the colours remain inside the shapes on screen provides an opportunity for rehearsal of this information prior to completing the task. This also means that participants do not need to attend for the full duration of the video to determine which colour the shapes should be; they could potentially only attend to the shapes towards the end of the video and still achieve the same outcome as a participant who watched consistently throughout. One aspect of the task which may highlight more differences between age groups is the order in which the shapes are filled in; the rectangle is coloured first, followed by the triangle, and then by the oval. In order to imitate this order, participants will need to attend more to the demonstration. Therefore, it is expected that participants who spend longer attending to the task will be more likely to imitate more precisely, copying the order of task completion as well as the colours used. Additionally, it is expected that participants who attend more to the demonstrator will be more likely to achieve higher scores on either the content or sequence memory tasks.

Functional and arbitrary tasks

When a more arbitrary, less goal-driven task was used, typically developing children were found to increase their focus towards the actions involved, as opposed to during a demonstration with a clear goal, in which their main focus was the goal itself (Wild et al., 2012). In this study, children viewed demonstrations of both goal-directed and goalless hand movements, in order to investigate the effect of the presence of a goal on participants' imitation. In the goalless condition, typically developing participants increased their attention to the movement itself, and imitated changes in movement kinematics, in contrast to the goal-directed condition. Additionally, research with adult participants has suggested that when there is no clear goal relating to a task, imitation of the actions themselves is seen as the overall goal, potentially meaning that increased overimitation is likely to be observed in these kinds of tasks (Schachner & Carey, 2013). This may be due to the fact that increased imitation tends to be observed in uncertain situations (Thelen et al., 1979).

Comparison of behaviour between tasks

As varied behaviour is observed as a result of differences in task type, it would be beneficial to include multiple task types in one study, so comparisons can be made between a child's behaviour when completing both arbitrary and goal-driven tasks. Research suggests that children are capable of flexibly overimitating, depending on context (Keupp, Behne, Zachow, Kasbohm, & Rakoczy, 2015; Legare et al., 2015). It could therefore also be expected that children display flexible overimitation during the completion of tasks with more or less obvious goal states. However, this may also be contingent on additional developing abilities, as discussed previously. It is expected that children, like adults, will display increased overimitation in the task with the least obvious goal state, as copying the actions themselves may become the goal when an obvious goal (such as retrieving a reward) is not apparent.

Consistency between demonstrations

As discussed in chapter 2, care must be taken to ensure consistency of demonstrations shown to participants, as the behaviour of a demonstrator towards a child during the demonstration of a task can be very influential in how they later choose to behave. For this reason, and to enable eye tracking to be conducted, video demonstrations will be used, with the same demonstrator performing multiple tasks in order to ensure consistency. Of particular importance may be whether the demonstrator faces the camera both at the beginning and end of the demonstration, as Nielsen (2006) describes the way in which ostensive cues inform children that they must pay attention as they are about to be taught something new. Eye contact with a participant appears to be of particular importance in inducing imitation in typically developing individuals (Gergely, 2006; Hoehl et al., 2014; Nielsen, 2006; Wang et al., 2011), and for this reason, the demonstrator will face the camera at the beginning of each demonstration video.

3.1.3. Aims of this study

In experiment one, differences in eye tracking patterns and in task performance were identified between children and adults, with adults focusing their attention more on relevant aspects of the box demonstration. However, in this initial experiment, an attempt was made to induce behaviour by giving specific instructions to encourage overimitation or emulation. Although this did not appear to be successful with child participants in terms of behavioural outcomes, it is still possible that task performance was affected in some way by the instructions given, even if this was not obviously apparent in the results obtained. Therefore, in this experiment, minimal instructions will be given prior to task completion in order to encourage participants to behave in a more natural manner. Participants will also undergo eye tracking, as in experiment one, to determine whether gaze patterns change with age in terms of attention directed towards relevant aspects of the demonstration, as well as whether differences can be observed between participants who overimitate and those who emulate. Due to the poor calibration rates obtained with the Eye Tribe tracker, an Eyelink 1000 tracker was used instead. This is a desktop tracker that uses a target sticker, which is placed on the participant's forehead, and therefore allows for more movement than the Eye Tribe tracker. It was hoped that this would increase calibration success, particularly in younger participants who may find it more difficult to remain still.

As previously described, differences in copying behaviours have been observed in both children and adults, depending on the type of task used. For this reason, three tasks were used: the puzzle box task used in experiment one, with an additional irrelevant action added, a tool building task, and a more arbitrary colouring in task. All participants completed all three tasks, meaning that within subjects comparisons could be made in order to investigate whether participants overimitate flexibly. It was expected that participants would overimitate most on the arbitrary colouring task, as there is no clear goal, and therefore participants could potentially interpret the actions involved in task completion themselves as the goal (Schachner & Carey, 2013). However, this could potentially be complicated by the fact that there is a defined end state. Development of an entirely abstract, goalless task was not possible, primarily due to the wide age range of participants to be tested, and the difficulty this may present. This was particularly true with the youngest age groups, as children struggle with abstract concepts even up to the age of 12 (Setti & Maurizzi, 2014). Overimitation should also be high in the tool building task; as previously stated, tool manufacture is an ill-structured problem (Cutting et al., 2014), and therefore participants may choose to copy the demonstrator more closely rather than having to find a solution of their own, particularly as they have seen the demonstrator successfully completing the task. With young children, this should even be the case for causally irrelevant actions, as they may find the task too perceptually opaque

(Hernik & Csibra, 2009). However, it could also be expected that, as found by Flynn & Whiten (2008), younger children are more likely to imitate the overall structure of the hook tool (general shape, order in which it is built), whereas older children and adults may be more likely to copy more intricate details, such as the exact pieces used by the demonstrator. The puzzle box task has the potential to result in the most variable amount of overimitation, as most of the irrelevant actions will be obviously unrelated to task completion, especially for older participants. Additionally, as the inner workings of the puzzle box are visible, participants had more causal information as to how the levers operate, and what effects the irrelevant actions have, and therefore may have chosen to omit actions that very clearly serve no purpose (McGuigan et al., 2007b).

Copying behaviours in a wider range of age groups were investigated in this experiment, in order to determine whether a developmental trajectory in copying fidelity from childhood to adulthood can be observed. Participants aged 3 to 21 years old took part, in order to allow for comparisons between age groups. Additional measures of developing abilities such as memory and theory of mind were administered, in order to determine whether there is a relationship between these abilities and changes in copying behaviour. It was expected that 3 and 4 year olds will perform poorly in these tasks, and would also be more likely to use an emulative strategy than older participants (Dickerson et al., 2008; Flynn & Whiten, 2008). It was also expected that overimitation will rise with age (Marsh, Ropar, Hamilton, Erdohegyi, & Csibra, 2014), and peak in adolescence, in line with conformity (Coates et al., 1982). There is some conflict within the literature about the ways in which adults overimitate, with some suggestion that adult participants actually overimitate equally as much as, or more than children (McGuigan et al., 2011b; Whiten et al., 2016). However, this behaviour has been found to be variable, dependent on the social context of the task, suggesting that overimitation in adults may be a result of conformity (McGuigan, Gladstone, Cook, Macris, & Keil, 2012). If this is the case, then overimitative behaviour in adults may still be present, but should be less than that seen in adolescent participants.

As ostensive cues have been highlighted to be important in eliciting overimitation and indicating that the transmission of information is about to begin (Gergely, 2006; Wang et al., 2011), the demonstrator will make eye contact with the camera at the beginning and end of the demonstration video. However, there will be no verbal cues given, as these may be too confusing for younger participants who have less developed language abilities than older children and adults. Additionally, the use of verbal cues has not been found to change participants' overimitative behaviour (Hoehl et al., 2014b). If there is a social component to participants' imitation of irrelevant actions, then it would be expected that participants who look at the demonstrator's face at the beginning of the video, during the transmission of the ostensive cue (i.e. eye contact), might feel increased affiliation with the demonstrator and therefore may be more likely to copy their behaviour (Vivanti et al., 2017). Previous eye tracking research has determined that attention to visual stimuli increases with age (Anderson & Levin, 1976; Roebers et al., 2010), and therefore it is likely that participants' overall attention to the demonstration video will increase with age. Additionally, the ability to attend to relevant stimuli increases with age (Hagen & Hale, 1973), so it would be expected that older participants will spend more time attending to the task as opposed to less relevant aspects of the demonstration video. If imitation occurs as a result of a desire to conform and/or the wish to demonstrate affiliation with an individual, then it would also be expected that participants who look at the demonstrator's face at the start of the video are more likely to later imitate actions performed by them.

3.2. Method (all tasks)

3.1.1. Participants

125 typically developing children aged 3-17 were recruited from the local area via social media and word of mouth. 25 adult participants, aged 18-21, were also recruited. All adult participants were undergraduate students studying psychology at Newcastle University. Students received course credit for their participation. Child participants received a certificate and small gift for their participation. Prior to completion of the research, written consent was provided by parents or participants themselves as appropriate; for children under 13, verbal consent was obtained. For participants aged 13-17, written consent was obtained from the participant alongside parental consent. For the purposes of initial analysis, participants were divided into the age groups indicated in table 3.1:

Table 3.1: Mean ages for participants in each of the assigned age groups

Age Group	N	Mean	Std. Deviation
3-5	52	3.90	0.823
6-8	31	7.06	0.892
9-11	22	9.73	0.883
12-17	20	12.75	0.851
18+	25	18.92	0.812
Total	150	9.09	5.386

All participants completed all three tasks (box, hook and colouring) during their visit, although eye tracking data was not obtained for all participants in all tasks; this is described further in each of the task subsections. Unless otherwise stated, behavioural results were calculated for all participants, whether eye tracking was successfully obtained or not.

3.1.2. Materials and apparatus

Eye Tracking

An Eyelink 1000 desktop eye tracker (SR Research) with a 16mm lens was used to collect monocular gaze data at a sampling rate of 500Hz, with 0.5º average accuracy. This tracker does not require use of a chin rest, but instead a small target sticker is placed on the participant's forehead to allow their eye to be tracked. This means that participants are able to sit more comfortably whilst watching demonstration videos. The target sticker also allows for some movement, which is particularly important with younger participants who may find it extremely difficult to sit still whilst watching the demonstration videos, and who may have found the use of a chin rest either uncomfortable or intimidating, especially in an unfamiliar environment. All participants completed a five point calibration prior to viewing the demonstration videos for each task, and any participants who failed the calibration process were excluded from analysis for that particular task. A five point calibration was used as opposed to the standard nine point calibration as it is more suitable for younger participants

(Gredebäck, Johnson, & von Hofsten, 2010) and therefore was more likely to result in successful calibration for these age groups. OpenSesame software (Mathôt, Schreij, & Theeuwes, 2012) was used to design and display the task. Participants underwent eye tracking during demonstration videos only, and were free to move and interact with each of the tasks in between viewing the demonstrations. The same location and demonstrator was used for each of the demonstration videos in order to ensure as much consistency between tasks as possible.

Memory

Participants were randomly allocated to one of two memory tasks: a sequencing task and a content task. Memory tests were completed at the end of the testing session, after completion of all three tasks, meaning that there could potentially be a delay of around 30 minutes between viewing the demonstration video for one of the particular tasks and completing the memory test relating to it. Additionally, memory tests were completed for all three tasks (box, hook, colouring) at once, as it was felt that asking participants to complete the appropriate memory test immediately following its corresponding puzzle task might affect performance on subsequent tasks. Participants would be aware that they were going to be asked to remember specific aspects of the demonstration that they had viewed. This could have the potential to change participants' interpretation of the task goal, so that rather than retrieving a reward (for example), they may be artificially influenced to consider the goal to be to remember and then reproduce the actions observed in the demonstration.

In the sequencing task, participants were presented with four still images, taken from the demonstration video, and were asked to place these in the correct order. In the content task, participants were presented with two images that they had previously seen in the demonstration video, and two false images that were not previously included. Participants were asked to identify which images they had seen in the demonstration video, and which they had not. Images for each of the tasks will be described further in the subsequent subsections relating to each task.

Theory of Mind

Theory of mind was tested in all participants using a false belief task, adapted from Astington & Jenkins (1999). Although this task is different to a more traditional theory of mind task, such as the Sally Ann task (Baron-Cohen et al., 1985), it was deemed more appropriate for use with the age range of participants used in this experiment, including with adults, who might be reluctant to engage in tasks that appear to be more childish. In this task, participants were shown the image in figure 3.1:

Figure 3.1: False belief task image shown to participants

Participants were informed that the little girl in the picture had a pet cat who liked to hide from her. He could not be seen in the picture at the moment, because he had gone to hide already. The cat could either be hiding in the tree (at this point, the experimenter points to the tree), or in the shed (experimenter points to the shed). The participant was then asked where they thought the cat might be hiding. Participants could either select their answer by pointing at the picture of the shed or the tree, or by telling the experimenter which place they had chosen. Once a decision had been made, the experimenter then informed the participant that the little girl thought that the cat was hiding in the opposite place to that selected by the participant (i.e. if the participant had selected the tree, they were told that the little girl thought that the cat was in the shed). Participants were then asked where they thought the little girl would go to look first when she looked for the cat. If the participant chose the location where the girl thought the cat was hiding, they were described as passing the test. If they chose the same location that they had identified themselves, then they were described as failing.

3.1.3. Procedure

All testing was conducted in the eye tracking lab at Newcastle University. Parents remained in the room with their children at all times, although they were encouraged not to help their children with any of the tasks. Upon entering the room, the participant was informed that they would be watching some videos and completing some puzzles whilst their eyes were tracked by a camera. In the case of child participants, both the child and parent(s) or guardian(s) were shown around the eye tracking room, and the way in which the eye tracking camera worked was explained. Parents and children were given the opportunity to ask any questions before verbal consent was sought from the child (along with written consent where appropriate), and written consent was completed by the parent. In the case of adult participants, the same procedure was followed. However, adult participants completed their own temperament questionnaire prior to beginning calibration, whereas parents completed the appropriate temperament questionnaire on behalf of their child as their child was watching the demonstration videos and completing the tasks.

Prior to beginning the task, the participant was seated in front of a computer monitor, which was positioned at a distance of approximately 60cm. A dark curtain was drawn to the right of the participant, in order to prevent them from seeing any of the puzzle tasks whilst the demonstration videos were playing. The eye tracking computer was positioned to the left of the participant, but was behind a barrier, which was covered with a black cloth in order to reduce the distraction to the participant as much as possible. In order to increase pupil size and for ease of calibration, the lights in the room were dimmed, as long as the child remained comfortable. A target sticker was placed on the participant’s forehead, and a five point calibration was completed. Participants were advised to remain as still as possible whilst watching the videos. The overall procedure for the testing session is outlined in figure 3.2. Participants were free to withdraw at any time, meaning that not all participants completed all three tasks. At the end of the testing session, the purpose of the study was explained to participants and parents or guardians, and child participants received a certificate and a small toy for taking part. Adult participants received course credit in return for their participation.

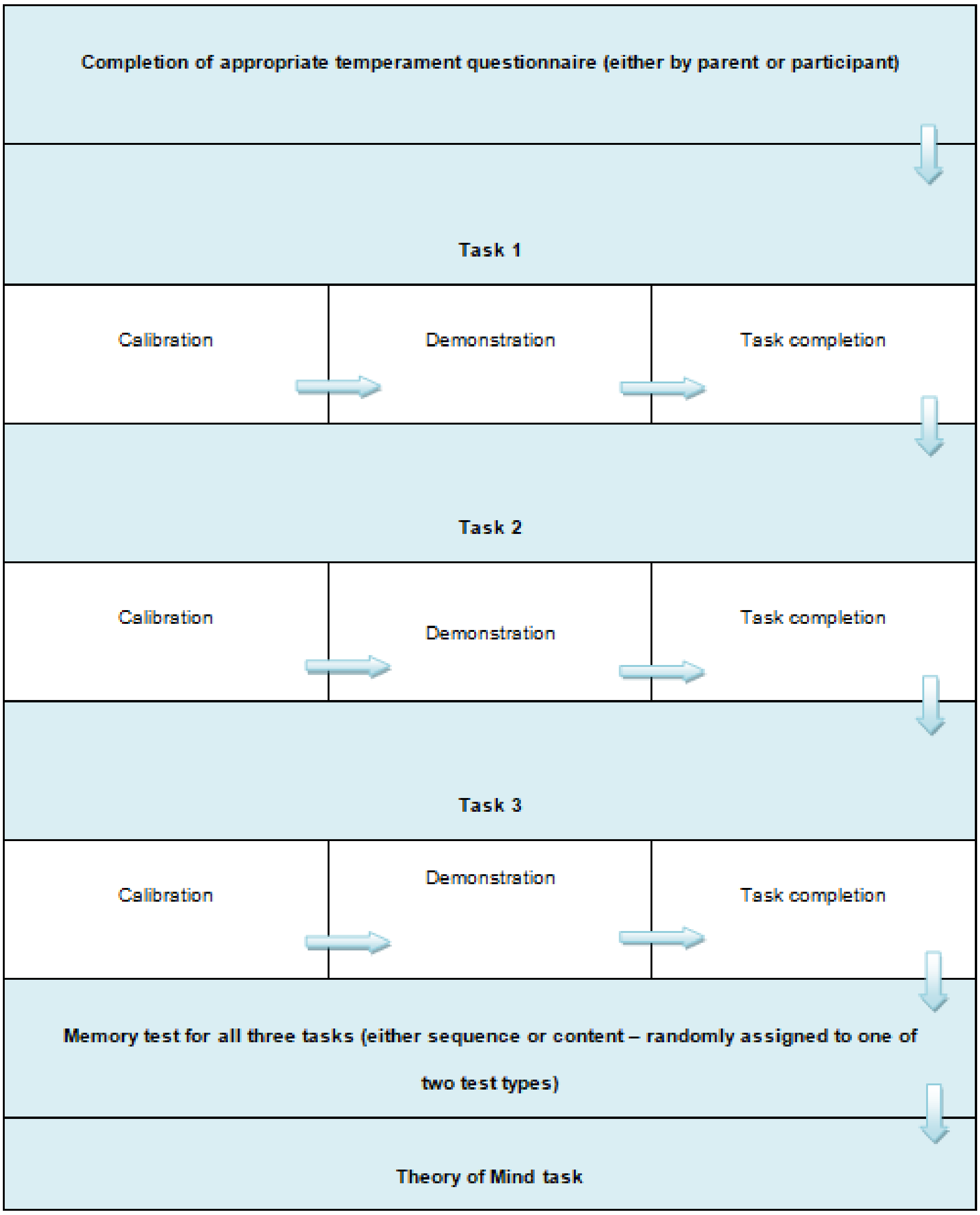


Figure 3.2: Testing procedure completed by all participants

3.1.4. Data Analysis

All data has been checked for homogeneity of variance. Where this assumption has been violated, non-parametric tests have been used instead.

Eye Tracking

Participants who failed the calibration procedure were excluded from analysis of eye tracking data. The demonstration videos were split into individual frames, and fixation data was extracted using SR Research Data Viewer. A MATLAB script was designed to display the video frames one by one, with any corresponding fixations overlaid onto them. As each fixation had a corresponding start and end time, the duration of fixations in each relevant area of interest was calculated per trial for each task. Areas of interest for each of the tasks were identified as the demonstrator's face, the task, and the background of the demonstration video.

Behavioural scoring

Scoring for individual tasks is described further in the subsections relating to each puzzle.

Memory and False Belief Scoring

In the sequence task, participants were given a score of 1 for each image that they placed in the correct order in the sequence, resulting in a maximum possible score of 4; e.g. images placed in the order ABCD would score 4, whereas images placed in the order ACBD would score 2. In the content task, participants were given a score of 1 for each image that they correctly identified as true or false, resulting in a maximum possible score of 4. In the false belief task, participants received a score of 1 if they passed, and 0 if they failed.

3.3. Method (Box Task)

3.3.1. Participants

Eye tracking data for this task was collected for 122 participants. For the purposes of initial analysis, participants were divided into the age groups described in table 3.3. However, analysis of behavioural data was conducted for all participants, even those for whom eye tracking data was not obtained.

Table 3.2: Mean ages for participants in each age group (box task)

Age Group	N	Mean	Std. Deviation
3-5	43	3.86	0.833
6-8	29	7.07	0.884
9-11	20	9.65	0.875
12-17	13	12.69	0.947
18+	17	18.92	0.812
Total	122	8.76	5.421

3.3.2. Materials and apparatus

The box task consisted of the same clear Perspex box as used in experiment one, measuring 20cm tall, 30cm wide and 16cm deep (see Figure 3.33.3). The box contained levers on the left and right hand sides which could be operated in any order to release a reward. Both the internal and external mechanisms of the task were visible to participants throughout the demonstration video. Three external levers operated the internal mechanisms of the box. On the left hand side, a wheel could be rotated either forwards or backwards to turn the platform containing the reward. On the lid of the box, a distracter lever could be moved up or down but this had no effect on the outcome of the task. On the right hand side of the box, a lever could be moved up or down, which moved the attached platform up and down. Upon moving the left and right levers, the reward would be tipped down a chute at the bottom of the box. This chute was designed to be small enough that participants could not use it to retrieve the reward without using the levers by using their hand to reach into the box. The reward consisted of a small plastic egg which could be swapped with the experimenter for a sticker. The lid of the box was held closed using a strong magnet, and a key was required to release the lid to reset the task. This key was not accessible by participants. The purpose of this was to prevent young participants from simply opening the box to retrieve the reward.

Figure 3.3: Box apparatus used in both demonstration video and task

Alterations from experiment one

There were only two irrelevant actions included in the box task in experiment one: lifting the irrelevant lever on the lid of the box, and knocking on the lid of the box. Whilst the lever action could be interpreted as functional, and therefore necessary for task completion, the knock action was quite obviously not functional. It may have been the case that these two actions that could be interpreted to be at two extremes, and for this reason, may not have allowed for a full interpretation of high or low fidelity copying behaviour. Therefore a third, more ambiguous action, the turn action, was added in order to investigate copying fidelity in more depth. This action is described in more detail below.

Eye Tracking

Following calibration, participants viewed a 16000ms demonstration video followed by a 3000ms fixation cross three times (see Figure 3.4).

Figure 3.4: Trial sequence shown to participants; participants viewed the same demonstration video (16000ms) three times, followed by a 3000ms fixation cross. The green fixation cross indicated that the trial had ended.

Demonstration video

A female demonstrator was used for this task. She faces the camera at the beginning of the video, in order to make eye contact with the participant to signal that the demonstration is about to take place. The demonstrator first tips the lever on the right hand side of the box to move the lower platform into a downwards position. This does not alter the position of the reward. She then performs three unnecessary actions. First, she knocks on the lid of the box.. Next, the demonstrator pulls the lever attached to the lid of the box, which does not have any function in solving the task. She then turns the box around 360º clockwise. Finally, the demonstrator performs a final necessary action, rotating the wheel on the left hand side of the box, tipping the platform that releases the prize, which falls onto the lower lever, and then downwards through the chute at the bottom of the box. The demonstrator picks up the egg, and then faces the camera again and smiles, whilst holding the reward (see Figure 3.5). As in experiment one, the irrelevant actions were included in between the two relevant actions due to the way that the box functions – the lever to release the reward must be operated last.

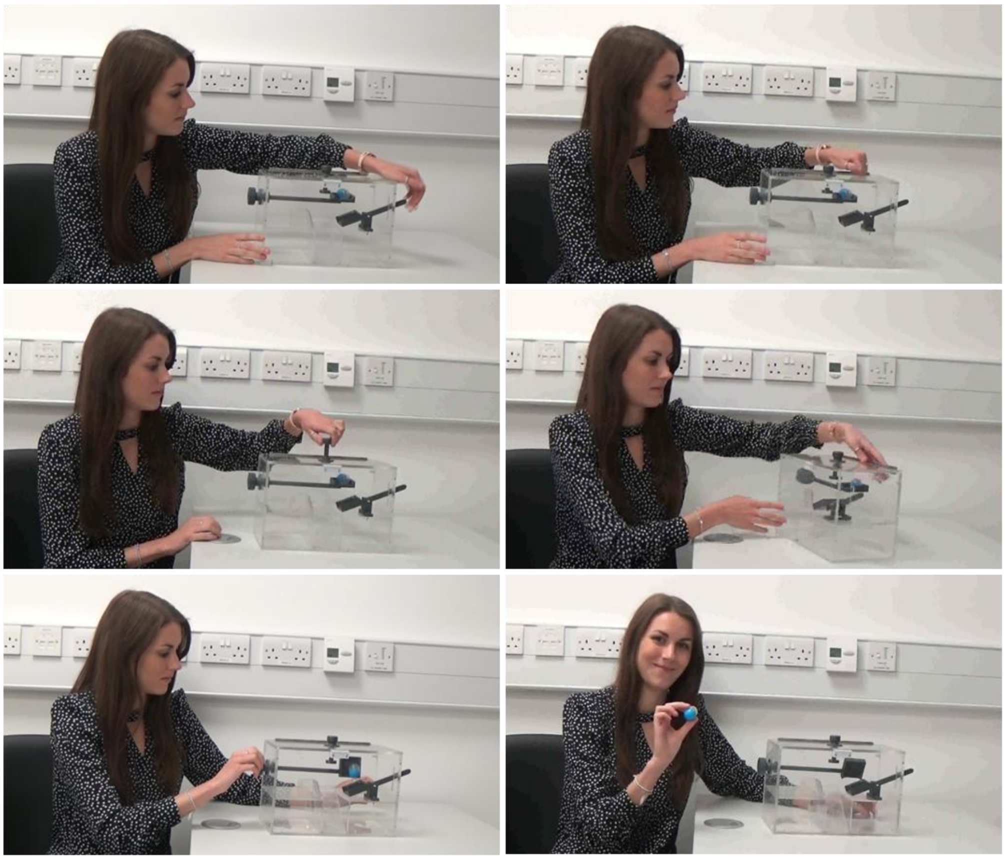


Figure 3.5: Actions performed during video demonstration; the demonstrator tips the right lever, knocks on the lid of the box, lifts the irrelevant lever, turns the box around clockwise, turns the left lever and retrieves the reward as it falls out of the chute at the bottom of the box.

Memory Tests

The true images used in the content memory task are shown in Figure 3.6, and the false images are shown in Figure 3.7. In this task, participants were required to indicate which of the images they had seen in the demonstration video, and which they had not.

Figure 3.6: True images shown to participants in content memory task. The demonstrator knocks on the lid of the box and lifts the lever on the lid of the box. Both of these actions were irrelevant to task completion.

Figure 3.7: False images shown to participants in content memory task. The demonstrator walks a toy hippo along the lid of the box, and lifts the box in the air to try to tip out the reward.

The images shown to participants who completed the sequence memory task are indicated in Figure 3.8. These images were shuffled, and participants were asked to place them back into the correct order.

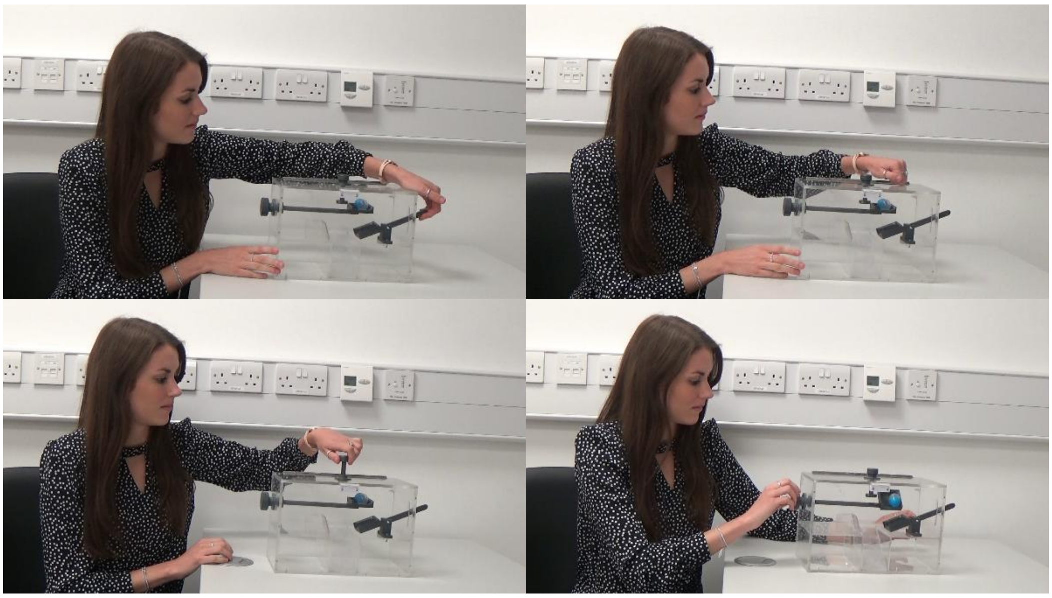


Figure 3.8: Images shown to participants during sequence memory task, presented in the correct sequence. The demonstrator tips the right lever, knocks on the lid of the box, lifts the irrelevant lever and then turns the left lever.

3.3.3. Data Analysis

Eye Tracking

As previously described, areas of interest for this task were identified as the demonstrator's face, the internal mechanisms (i.e. the levers) of the box, the empty space inside the box, and the demonstrator's hands.

Task Scoring

Participant data was scored depending on the number of irrelevant actions produced and the order in which the relevant actions were copied. Within this task, there were two possible irrelevant actions and two potential measures of the order of actions completed: the order in which participants completed the relevant actions (left lever, right lever) only, and the order in which they completed all actions. Participants were awarded a score of 1 for each of these elements that they completed, meaning that the maximum possible score for this task was 5. A quotient score was then calculated for each participant, by taking their actual score on the task and dividing it by the maximum possible score. A quotient score closer to 1 would indicate higher fidelity copying (overimitation), whereas a quotient score closer to 0 would indicate lower fidelity copying (emulation). As well as quotient score, participants' performance of individual aspects of the task, such as the irrelevant actions used, was also examined, in order to gain a better understanding of more nuanced behaviour during task completion.

3.4. Method (hook task)

3.4.1. Participants

Eye tracking data for this task was collected for 125 participants in total. For the purposes of initial analysis, participants were divided into the age groups shown in table 3.4. However, analysis of behavioural data was conducted for all 124 child and 27 adult participants, even those for whom eye tracking data was not obtained.

Table 3.3: Mean age of participants in each assigned age group (hook task)

Age Group	N	Mean Age	Std. Deviation
3-5	48	3.88	0.815
6-8	29	7.00	0.886
9-11	18	9.72	0.958
12-17	14	12.71	0.994
18+	16	18.81	0.834
Total	125	8.34	5.047

3.4.2. Materials and Apparatus

In this task, participants were presented with an upright plastic tube, measuring 25.5cm tall by 8cm in diameter, which covered a yellow bucket (see Figure 3.9). Inside the bucket was a sticker, which could be seen by participants if they looked down into the tube. Tape was placed over the sides of the opening at the top of the tube to prevent participants from reaching in to retrieve the bucket.

Figure 3.9: Tube and reward presented to participants: side view (L), top view (R).

The bucket could be retrieved from the tube only by using plastic pieces, which could be connected to form a hook. Participants were provided with identical pieces to those used by the demonstrator, alongside a number of distracter pieces (see Figure 3.10). This task was designed so that a varying range of copying strategies could be measured. Participants who imitated very precisely would copy not only the general overall shape of the hook tool, but would also use identical pieces to those used by the demonstrator, and would build the hook in the same order.

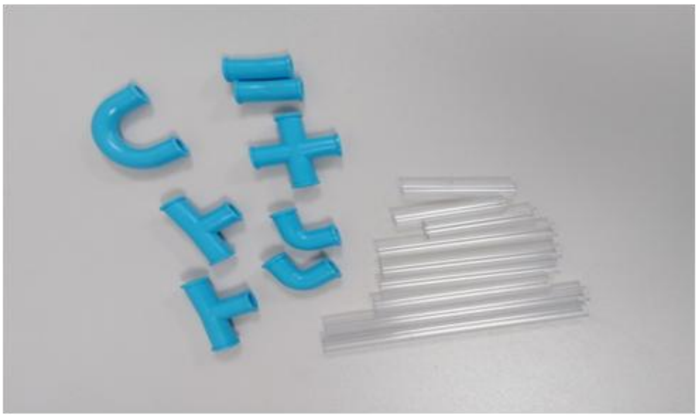


Figure 3.10: Pieces provided to participants in order to build hook to retrieve reward

The completed hook, as shown in the demonstration video, can be seen in Figure 3.11.

Figure 3.11: Completed hook tool as shown in demonstration video

Eye Tracking

Following calibration, participants viewed a 69000ms demonstration video followed by a 3000ms fixation cross three times (see Figure 3.12).

Figure 3.12: Trial sequence shown to participants; participants viewed the same demonstration video (69000ms) three times, followed by a 3000ms fixation cross.

Demonstration video

A female demonstrator was used for this task. She faces the camera at the beginning of the video. The demonstrator builds the hook from the bottom up, by adding a small blue hook piece to a long clear plastic piece. She then adds a straight blue connecting piece, along with another long clear piece. She then taps the partially built hook tool onto the plastic tube containing the reward bucket. Finally, the demonstrator adds a handle and uses the completed hook to retrieve the bucket, which contains the reward. She then removes the sticker from the bucket, faces the camera again and smiles, whilst holding the reward (see Figure 3.13).

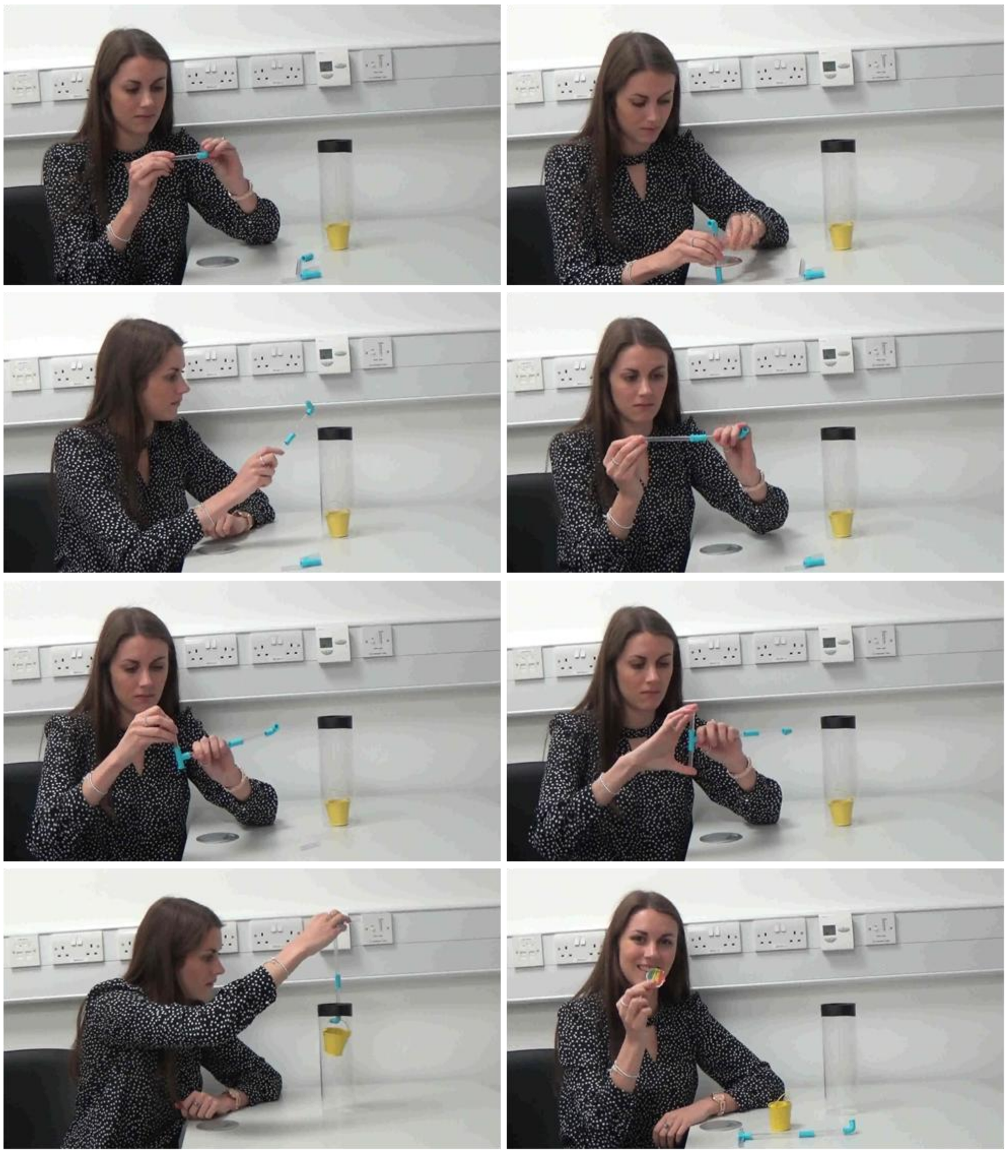


Figure 3.13: Sequence of actions shown to participants during the demonstration video.

Memory Tests

The true images used in the content memory task are shown in Figure 3.14, and the false images are shown in Figure 3.15. In this task, participants were required to indicate which of the images they had seen in the demonstration video, and which they had not.

Figure 3.14: True images shown to participants in content memory task – the demonstrator adds a handle to the hook (L) and taps the hook onto the plastic tube (R).

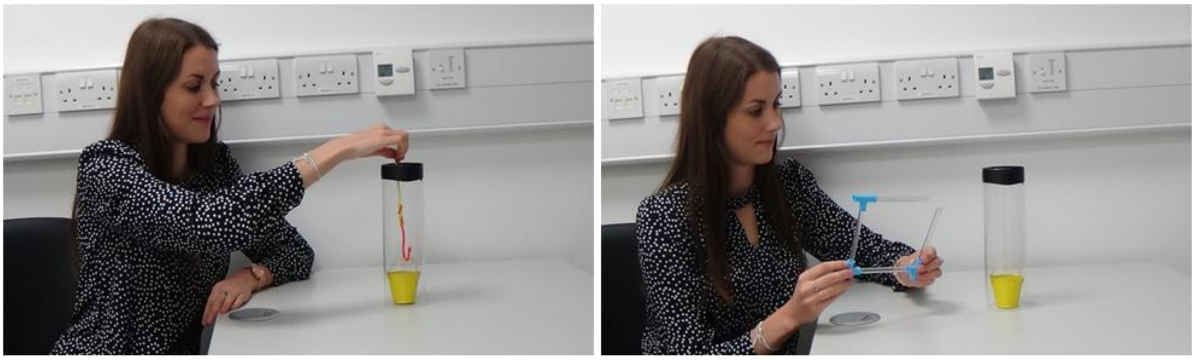


Figure 3.15: False images shown to participants in content memory task – the demonstrator uses a hook made of different materials to those used in the demonstration video (L) and uses the pieces to build a square instead of a hook shape (R).

The images shown to participants who completed the sequence memory task are indicated in Figure 3.16. These images were shuffled, and participants were asked to place them back into the correct order.

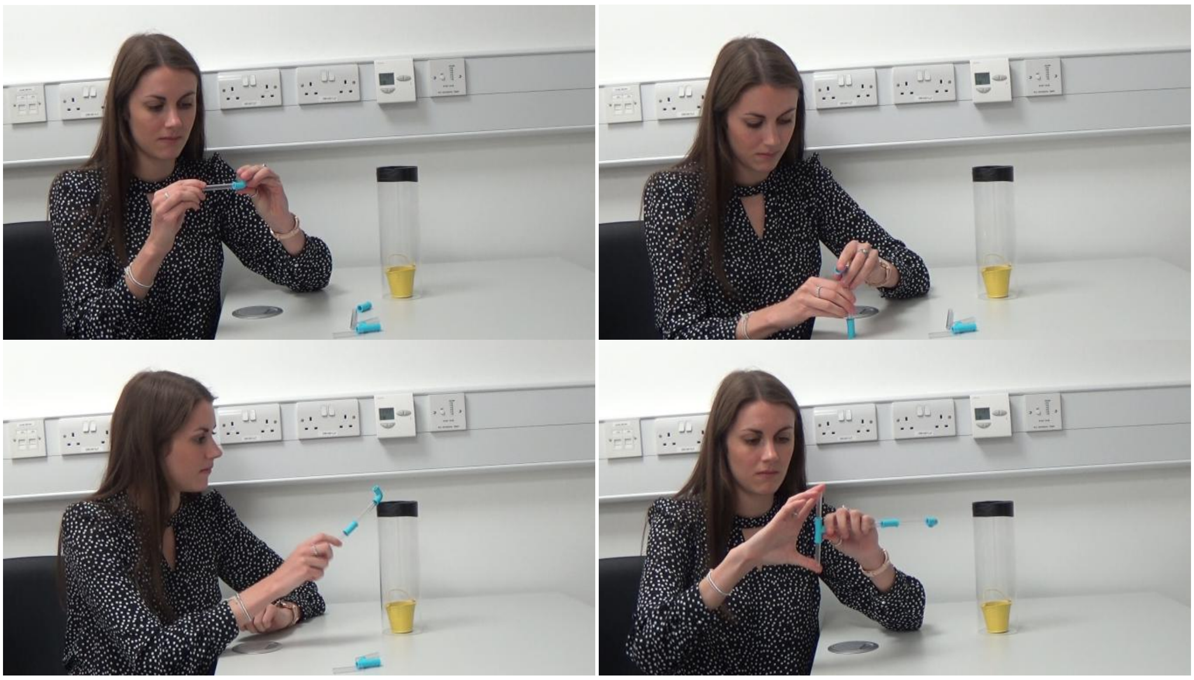


Figure 3.16: Images shown to participants during sequence memory task, presented in the correct sequence. The demonstrator links two pieces together, adds the hook shape onto one end, taps the incomplete hook tool on the plastic tube, and finally adds a handle to the hook.

3.4.3. Data analysis

Eye Tracking

As previously described, areas of interest for this task were identified as the demonstrator's face, the task (referring to the hook, the tube and the reward), and the background of the demonstration video.

Behavioural Scoring

Participant data was scored depending on the number of irrelevant actions produced and the order in which the relevant actions were copied. Within this task, there were two possible irrelevant actions, as well as three potential measures of copying fidelity: the precise pieces used by the participant to build the hook, the order in which the hook was built (top to bottom, or bottom to top) and the overall shape of the hook itself. Participants were awarded a score of 1 for each of these elements that they completed, meaning that the maximum possible score for this task was 5. A quotient score was then calculated for each participant, by taking their actual score on the task and dividing it by the maximum possible score. A quotient score closer to 1 would indicate higher fidelity copying (overimitation), whereas a quotient score closer to 0 would indicate lower fidelity copying (emulation). As well as quotient score, participants' performance of individual aspects of the task, such as the irrelevant actions used, was also examined, in order to gain a better understanding of more nuanced behaviour during task completion. Some example hooks produced by participants that differed to the one produced in the demonstration videos can be seen in Appendix C.

3.5. Method (colouring task)

3.5.1. Participants

Eye tracking data for this task was collected for 119 participants in total, divided into the age groups indicated in table 3.4. However, analysis of behavioural data was conducted for all 124 child and 27 adult participants, even those for whom eye tracking data was not obtained.

Table 3.4: Mean age of participants in each assigned age group (colouring task)

Age Group	N	Mean	Std. Deviation
3-5	46	3.85	0.816
6-8	30	7.03	0.890
9-11	19	9.63	0.831
12-17	11	12.73	1.009
18+	13	18.92	0.862
Total	119	8.04	4.830

3.5.2. Materials

Colouring Task

Participants were presented with a colouring sheet containing three shapes, as shown in figure 1, as well as felt tip pens, including the four colours used by the demonstrator (green, yellow, purple, blue) and four distracter colours (red, pink, orange, black). In contrast to the box and hook tasks, this task did not have any obvious reward (e.g. the ball in the box task, or the sticker in the hook task). The colouring sheet was held onto a vertical magnetic board using two triangle shaped magnets, one in each of the bottom corners; these can be seen in Figure 3.17.

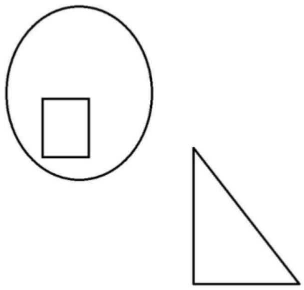


Figure 3.17: Colouring sheet given to participants

The picture, as completed by the demonstrator, is shown in Figure 3.18.

Figure 3.18: Completed colouring task as shown in demonstration video

Eye Tracking

Following calibration, participants viewed a 105000ms demonstration video followed by a 3000ms fixation cross three times (see Figure 3.19).

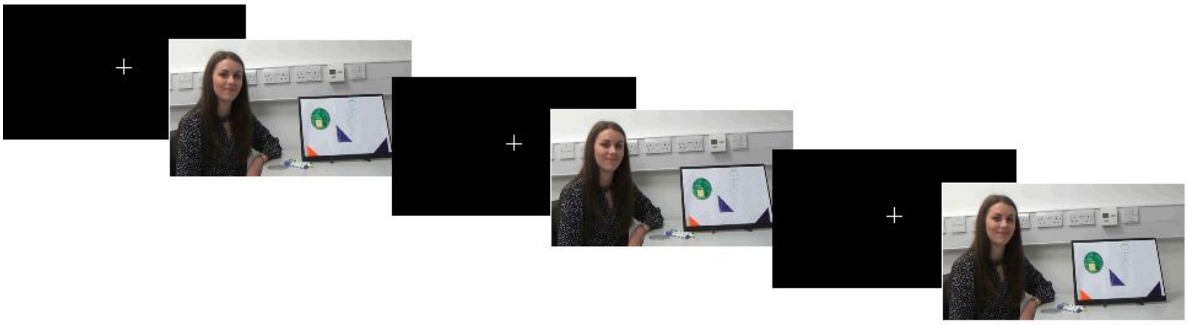


Figure 3.19: Trial sequence shown to participants

Demonstration video

A female demonstrator was used for this task. She colours the rectangle in yellow, draws a blue stick figure to the right of the shapes, colours the triangle in purple, and the oval in green. At the end of the video, she faces the camera and smiles.

Memory Tests

The true images used in the content memory task are shown in Figure 3.20, and the false images are shown in Figure 3.21. In this task, participants were required to indicate which of the images they had seen in the demonstration video, and which they had not.

Figure 3.20: True images shown to participants in content memory task

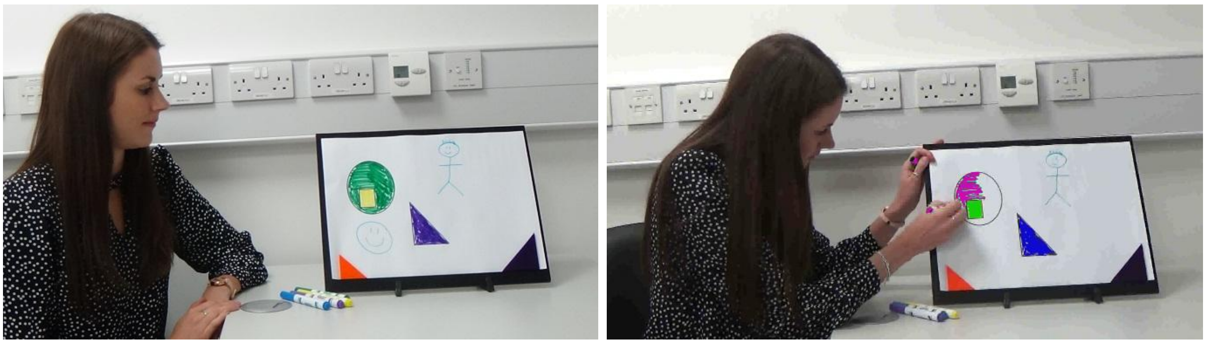


Figure 3.21: False images shown to participants in content memory task

The images shown to participants who completed the sequence memory task are indicated in Figure 3.22:. These images were shuffled, and participants were asked to place them back into the correct order.

Figure 3.22: Images shown to participants in sequence memory task

Eye Tracking

As previously described, areas of interest for this task were identified as the demonstrator’s face, the task (referring to any area of the colouring sheet), and the background of the demonstration video.

3.5.3. Data Analysis

Behavioural Scoring

Participant data was scored depending on the number of irrelevant actions produced and the order in which the relevant actions were copied. Within this task, there was one possible irrelevant action (the stick figure) and five potential measures of copying fidelity: the order in which participants coloured the shapes, and the colours used for each shape (rectangle, stick figure, triangle, oval). Participants were awarded a score of 1 for each of these elements that they completed, meaning that the maximum possible score for this task was 6. A quotient score was then calculated for each participant, by taking their actual score on the task and dividing it by the maximum possible score. A quotient score closer to 1 would indicate higher fidelity copying (overimitation), whereas a quotient score closer to 0 would indicate lower fidelity copying (emulation). As well as quotient score, participants' performance of individual aspects of the task, such as the irrelevant actions used, was also examined, in order to gain a better understanding of more nuanced behaviour during task completion.

Inter-rater reliability

Inter-rater reliability was tested for each task, based on a second individual blind coding of 15% of the data collected. Cohen's Kappa indicated good agreement between raters for the box task (K=.696, p=.001), the hook task (K=.426, p=.001) and the colouring task (K=.418, p<.001).

3.6. Box Results

3.6.1. Developmental results

Theory of Mind

Comparisons were made between performance on the tasks measuring developmental abilities (theory of mind, sequence memory) and copying performance.

All participants completed the theory of mind task at the end of the testing session. The overall results obtained for this task are indicated in table 3.5, which indicates the number of participants in each age group that either passed or failed the false belief task. Participants did not begin to perform at ceiling on this task until the age of

9. However, more than 85% of 6-8 year old participants were able to pass the task, as opposed to 51% of 3-5 year olds. This finding was consistent with previous research, which suggests that the ability to pass false belief tasks begins to develop between the age of 4 and 5.

Table 3.5: Number of participants in each age group that passed and failed the theory of mind task

Age	Pass	Fail	Total
3-5	24	23	47
6-8	26	4	30
9-11	20	0	20
12-17	20	0	20
18+	25	0	25
Total	115	27	142

Since a significant difference in theory of mind scores was found between age groups, a binary logistic regression was used to compare ToM scores between copying groups, with the predictors being age and copying type. The regression was found to be significant; x2(2)=51.52, $p<.001$, and the model explained 49.2% of the variance in task performance. As expected, age was found to be a significant predictor of ToM ability ($p<.001$). However, quotient score was not ($p=.307$). However, as indicated in figure 63, all participants who achieved a quotient score of .80 or 1.00 passed the false belief task.

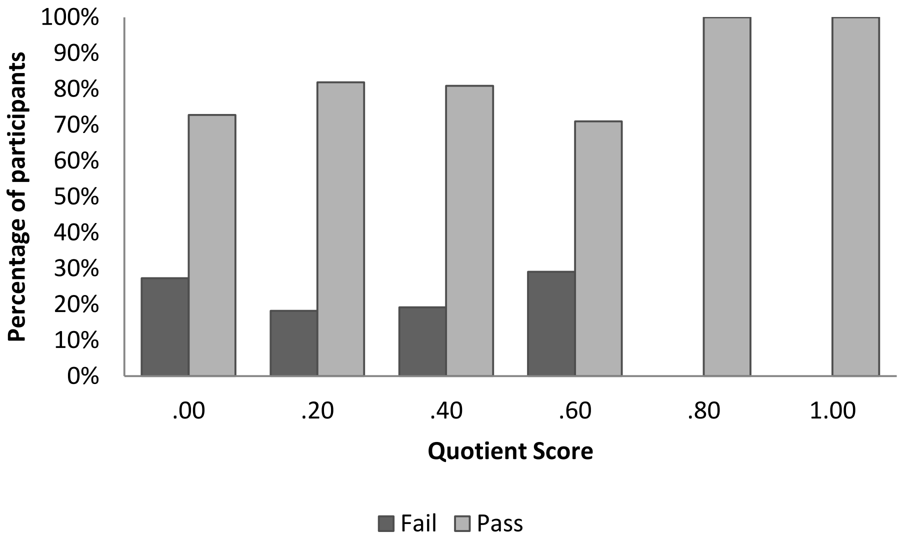


Figure 3.23: Percentage of participants in each quotient score group that passed or failed the false belief task.

Sequence memory

Participants' performance on the sequence memory task can be seen in table 3.6.

Table 3.6: Mean sequence memory score for participants in each age group (box task)

Age Group	N	Mean	Std. Deviation
3-5	23	1.35	1.369
6-8	16	3.44	1.031
9-11	13	3.23	1.301
12-17	13	4.00	0.000
18+	16	3.56	0.964
Total	**81**	**2.93**	**1.481**

As expected, the youngest age group performed poorly on this task in comparison to other age groups. When using a one-way ANOVA to examine this data, Levene's test indicated unequal variances (F=8.543, p<.001), and therefore a Kruskal Wallis test was used for analysis instead. This indicated a significant difference between age groups: $x^2(4)=38.19$, p<.001. The mean ranks are indicated in table 3.7. Post hoc pairwise comparisons indicated a significant difference between the sequence memory score obtained by 3-5 year olds and 6-8 year olds (p<.001), 9-11 year olds (p<.001), 12-17 year olds (p<.001) and 18+ year olds (p<.001), with 3-5 year olds achieving the lowest score.

Table 3.7: Mean ranks for sequence memory scores

Age	N	Mean Rank
3-5	23	19.39
6-8	16	47.59
9-11	13	44.96
12-17	13	56.50
18+	16	49.66

Interestingly, participants' performance on this memory task did not rise consistently with age. The results appear to suggest that 3-5 year olds performed particularly poorly, but memory performance did increase rapidly above the age of six, as expected, and remained high in the older age groups.

Since the number of participants in some of the quotient score groups was low, as indicated in table 3.8, a Spearman's rho correlation was used to examine the relationship between quotient score and mean sequence memory score. There was no significant relationship found: r_s=.140, p=.218.

Table 3.8: Mean sequence memory score for participants (box task)

Quotient Score	N	Mean	Std. Deviation
.00	8	1.50	1.773
.20	18	3.61	0.916
.40	31	2.74	1.527
.60	14	2.79	1.578
1.00	8	4.00	0.000

Content memory

As demonstrated in

Table 3.9, the majority of participants performed at ceiling in the content memory task, suggesting that it was too easy for participants in all age groups. As would be expected from these results, there were no significant differences in performance on this task between age groups; F(4,56)=.74, p=.569 Of the 57 participants assigned to this group, three gave one incorrect answer during this task, with the remaining 54 achieving the maximum possible score. No participants in any group scored lower than 3 on this task.

Table 3.9: Mean content memory task scores for participants in each age group

Age Group	N	Mean	Std. Deviation
3-5	20	3.90	0.308
6-8	14	4.00	0.000
9-11	7	4.00	0.000
12-17	7	4.00	0.000
18+	9	3.89	0.333
Total	57	3.95	0.225

Similarly, there was no significant correlation found between quotient score and content memory score, using Spearman's rho: r_s=.144, p=.282.

3.6.2. Behavioural data

As described previously, participants were allocated a quotient score based on their task performance, with a score closer to 1 indicating more faithful copying from the demonstration video. The mean quotient scores for each age group are indicated in table 3.10, and were compared using an ANOVA, which indicated a significant difference in mean quotient score between age groups: F(4,138)=6.30, p<.001. Post-hoc Bonferroni tests indicated a significant difference between 3-5 year olds (M=.39) and 6-8 year olds (M=.58, p=.014), 6-8 year olds and adults (M=.27, p<.001), 9-11 year olds (M=.52) and adults (p=.018), and 12-16 year olds (M=.50) and adults (p=.031).

Table 3.10: Mean quotient scores for participants in each age group (box task)

Age	N	Mean	Std. Deviation
3-5	45	0.39	0.227
6-8	30	0.58	0.259
9-11	19	0.52	0.300
12-17	20	0.50	0.247
18+	25	0.27	0.251
Total	139	0.44	0.271

Figure 3.24 demonstrates what percentage of participants achieved each of the six quotient scores (0, 0.2, 0.4, 0.6, 0.8, and 1.0) within each age group. A significant difference between groups was observed: $x^2(24)=62.03$, $p<.001$. Of all groups, 6-8 year old participants were the most likely to overimitate.

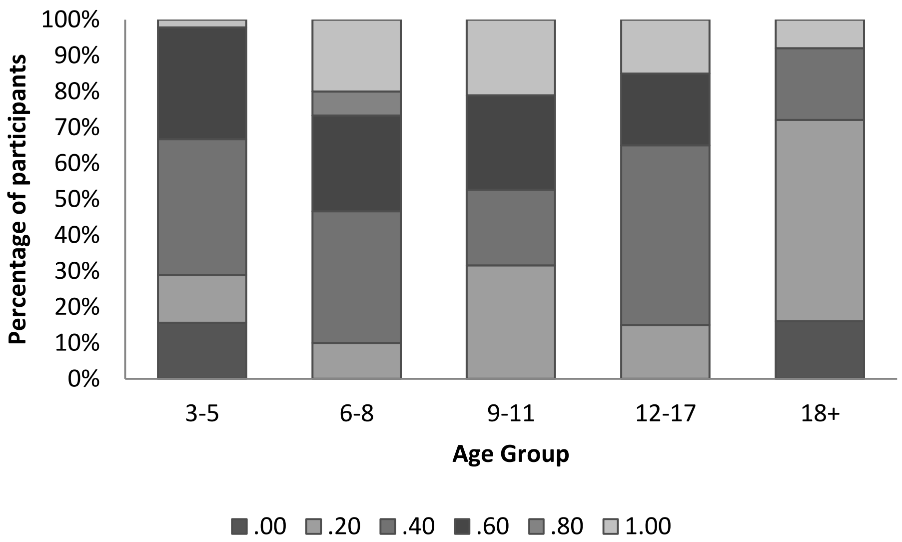


Figure 3.24: Percentage of participants in each age group that achieved each of the possible quotient scores (.00 (N=6), .20 (N=13), .40 (N=38), .60 (N=26), .80 (N=2), 1.00 (N=12))

Although the fidelity quotient allows for a broad measure of copying fidelity across age groups, it does not indicate whether participants are more likely to copy some actions than others. For this reason, each of the irrelevant actions was investigated separately in order to determine whether imitation of particular actions increased or decreased with age. In experiment one, participants were observed to copy the lever action more than the knock action, perhaps because the lever appeared more likely to be a functional component of the task. Since an additional irrelevant action, the turn action, was added in this experiment, comparisons between the individual irrelevant actions copied were made in order to determine what proportion of participants copied each of these three actions, and whether a similar pattern to experiment1 could be observed.

Figure 3.25 indicates the number of participants in each age group that imitated the knock action.

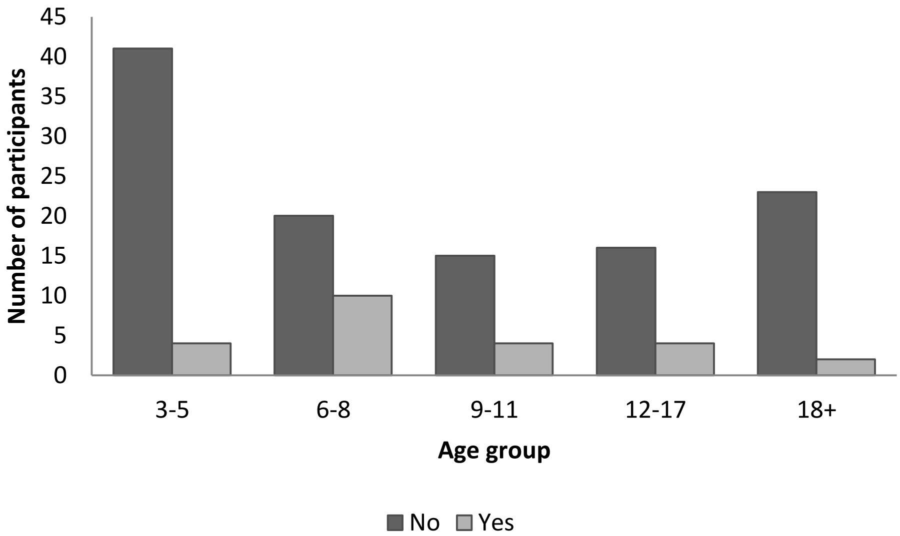


Figure 3.25: Number of participants in each age group that imitated the irrelevant knock action

A chi-squared test indicated a difference that was close to significance; $x^2(4)=9.43$, p=.051. 6-8 year old participants were most likely to imitate the knock action, and this then decreased with age. Interestingly, the proportion of participants in the youngest and oldest age groups that imitated the knock action was very similar (4% and 2%), although there may be differing reasons why each of those age groups chose to copy the action.

Figure 3.26 indicates the number of participants in each age group that chose to copy the irrelevant lever action.

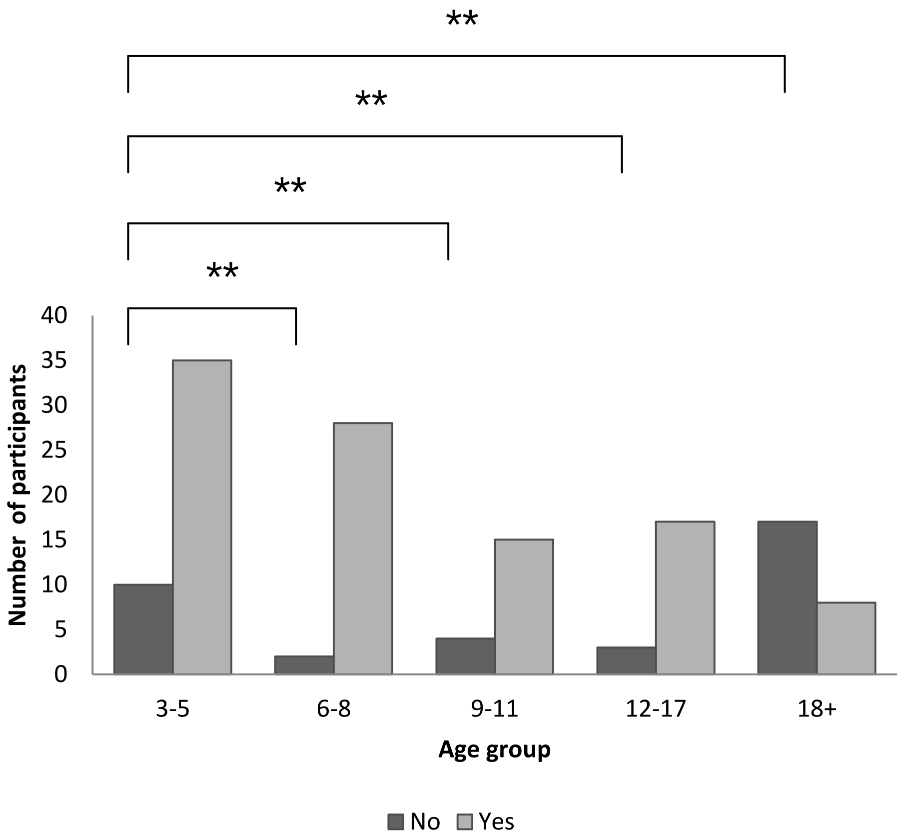


Figure 3.26: Number of participants in each age group that imitated the irrelevant lever action. * indicates significance at the .05 level, ** indicates significance at the .01 level.

Participants in all age groups were more likely to copy the lever action than the knock action. This is consistent with findings from experiment one. The difference between age groups was observed to be significant, using a chi-square test: $x^2(4)=30.66$, $p<.001$. Post-hoc pairwise comparisons indicated a significant difference between 3-5 year olds and 6-8 year olds ($p=.001$), 9-11 year olds ($p<.001$), 12-17 year olds ($p=.020$) and adults ($p=.004$). The majority of participants in all age groups other than the adult age group copied the lever action. However, only 32% of adults chose to copy the action. This may suggest that adults understood that the top lever was not functional, whereas child participants did not.

Figure 3.27 shows the number of participants in each age group that imitated the turn action.

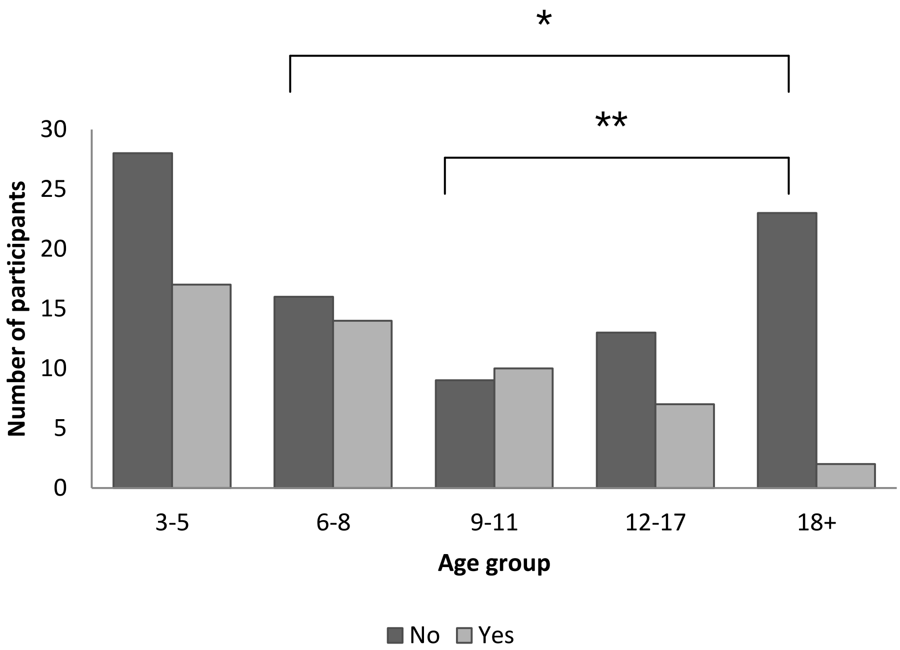


Figure 3.27: Number of participants in each age group that imitated the irrelevant turn action. * represents significance at the .05 level. ** represents significance at the .01 level.

The difference between age groups was found to be significant: $x^2(4)=12.33$, p=.015. Post-hoc pairwise comparisons indicated a significant difference between 6-8 year olds and adults (p=.017) and between 9-11 year olds and adults (p=.010), but not between any of the remaining age groups. The pattern observed for this action was similar to that seen with the knock action: the majority of participants in all of the child age groups copied the action, in comparison to only 8% of adult participants. Again, this may indicate that adults may simply understand more about the functionality of the task and the purpose of the actions.

The order in which participants completed the relevant actions was also investigated, and is shown in figure 3.28.

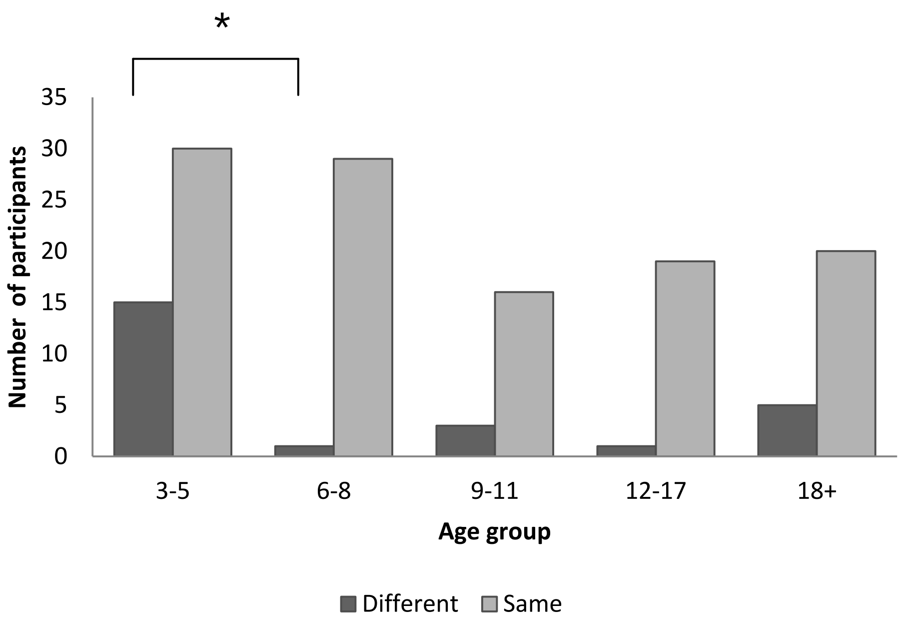


Figure 3.28: Number of participants that completed the relevant actions in the same order as the demonstrator. * indicates significance at the .05 level.

A chi-squared test indicated a significant difference between the number of participants in each age group that completed the relevant actions in the same order as the demonstrator: $x^2(4)=13.97$, $p=.007$. The proportion of participants in each age group that completed the relevant actions in order was then compared, and a significant difference was observed between 3-5 year olds and 6-8 year olds ($p=.019$), with 3-5 year olds being significantly more likely to use a different order than that demonstrated in the video. The majority of participants in all age groups copied the relevant actions in the same order as the demonstrator. However, participants in the 3-5 year old age group were the most likely of all age groups to use the opposite order, followed by the adult age group. As before, it may be the case that the youngest children and adults used the opposite order for different reasons, although this is not immediately clear from the results.

3.6.3. Eye Tracking Data

Eye tracking data was collected for 122 participants in total. The number of participants in each age group is indicated in table 3.11.

Table 3.11: Mean age of participants in each assigned age group (box task)

Age Group	N	Mean	Std. Deviation
3-5	43	3.86	0.833
6-8	29	7.07	0.884
9-11	20	9.65	0.875
12-17	13	12.69	0.947
18+	17	20.00	0.000
Total	122	8.76	5.421

Areas of Interest

Four areas of interest (AOIs) were identified for analysis purposes: the demonstrator's face, the internal mechanisms (i.e. the levers) of the box, the empty space inside the box itself, and the demonstrator's hands. These are the same areas of interest that were used in experiment one.

Trial Differences

No significant differences were observed in the mean looking time at any of the three areas of interest between trials 1-3; face: $F(2,370)=.004$, $p=.996$; task: $F(2,367)=1.986$, $p=.139$; background: $F(2,367)=1.741$, $p=.177$. For this reason, unless specified, analysis of eye tracking data was conducted on the data from the initial trial, as this data is likely to be less subject to loss of data due to participants' inability to sustain their attention to the demonstrations.

Attention to demonstration videos

Total looking time at the demonstration video was examined to determine whether there were any differences in ability to sustain attention between the participant age groups. No significant difference in total looking time between age groups was found; $F(4,121)=.671$, $p=.613$, suggesting that differences in copying behaviours between age groups were not due to reduced/increased attention to the demonstration.

Overall looking time at areas of interest

Although the mean looking time at irrelevant areas of the demonstration video decreased with age as expected, this difference was not found to be significant: $F(4,121)=1.63$, $p=.173$. The mean looking times at these areas for each age group can be seen in table 3.12.

Table 3.12: Mean looking time at irrelevant (background) areas of the demonstration video for each age group

Age Group	N	Mean	Std. Deviation
3-5	43	3974.14	3051.45
6-8	29	3879.45	3610.27
9-11	20	3378.20	3195.97
12-17	13	2505.38	1940.83
18+	17	1420.00	6888.90
Total	122	3341.52	3912.36

The mean looking time at each of the relevant areas of interest was then analysed for each age group.

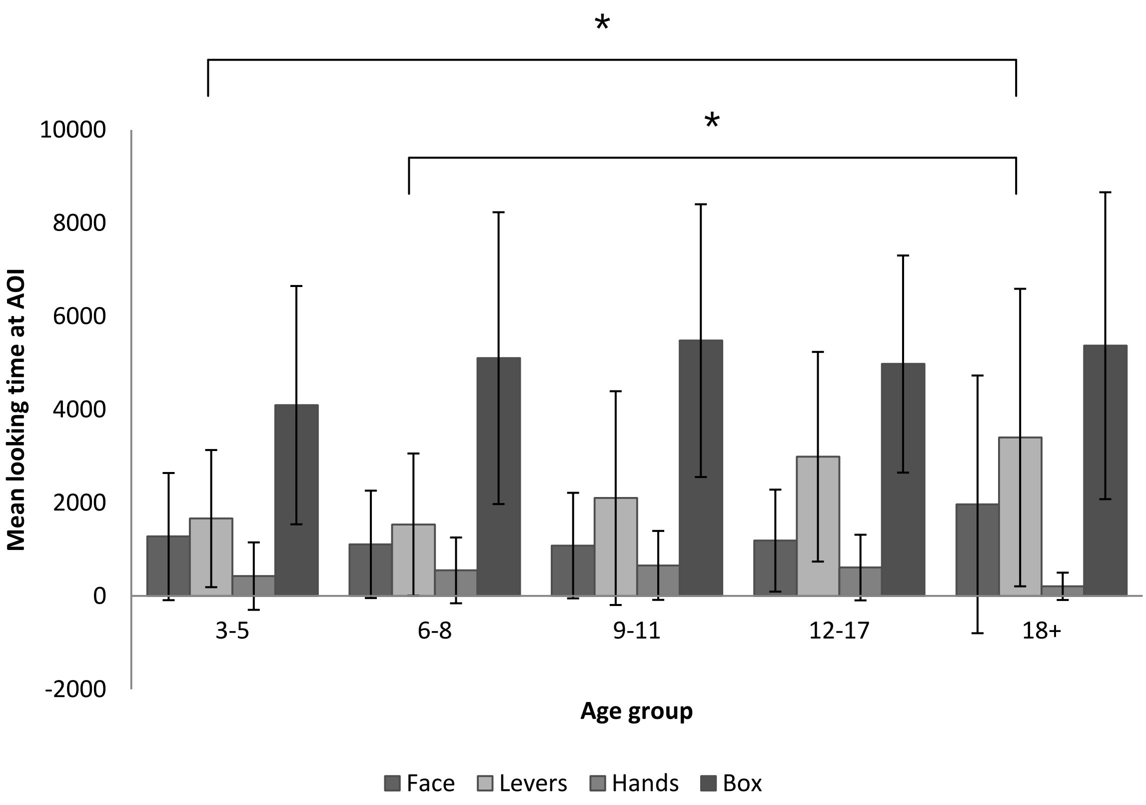


Figure 3.29: Mean proportion of looking time at each of the areas of interest (face, levers, hands, box) for each age group. Error bars represent standard deviation.

As indicated in figure 3.29, participants in all age groups spent the majority of their time looking at the box itself, rather than at the demonstrator's face or hands, or at the actual mechanisms of the box (the levers). Examples of each of these AOIs can be seen in Appendix A. However, no significant difference between age groups was found for the box AOI: F(4,121)=1.20, p=.313.

However, participants' attention to the levers did increase with age. This difference was found to be significant: F(4,121)=3.45, p=.011. Bonferroni post-hoc tests indicated a significant difference in mean looking time at the levers AOI between 3-5 year olds (M=1661) and adults (M=3399, p=.033), and between 6-8 year olds (M=1533) and adults (p=.031), potentially indicating that younger participants are less successful in attending to relevant areas of a demonstration. No significant difference between age groups was found for either the hands AOI: F(4,121)=1.29, p=.278 or the face AOI: F(4,121)=1.03, p=.394.

The mean looking times at each of the areas of interest were then compared between participants achieving each of the different quotient scores. These results are shown in figure 3.30.

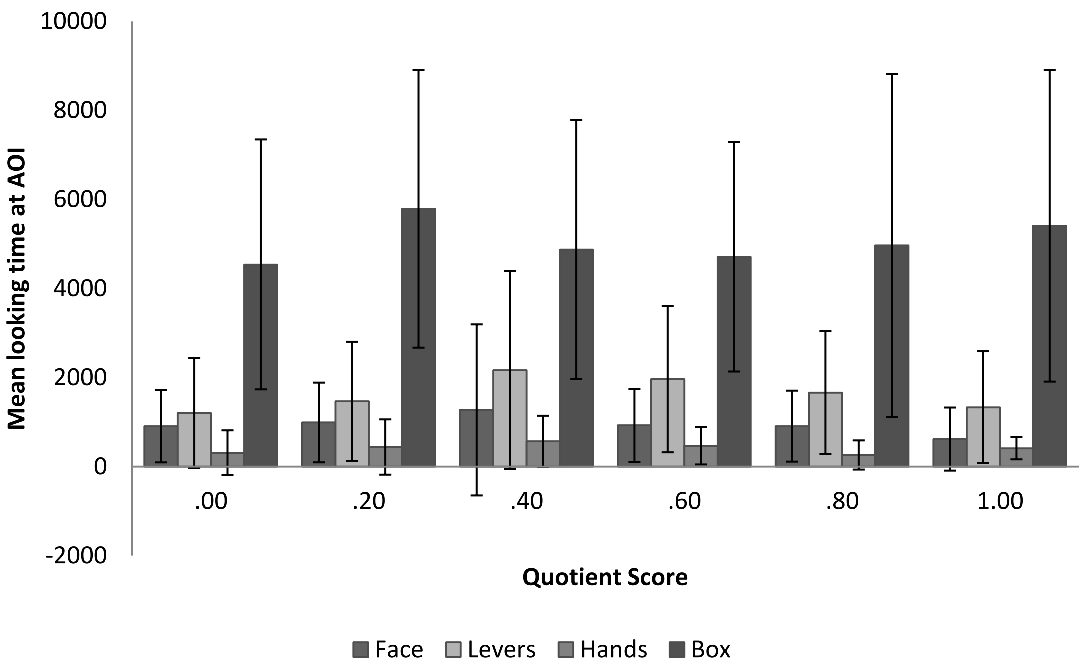


Figure 3.30: Mean looking time at each of the areas of interest (face, levers, hands, box) for each of the possible quotient scores (.00 (N=6), .20 (N=13), .40 (N=38), .60 (N=26), .80 (N=2), 1.00 (N=12)). Error bars represent standard deviation.

There was no significant difference observed between participants achieving each of the quotient scores for the face AOI: F(5,108)=.515, p=.764, the box AOI:

F(5,108)=.422, p=.833, the hands AOI: F(5,108)=.604, p=.697, or the levers AOI: F(5,108)=.873, p=.502. As indicated in figure 3.30, the looking times at each of the areas of interest remained broadly consistent between participants achieving the different quotient scores. This suggests that something other than looking patterns is driving participant behaviour in this task.

3.7. Hook results

3.7.1. Developmental abilities

Theory of Mind

Since a significant difference in theory of mind scores was found between age groups, a binary logistic regression was used to compare ToM scores between copying groups, with the predictors being age and copying type. The regression was found to be significant; $x^2(1)=43.40$, $p<.001$, and the model explained 52% of the variance in task performance. Age was found to be a significant predictor of ToM ability ($p<.001$), however, quotient score was not ($p=.091$). However, as indicated in figure 3.31, the majority of participants achieving a quotient score of 0 failed the theory of mind task.

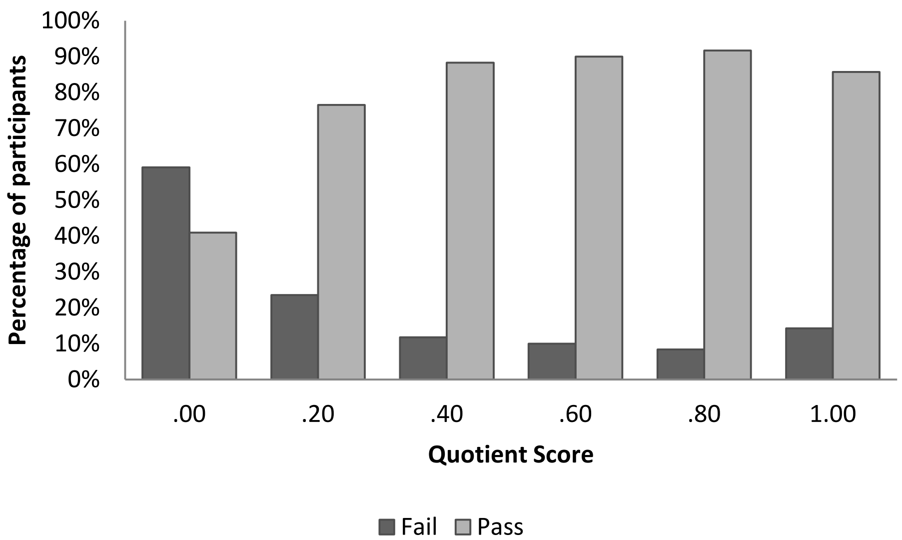


Figure 3.31: Percentage of participants achieving each of the possible quotient scores (.00 (N=22), .20 (N=18), .40 (N=17). ,60 (N=40), .80 (N=36), 1.00 (N=7)) that passed or failed the false belief task

Sequence memory

Participants' performance on the sequence memory task can be seen in table 3.13:

Table 3.13: Mean sequence memory score for each of the age groups (hook task)

Age group	N	Mean	Std. Deviation
3-5	23	1.96	1.296
6-8	16	3.88	0.500
9-11	13	3.85	0.555
12-17	13	4.00	0.000
18+	16	4.00	0.000
Total	81	3.37	1.167

An ANOVA indicated a significant difference between age groups: $F(4,80)=27.60$, $p<.001$. Bonferroni post-hoc tests indicated a significant difference between 3-5 year olds (M=1.96) and 6-8 year olds (M=3.88, $p<.001$), 9-11 year olds (M=3.85, $p<.001$), 12-17 year olds (M=4.00, $p<.001$) and adults (M=4.00, $p<.001$), with 3-5 year olds achieving a lower score in comparison to all other age groups as expected. All participants in the 12-17 and adult age group achieved the maximum possible score on this task.

Due to the small number of participants in some of the quotient score groups, as indicated in table 3.14, a Spearman's correlation was used to determine whether there was a relationship between quotient score and sequence memory score. A moderate positive correlation was found between quotient score and sequence memory score: $r_s = .44$, $p<.001$, indicating that participants with a higher quotient score were also likely to have a high sequence memory score.

Table 3.14: Mean sequence memory score for participants achieving each of the possible hook quotient scores (.00, .20, .40, .60, .80, 1.00)

Hook Quotient	N	Mean	Std. Deviation
.00	12	2.25	1.485
.20	13	3.38	0.961
.40	8	3.13	1.246
.60	24	3.63	1.013
.80	20	3.90	0.447
1.00	2	4.00	0.000
Total	79	3.41	1.115

Content memory

Participants' performance on the content memory task can be seen in table 3.15. As indicated, performance on this task was fairly stable across all age groups, potentially indicating that it was too easy for the majority of participants. No significant difference in content memory score was found between age groups: $F(4,56)=1.92$, $p=.120$.

Table 3.15: Mean content memory scores for participants in each age group (hook task)

Age Group	N	Mean	Std. Deviation
3-5	20	3.60	0.598
6-8	14	3.93	0.267
9-11	7	4.00	0.000
12-17	7	3.86	0.378
18+	9	3.89	0.333
Total	57	3.81	0.441

The mean content memory scores for participants in each quotient score group can be seen in table 3.16. Due to the small number of participants in some of the quotient score groups, a Spearman's correlation was used to determine whether there was a relationship between quotient score and content memory score. A moderate positive correlation was found between quotient score and content memory score: $r_s = .31$, p=.018, indicating that participants with a higher quotient score were also likely to have a high content memory score.

Table 3.16: Mean content memory score for participants achieving each of the possible hook quotient scores (.00, .20, .40, .60, .80, 1.00)

Quotient Score	N	Mean	Std. Deviation
.00	7	3.29	0.756
.20	4	3.75	0.500
.40	9	4.00	0.000
.60	16	3.75	0.447
.80	16	3.94	0.250
1.00	5	4.00	0.000
Total	57	3.81	0.441

3.7.2. Behavioural data

Task performance across each of the age groups was compared, to determine whether participants' behaviour changed with age as expected. The mean quotient scores for each age group are indicated in table 3.17, and were compared using an ANOVA, which indicated a significant difference in mean quotient score between age groups: F(4,138)=.833, p<.001. Post-hoc tests indicated a significant difference in mean quotient score between 3-5 year olds (M=.31) and 6-8 year olds (M=.53, p=.007), 9-11 year olds (M=.046, p<.001) and adults (M=.74, p<.001). There was also a significant difference between 6-8 year olds and adults (p=.039) and between 9-11 year olds and adults (p=.009).

Table 3.17: Mean hook quotient score for participants in each age group

Age	N	Mean	Std. Deviation
3-5	45	0.31	0.288
6-8	30	0.53	0.290
9-11	19	0.46	0.291
12-17	20	0.62	0.214
18+	25	0.74	0.180
Total	139	0.50	0.302

The number of participants in each of the age groups achieving each of the quotient scores was then examined. A significant difference of age was observed, using a chi-squared test: $x^2(20)=77.67$, $p<.001$. Figure 3.31 indicates the differences in task performance between age groups.

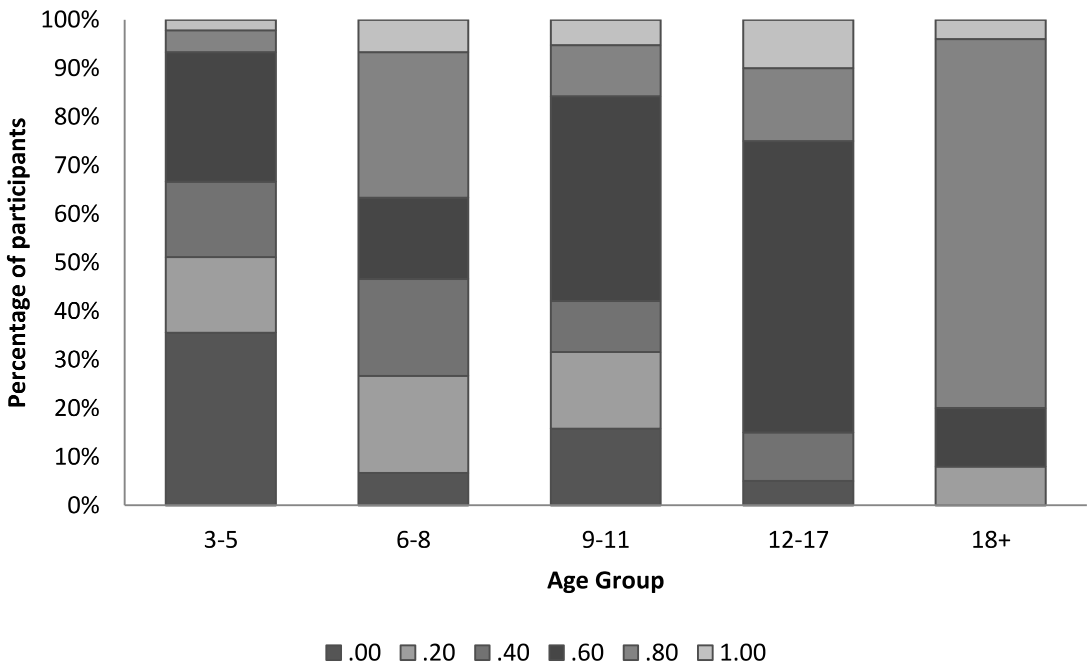


Figure 3.32: Percentage of participants in each age group that achieved each possible quotient score (.00 (N=22), .20 (N=18), .40 (N=17), .60 (N=40), .80 (N=36), 1.00 (N=7))

Copying fidelity increased with age, with the majority of 3-5 year olds achieving a score of 0 or .20, and the majority of adults achieving a score of .80. However, performance across the remaining child age groups was variable; 6-8 year olds appeared to be more likely to copy faithfully than both 9-11 year olds and 12-17 year olds.

As with the box task, the number of participants in each age group that copied the irrelevant actions (tap and handle) was investigated. No significant difference between age groups was observed for the tap action: $x^2(4)=1.85$, $p=.763$. Figure 3.33 shows the number of participants in each age group that copied and did not copy the tap action. Performance of this action was variable with age but did not appear to follow any kind of consistent pattern. Participants aged 6-8 were the most likely to copy the tap action, whereas adult participants were the least likely to perform it.

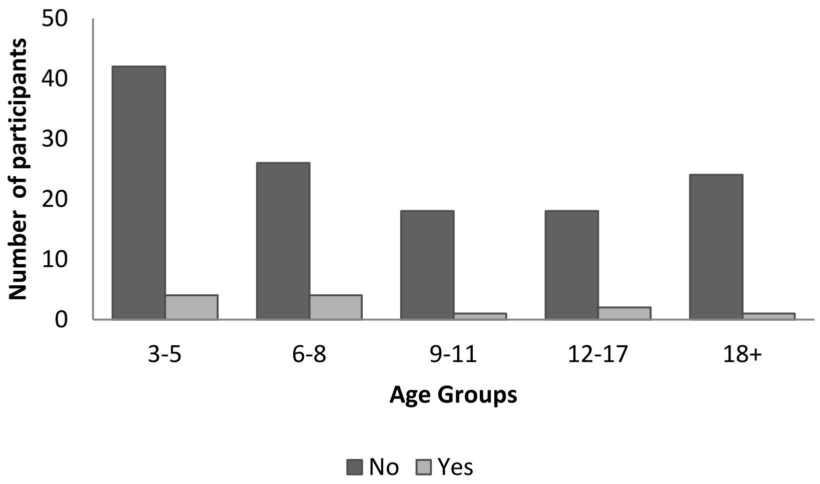


Figure 3.33: Number of participants in each age group that imitated the irrelevant tap action

Figure 3.34 indicates the number of participants in each age group that copied the irrelevant handle from the demonstrator. A significant difference was found between age groups: $x^2(4)=14.80$, $p=.005$, with 3-5 year old participants being least likely to imitate this action, and 12-17 year olds and adults being most likely to copy it. Pairwise comparisons indicated a significant difference between 3-5 year olds and adults ($p=.042$) , with adults being significantly more likely to copy the action in comparison to 3-5 year olds. No other significant differences between age groups were found.

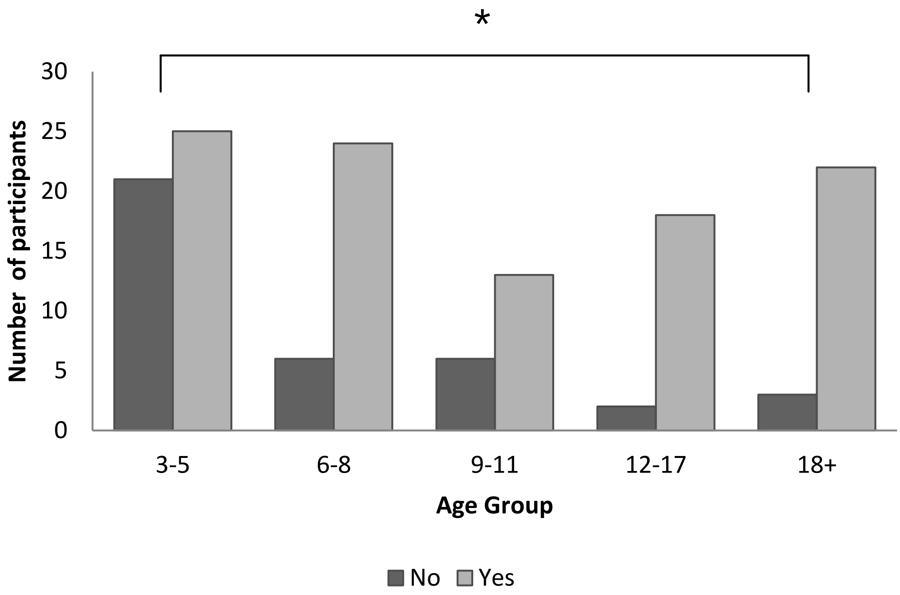


Figure 3.34: Number of participants in each age group that included the irrelevant hook handle. * indicates significance at the .05 level.

The order in which participants built the hook was also investigated. A significant difference between age groups was observed: $x^2(4)=33.15$, $p<.001$. Figure 3.35 indicates that participants' tendency to build the hook in the same order as the demonstrator increased consistently with age. Post-hoc pairwise comparisons indicated a significant difference between 3-5 year olds and 12-17 year olds ($p=.001$), and 3-5 year olds and adults ($p<.001$).

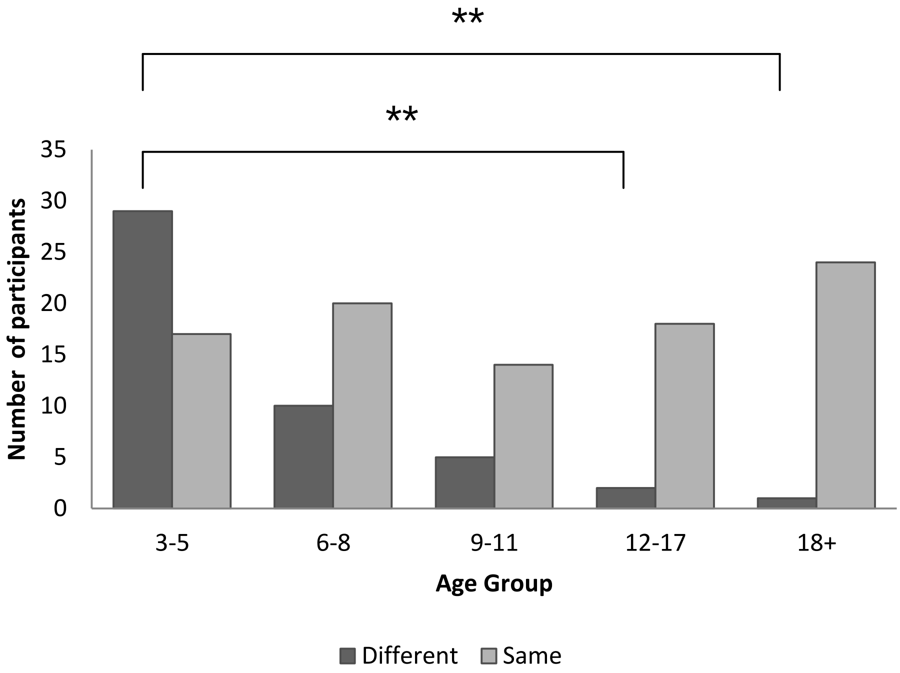


Figure 3.35: Number of participants that imitated the built the hook in the same order (bottom to top) as the demonstrator. ** indicates significance at the .01 level.

3.7.3. Eye Tracking Data

Areas of Interest

Three areas of interest (AOIs) were identified for analysis purposes: the demonstrator's face, the pieces used to build the hook, and the remainder of the task itself(the tube and the reward bucket) , and the background of the demonstration video.

Trial Differences

No significant differences between trials were observed in the mean looking time at the face AOI: F(2,378)=.152, p=.859, the tube AOI: F(2,378)=.554, p=.575, the pieces AOI: F(2,378)=.783, p=.458, or the hands AOI: F(2,378)=1.64, p=.196. For this reason, analysis was conducted on the mean fixation times at each AOI for the first trial only.

Overall looking time at areas of interest

The mean looking time at irrelevant aspects of the demonstration video was found to be fairly consistent across age groups, and there was no significant difference observed between groups: $F(4,122)=.438$, $p=.781$. The mean looking times at irrelevant aspects of the demonstration for each age group can be seen in table 3.18.

Table 3.18: Mean looking time at irrelevant (background) areas of the demonstration video for each of the age groups

Age Group	N	Mean	Std. Deviation
3-5	47	30825.30	13830.33
6-8	29	33786.55	12606.93
9-11	18	34178.11	15410.70
12-17	14	29655.29	12315.02
18+	15	32265.87	12939.70
Total	123	32056.64	13404.92

Mean looking times at each of the AOIs (face, hands, tube, pieces) for each age group are displayed in figure 3.36.

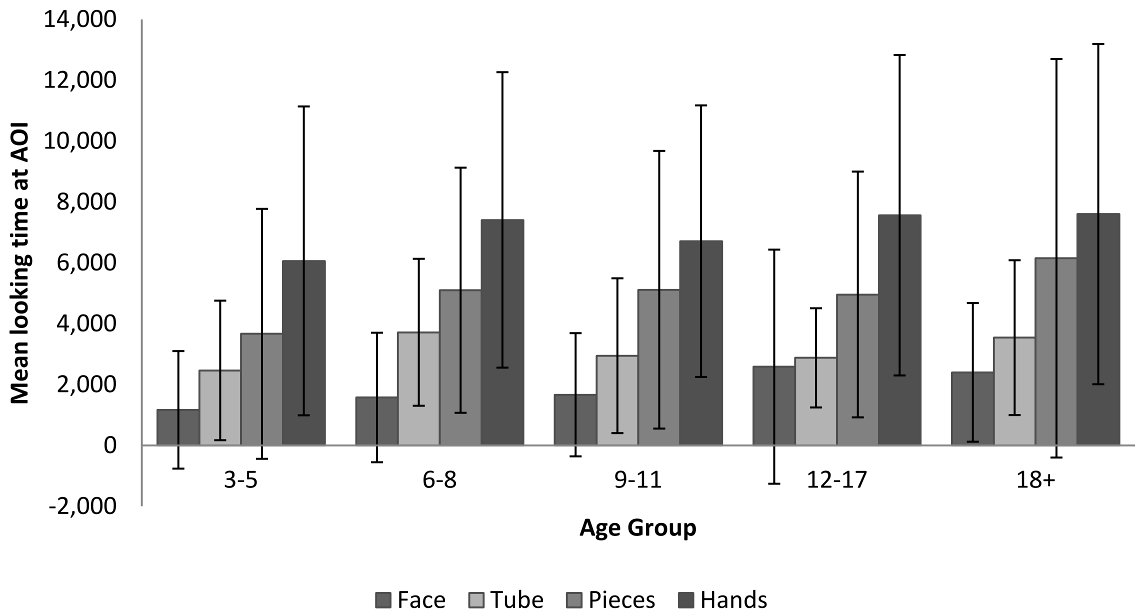


Figure 3.36: Mean looking time at each of the areas of interest (face, tube, pieces, hands) for participants in each age group. Error bars represent standard deviation.

No significant difference between groups was observed for the face AOI: F(4,122)=1.48, p=.212, the tube AOI: F(4,122)=1.52, p=.200, the pieces AOI: F(4,122)=1.11, p=.355, or the hands AOI: F(4,122)=.535, p=.710. Participants in all age groups spent the majority of the time looking at the demonstrator's hands, rather than the task itself, as indicated in figure 3.36 above.

Differences in looking time based on copying behaviours

In order to further explore whether participants' attention to the demonstration video influenced their behaviour, an ANOVA was used to examine the differences in the mean proportion of looking time at each of the four areas of interest (face, tube, pieces, hands) for participants achieving each of the quotient scores. There was no significant difference observed for the face AOI: F(5,114)=2.12, p=.068, the tube AOI: F(5,114)=1.02, p=.411, the pieces AOI: F(5,114)=2.09, p=.073, or the hands AOI: F(5,114)=1.66, p=.151. The mean looking time for participants in each quotient score group is indicated in figure 3.37. As before, participants in all groups spent the majority of their time looking at the demonstrator's hands.

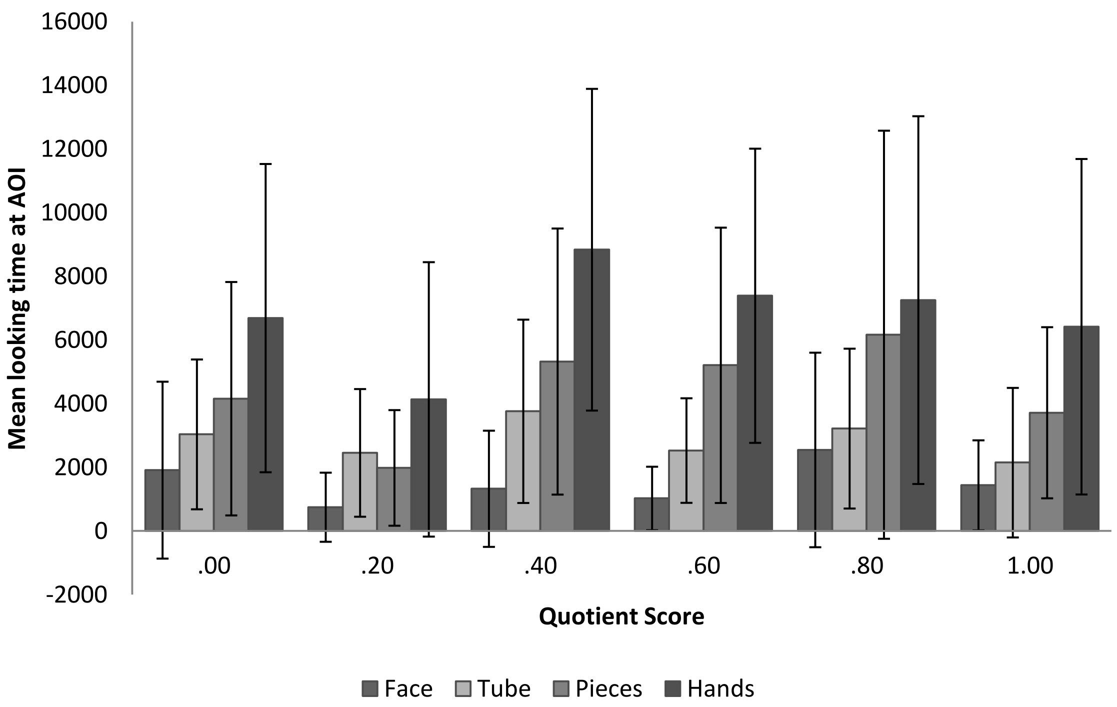


Figure 3.37: Mean looking time at each of the areas of interest (face, tube, pieces, hands) for participants achieving each of the possible quotient scores (.00 (N=21), .20 (N=17), .40 (N=17), .60 (N=17), .80 (N=29), 1.00 (N=6)). Error bars represent standard deviation.

3.8. Colouring Results

3.8.1. Developmental abilities

Theory of Mind

Since a significant difference in theory of mind scores was found between age groups (see section 3.6.1.), a binary logistic regression was used to compare ToM scores between copying groups, with the predictors being age and copying type. The regression was found to be significant; $x^2(1)=43.46$, $p<.001$, and the model explained 49% of the variance in task performance. As expected, age was found to be a significant predictor of ToM ability ($p<.001$). However, quotient score was not ($p=.854$). As indicated in figure 3.38 the majority of participants with a score of 0 failed the theory of mind task, but the majority of participants in all other quotient score groups passed.

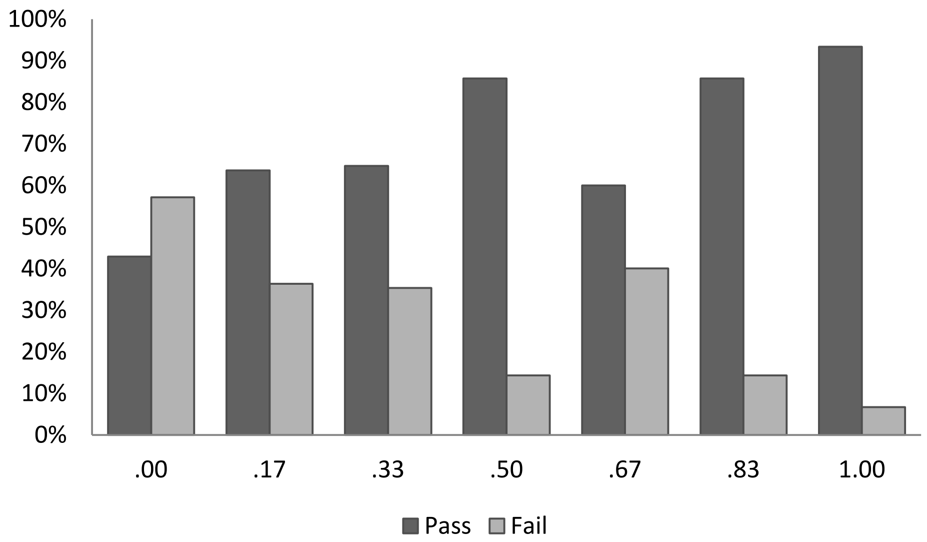


Figure 3.38: Percentage of participants achieving each of the possible quotient scores (.00, .17, .33, .50, .67, .83, 1.00) that passed or failed the theory of mind task

Sequence memory

A significant difference between participant age groups was found on the sequence memory task, using an ANOVA ($F(4,80)=22.27$, $p<.001$), with Bonferroni post-hoc tests indicating a significant difference between 3-5 year olds ($M=1.78$) and 6-8 year olds (M=3.69, p<.001), 9-11 year olds (M=4.00, p<.001), 12-17 year olds (M=4.00, p<.001) and adults (M=4.00, p<.001). No other significant differences were found between age groups. Table 3.19 indicates the mean score for each age group.

Table 3.19: Mean sequence memory score for each of the assigned age groups (colouring task)

Age	N	Mean	Std. Deviation
3-5	23	1.78	1.536
6-8	16	3.69	0.873
9-11	13	4.00	0.000
12-17	13	4.00	0.000
18+	16	4.00	0.000
Total	81	3.31	1.319

Comparisons were then made between copying groups' performance on the sequence memory task (see table 3.20). Due to the small number of participants in some of the quotient score groups, as indicated in table 3.20, a Spearman's correlation was used to determine whether there was a relationship between quotient score and sequence memory score. A moderate positive correlation was found between quotient score and sequence memory score: $r_s = .43$, p<.001, indicating that participants with a higher quotient score were also likely to have a high sequence memory score.

Table 3.20: Mean sequence memory score for participants achieving each of the potential quotient scores (.00, .17, .33, .50, .67, .83, 1.00)

Quotient Score	N	Mean	Std. Deviation
.00	5	2.00	1.414
.17	7	2.86	1.574
.33	9	2.67	1.323
.50	5	3.40	1.342
.67	3	3.00	1.732
.83	17	3.76	0.664
1.00	33	3.79	0.696
Total	79	3.41	1.115

Content memory

Participants' performance on the content memory task can be seen in table 3.21.

Table 3.21: Mean content memory score for participants in each of the assigned age groups (colouring task)

Age	N	Mean	Std. Deviation
3-5	20	3.30	0.923
6-8	14	3.86	0.363
9-11	7	3.71	0.756
12-17	7	3.86	0.378
18+	9	4.00	0.000
Total	57	3.67	0.690

Levene's test indicated unequal variances (F=14.54, p<.001) and therefore a Kruskal-Wallis test was used instead of an ANOVA. No significant difference in mean content memory score was observed between age groups; $x^2(4)=8.17$, p=.086, although 3 year olds had a lower mean rank (23.20) than all other age groups. The mean ranks for each age groups are displayed in table 3.22.

Table 3.22: Mean rank (sequence memory) for each age group

Age	N	Mean Rank
3-5	20	23.20
6-8	14	31.43
9-11	7	30.57
12-17	7	31.43
18+	9	35.00

Comparisons were then made between copying groups' performance on the content memory task. Due to the small number of participants in some of the quotient score groups, as indicated in table 3.23, a Spearman's rho correlation was used to determine whether there was a relationship between quotient score and content memory score. A moderate positive correlation was found between quotient score and sequence memory score: $r_s = .36$, p=.008, indicating that participants with a higher quotient score were also likely to have a high content memory score.

Table 3.23: Mean content memory score for participants achieving each of the possible quotient scores (.00, .17, .33, .50, .67, .83, 1.00)

Quotient Score	N	Mean	Std. Deviation
.00	2	2.50	0.707
.17	4	3.00	1.155
.33	7	3.43	0.976
.50	2	4.00	0.000
.67	2	3.50	0.707
.83	11	3.73	0.647
1.00	26	3.85	0.464
Total	54	3.65	0.705

3.8.2. Behavioural data

Task performance across each of the age groups was compared, to determine whether participants' behaviour changed with age as expected. Figure 3.39 indicates the percentage of participants in each age group that achieved each of the possible quotient scores. A significant difference was observed between age groups in terms of the number of participants in each age group that achieved each quotient score: $x^2(24)=6.48$, p=<.001. The mean quotient score was found to increase steadily with age, and an ANOVA indicated a significant difference in mean quotient score between age groups: F(4,132)=19.37, p<.001. Post-hoc tests indicated a significant difference in mean quotient score between 3-5 year olds (M=.44) and 6-8 year olds (M=.75, p<.001), 9-11 year olds (M=.80, p<.001), 12-17 year olds (M=.94, p<.001) and adults (M=.94, p<.001). 76% of adult participants achieved the maximum score of 1 in comparison to only 15% of 3-5 year olds. Additionally, the majority of 3-5 year olds (32%) achieved a mean quotient score of .33, whereas the majority of 6-8 year olds (38%) achieved a score of 1.00, indicating that copying fidelity rose sharply with age.

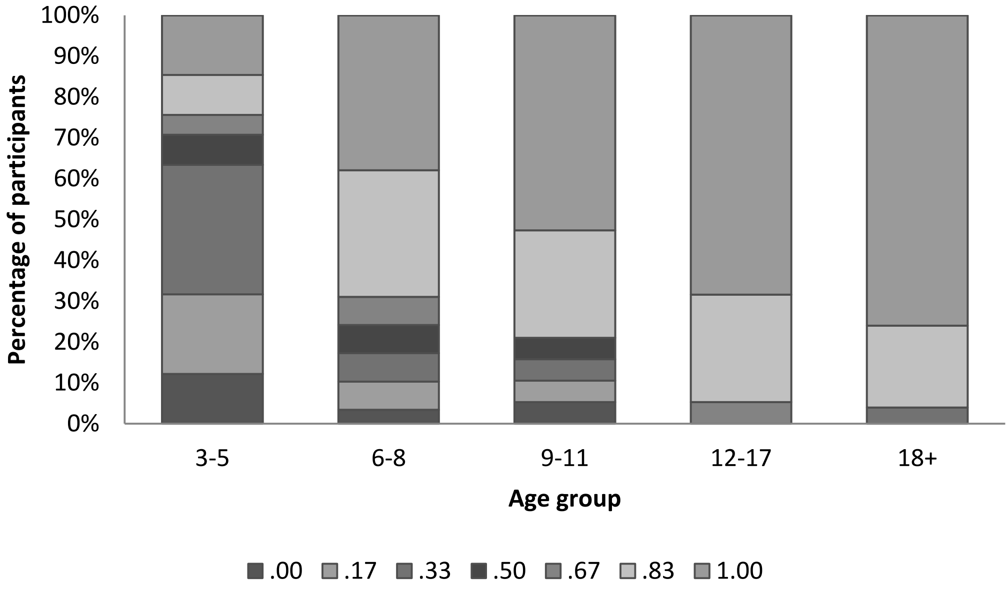


Figure 3.39: Percentage of participants in each age group that achieved each of the possible quotient scores (.00 (N=7), .17 (N=9), .33 (N=15), .50 (N=6) .67 (N=3), .83 (N=21), 1.00 (N=44))

In order to further investigate differences in copying fidelity between groups, the number of participants imitating the colours used, the irrelevant drawing, and the order of shapes coloured was examined. The percentage of participants that used the same colours as the demonstrator to complete the task is shown in figure 3.40.

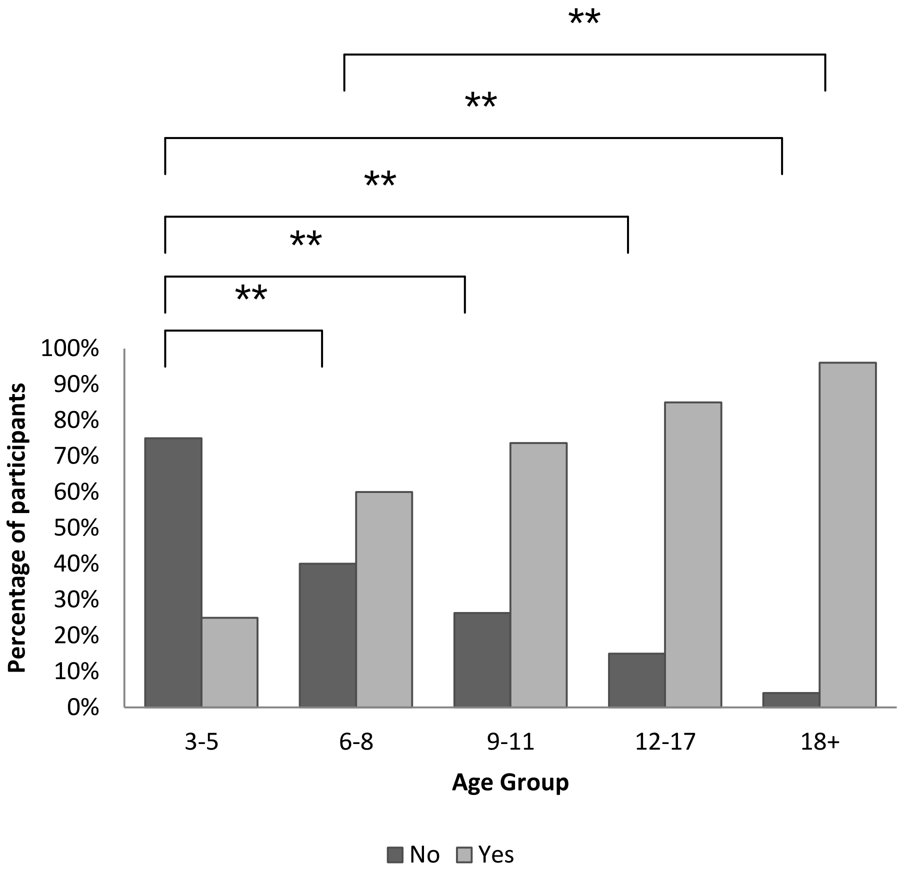


Figure 3.40: Percentage of participants in each age group that used the same colours to complete the picture as those used by the demonstrator. ** indicates significance at the .01 level.

There was a significant difference observed between age groups: $x^2(4)=42.93$, $p<.001$, again with overimitation increasing with age. Post-hoc pairwise comparisons indicated a significant difference between 3-5 year old participants and 6-8 year olds ($p=.020$), 9-11 year olds ($p=.002$), 12-17 year olds ($p<.001$), and adults ($p<.001$). There was also a significant difference between 6-8 year olds and adults ($p=.018$). Only 25% of 3-5 year olds used the same colours as the demonstrator, in comparison to 96% of adults.

To investigate this further, the colours used for each individual shape were investigated, in order to determine whether participants were more likely to copy the colours used for some shapes more than others. The order of shapes coloured by the demonstrator was: rectangle (yellow), triangle (purple), drawing the stick man (blue), oval (green). The percentage of participants in each age group that copied the colour used for each shape are shown in table 3.24. Imitation of the colours used by participants in the 3-5 age group was low across all shapes, as would be expected given the results shown in figure 3.40. Interestingly, the triangle was the shape least likely to be filled in using the same colour as that shown by the demonstrator overall, followed by the drawing. The rectangle colour was the most likely to be copied, followed by the oval. As the rectangle was completed first, and the oval last, this may suggest that both primacy and recency effects were present for all participant age groups.

Table 3.24: Percentage of participants that used the same colour as the demonstrator for each of the shapes in the colouring task.

Age Group	Rectangle	Drawing	Triangle	Oval
3-5	44	29	37	46
6-8	83	69	66	79
9-11	84	79	74	79
12-17	100	100	95	100
18+	100	96	96	100

Post-hoc pairwise comparisons indicated a significant difference between the number of participants aged 3-5 (44%) and 6-8 (83%) that used the correct colour for the rectangle shape (p=.003). There was also a significant difference observed between 3-5 year olds and 9-11 year olds (84%, p=.010). For the drawing, there was a significant difference between 3-5 year olds (29%) and 6-8 year olds (66%, p=.006), 9-11 year olds (74%, p=.002) and adults (96%, p<.001). For the triangle shape, there was a significant difference between 3-5 year olds (37%) and 12-17 year olds (95%, p<.001) and adults (96%, p<.001). Finally, for the oval shape, there was a significant difference between 3-5 year olds (46%) and 6-8 year olds (79%, p=.017). These results indicate that imitation of the colours used rose sharply above the age of 5.

The percentage of participants that included the irrelevant drawing was then investigated, and is shown in figure 3.41.

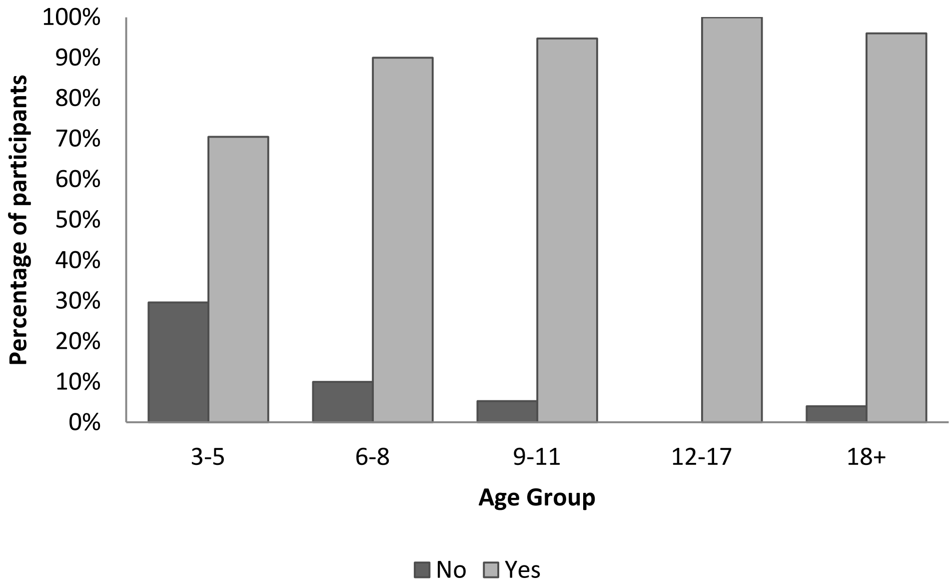


Figure 3.41: Percentage of participants in each age group imitating the irrelevant drawing

There was a significant difference observed between age groups: $x^2(4)=16.63$, $p=.002$. However, post-hoc pairwise comparisons did not indicate any significant differences between individual age groups. 30% of 3-5 year old participants did not include the drawing, in comparison to 4% of adults, and no 12-17 year olds.

The order in which participants coloured the shapes was also investigated. The percentage of participants that used the same order as the demonstration is indicated in figure 3.42.

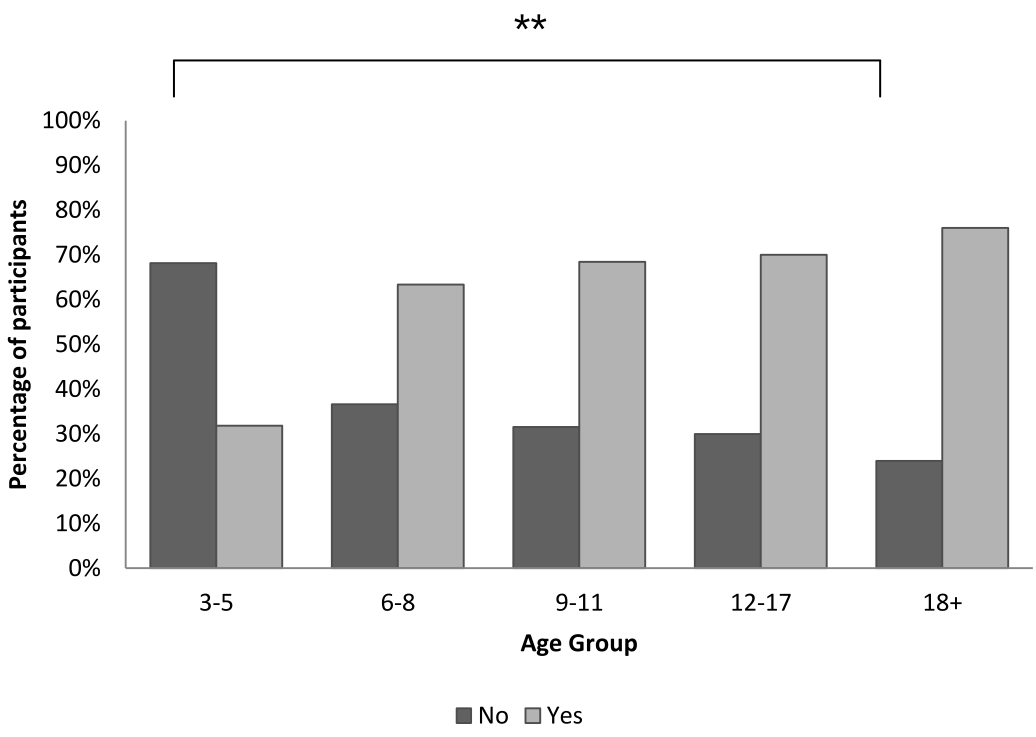


Figure 3.42: Percentage of participants in each age group that coloured the shapes in the same order as the demonstrator. ** indicates significance at the .01 level.

A significant difference between groups was observed: $x^2(4)=17.97$, p=.001. Post-hoc pairwise comparisons indicated a significant difference between 3-5 year olds and adults (p=.006), but not between any other age groups. 68% of 3-5 year olds used a different order to colour the shapes than the one shown in the demonstration video. However, the majority of participants in all other age groups coloured the shapes in the same order as the demonstrator, whether they used the same colours as the demonstrator or not. It is interesting to note that even in the adult age group, 24% of participants used a different order to that used by the demonstrator.

3.8.3. Eye Tracking Data

Areas of Interest

Three areas of interest (AOIs) were identified for analysis purposes: the demonstrator's face, the task itself (the colouring sheet), and the background of the demonstration video. Unlike in previous tasks, it was deemed too difficult to separate fixations between the demonstrator's hand, the pen she was holding, and the board itself, and so these AOIs are all combined into the "task" AOI.

Trial Differences

The total looking time at the demonstration video was compared across trials. Levene's test indicated unequal variances (F=5.640, p=.004) and therefore a Kruskal-Wallis test was used instead of an ANOVA for comparison between trials. No significant difference in looking times between trials were found; $x^2(2)=4.972$, p=.083, and therefore data from the initial trial was used for analysis of eye tracking data unless otherwise stated.

Overall looking time at areas of interest

The mean looking time for each age group at the two AOIs is displayed in figure 3.43.

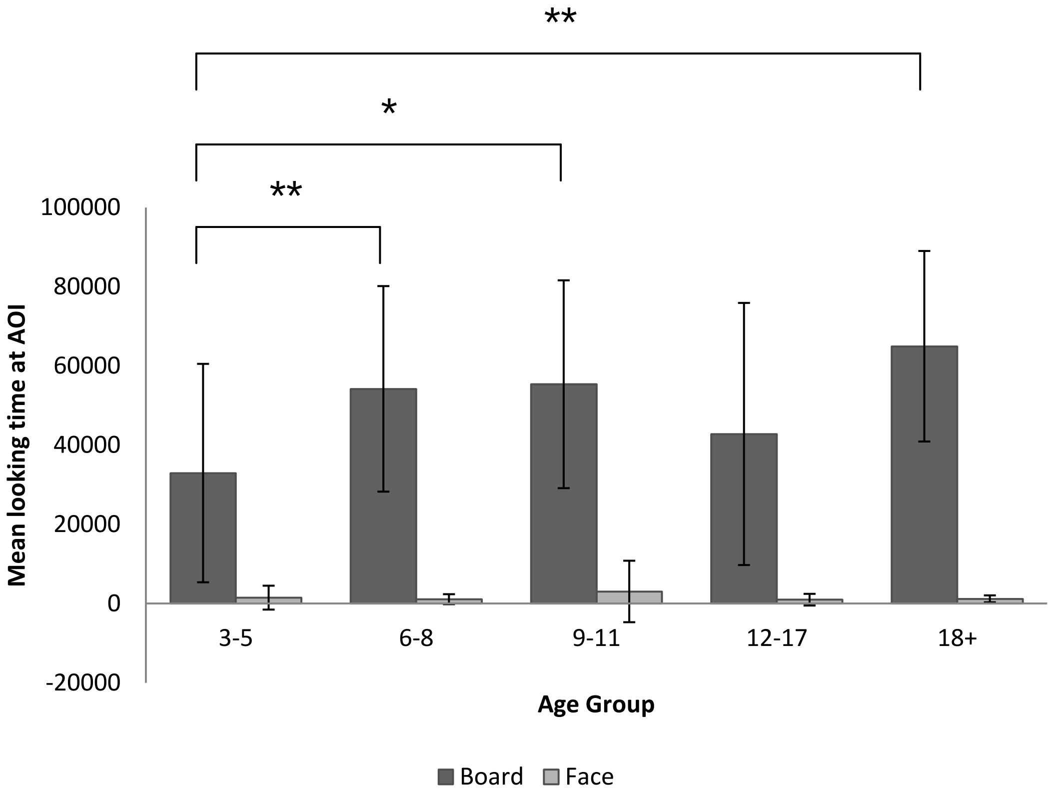


Figure 3.43: Mean looking time at each of the areas of interest (board, face) for participants in each age group. Error bars represent standard deviation. * indicates significance at the .05 level. ** indicates significance at the .01 level.

There was a significant difference in the mean looking time at the board AOI: F(4,118)=5.53, p<.001. Post-hoc Bonferroni tests indicated a significant difference between 3-5 year old participants (M=32919) and 6-8 year old participants (M=54178, p=.011), 9-11 year old participants (M=55339, p=.030) and adults (M=64933, p=.003). Adults spent the longest time attending to the board AOI of all age groups. There was no significant difference observed between groups for the face AOI: F(4,118)=.986, p=.418.

Differences in looking time based on copying behaviours

The mean looking times at the three AOIs by participants achieving each of the potential quotient scores in the colouring task are shown in figure 3.44.

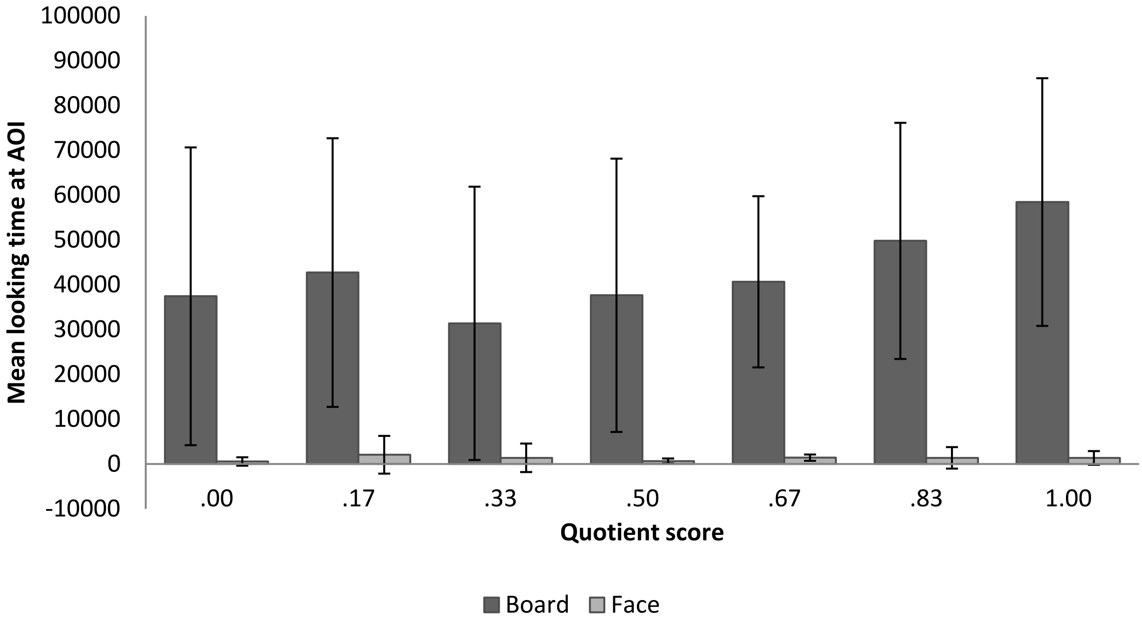


Figure 3.44: Mean looking time at each of the areas of interest (board, face) for participants achieving each of the possible quotient scores (.00 (N=7), .17(N=9), .33 (N=15), .50 (N=6), .67 (N=3), .83 (N=21), 1.00 N=44))

As indicated in figure 3.44, the number of participants in each quotient group was extremely variable. Initially, the intention was to analyse this data using an ANCOVA, with age group as a covariate. This analysis was performed, but results must be interpreted with caution. For the board AOI, there was no significant difference of age, F(1,104)=1.75, p=.189. There was no also significant difference between quotient scores: F(6,104)=.87, p=.520. For the face AOI, there was no significant effect of age: F(1,104)=.74, p=.391, nor was there a significant difference between quotient scores: F(6,104)=.435, p=.854.

In order to further investigate any relationship between quotient scores and mean looking times at each of the AOIs, Spearman's rho correlations were performed. For the board AOI, a moderate significant positive correlation was found between looking time at the board and quotient score: r_s=.32, p=.001, indicating that participants with higher quotient scores were more likely to attend more to the board. For the face AOI, a weak significant positive correlation was found between looking time at the demonstrator's face and quotient score: r_s=.21, p=.032, indicating that participants who spent more time looking at the demonstrator's face were also likely to have higher quotient scores.

The time spent looking at irrelevant information (i.e. the background) during the demonstration video was compared across age groups. The mean looking time for each age group at the background is indicated in table 3.24.

Table 3.24: Mean looking time at irrelevant (background) areas of the demonstration video for each age group

Age	N	Mean	Std. Deviation
3-5	46	72604	27617
6-8	30	51738	26436
9-11	19	48625	24669
12-17	11	63252	33762
18+	13	40883	24233
Total	119	59185	29197

The difference between age groups was found to be significant: F(4,118)=5.36, p<.001. Post-hoc tests indicated a significant difference in the mean looking time at the background between 3-5 year olds (M=72604) and 6-8 year olds (M=51738, p=.014), 9-11 year olds (M=48265, p=.016) and adults (M=40883, p=.003), with 3-5 year olds attending to the background of the demonstration video significantly more than older age groups.

3.9. Task Comparisons

The mean quotient score for each task is shown for participants in each age group in table 3.26.

Table 3.26: Mean quotient score for the box, hook and colouring tasks for each participant age group

Age Group	N	Box Quotient Score	Std. Deviation	Hook Quotient Score	Std. Deviation	Colour Quotient Score	Std. Deviation
3-5	43	0.39	0.222	0.32	0.290	0.44	0.329
6-8	30	0.58	0.259	0.53	0.290	0.75	0.298
9-11	19	0.52	0.300	0.46	0.291	0.80	0.312
12-17	20	0.50	0.247	0.62	0.214	0.94	0.995
18+	25	0.27	0.251	0.74	0.180	0.94	0.143
Total	137	0.44	0.270	0.51	0.300	0.72	0.335

For each age group, a Spearman correlation was performed to compare quotient scores across the three tasks, in order to determine whether participants' task performance was consistent, or whether they were able to use flexible copying strategies. These results can be seen in table 3.27.

Table 3.27: Correlations between quotient scores in each of the three tasks

	3-5 year olds			6-8 year olds			9-11 year olds			12-17 year olds			18+ year olds		
	Box	**Hook**	**Colour**	**Box**	**Hook**	**Colour**	**Box**	**Hook**	**Colour**	**Box**	**Hook**	**Colour**	**Box**	**Hook**	**Colour**
Box	1.00			1.00			1.00			1.00			1.00		
Hook	0.012*	1.00		0.08	1.00		0.001*	1.00		0.31	1.00		0.22	1.00	
Colour	0.20	0.16	1.00	0.18	0.14	1.00	0.45	0.16	1.00	1.00	0.10	1.00	0.47	0.06	1.00

For 3-5 year old participants, a weak positive correlation was found between their box and hook quotient scores. This was found to be significant: $r_s(44)=.376$, $p=.012$, indicating that participants who scored highly on the box task also achieved a high score on the hook task. There was a very weak positive correlation between 3-5 year old participants' box and colouring scores, but this was not significant: $r_s(40)=.204$, $p=.206$. Similarly, there was a weak positive correlation between their hook and colouring scores, but this was not significant: $r_s(41)=.227$, $p=.159$.

For 6-8 year old participants, a weak positive correlation was found between their box and colouring quotient scores. This was not found to be significant: $r_s(29)=.257$, p=.178. A weak positive correlation was also found between participants' hook and colouring scores, and this was again not found to be significant: $r_s(30)=.280$, $p=.142$, indicating that again, scores remained consistent between the hook and colouring tasks. Finally, a weak positive correlation was found between the box and hook tasks, but this was not found to be significant: $r_s(30)=.323$, $p=.082$.

For 9-11 year old participants, a strong positive correlation was found between their box and hook scores. This was found to be significant: $r_s(19)=.829$, $p<.001$, indicating that participants' scores remained consistent across these tasks. A weak positive correlation was found between participants' box and colouring scores, but this was not significant: $r_s(19)=.185$, $p=.447$. Finally, there was a weak positive correlation between participants' hook and colouring scores. This was not significant: $r_s(19)=.336$, $p=.159$, suggesting that participants' task performance was consistent across these tasks.

For 12-17 year olds, a moderate positive correlation between participants' hook and colouring scores was found. This was not found to be significant: $r_s(19)=.386$, $p=.103$. There was also a very strong positive correlation between participants' box and colouring scores, but again, this was not significant: $r_s(20)=1.000$. Similarly, there was a weak positive correlation between participants' box and hook scores, but this was not significant: $r_s(20)=.241$, $p=.306$.

For adult participants, there was a weak positive correlation between their box and hook quotient scores, but it was not significant: $r_s(25)=.235$, $p=.222$. Similarly, there was a weak positive correlation between the box and colouring quotient scores, and it was not significant: $r_s(25)=.153$, $p=.465$. Finally, there was a weak positive correlation between the hook and colouring scores, and it was not significant: $r_s(25)=.381$, $p=.060$. With no significant correlations between any of the task scores, adults were the group that appear to have displayed the most variable task performance, since high or low performance on one task did not appear to be related to a similar performance on either of the two remaining tasks.

Participants achieved the highest overall mean quotient score in the colouring task, followed by the hook task, and then by the box task, despite the fact that the colouring task had the highest number of actions. This is indicated in figure 3.45. A repeated measures ANOVA was used to further investigate these differences: There was a significant effect of age: $F(1,127)=14.11$, $p<.001$, and a significant effect of task: F(1,127)=107.35, p<.001. Additionally, there was a significant interaction between age and task: F(1,127)=15.93, p<.001.

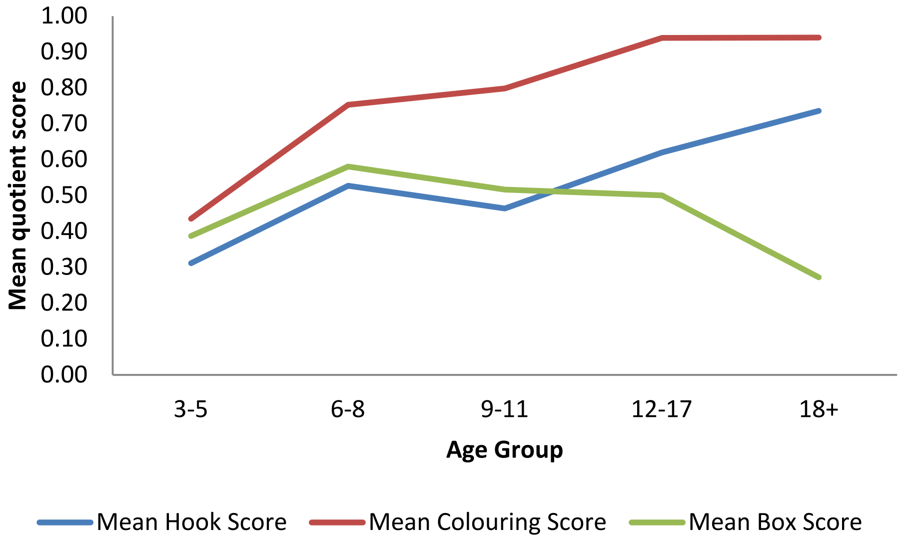


Figure 3.45: Mean score on the hook, colouring and box tasks for participants in each of the assigned age groups (3-5, 6-8, 9-11, 12-17, 18+)

Interestingly, the box task does not follow the same pattern of an increase in overimitation across age groups. In fact, adults were the least likely group to copy faithfully in this task, in contrast to the hook and colouring tasks, in which they achieved the highest scores.

The order in which participants completed each of the tasks was then investigated. The percentage of participants in each age groups that completed each of the actions in the tasks in the same order as the demonstrator is indicated in table 3.26.

Table 3.27: Percentage of participants in each age group that completed each of the tasks in order

Age Group	Box	Hook	Colour
3-5	67	37	32
6-8	97	67	63
9-11	84	74	68
12-17	95	90	70
18+	80	96	76
Total	82	66	57

As indicated in table 3.27, participants in the 3-5 year old age group were least likely to copy the order of actions across all three tasks. However, participants in all child age groups were most likely to copy the order of actions in the box task. This is likely to be the case as imitating the order of relevant actions only involved copying two actions in the box task, whereas in the hook and colouring task, more actions were involved, meaning that it was easiest to copy the order of (relevant) actions in the box task. In the box task, imitation of the order of relevant actions remained high across all age groups other than the 3-5 year old group, whereas imitation of the order of actions in the hook and colouring tasks increased with age, likely due to the difference in complexity. Interestingly, the colouring task showed the lowest percentages of participants imitating the order shown in the demonstration video. This may be due to the complexity of imitating this task in order. Imitating the order of relevant actions in the box and hook task involved a small number of actions (right lever followed by left lever in the box task, and building the hook from bottom to top in the hook task), whereas in the colouring task, participants had to imitate four actions in order.

3.10. Discussion

This experiment aimed to build on findings from experiment one, in which 3-5 year old participants were provided with specific instructions in order to induce copying behaviours in two tasks. Participants in the video control condition, in which no specific instructions were given, appeared to behave differently, depending on the task, with participants being more overimitative in a puzzle box task, and imitative in a puzzle board task. Additionally, task performance differed between age groups, with adults achieving higher mean quotient scores than children, indicating more faithful copying. In order to further understand why these differences may have been present, a number of aims for this experiment were identified:

- To use multiple tasks to investigate context-dependent flexible copying behaviours
- To once again use eye tracking in order to determine what participants attend to during a demonstration, with particular reference to affordances of the puzzle tasks, as well as the demonstrator’s movements when completing the tasks.

- To determine at what age participants' copying behaviours become more adult-like, and to see whether a developmental trajectory in copying behaviours in participants aged 3-21
- To further investigate underlying developmental factors, such as memory and theory of mind, that may subsequently impact participants' copying strategies.

3.10.1. Summary of findings

- Theory of mind scores increased with age, with all participants above the age of nine passing the task.
- Sequence memory scores increased with age across all three tasks.
- Content memory tasks were too easy, meaning that participants' scores quickly achieved ceiling levels.
- The mean quotient scores increased with age for the hook and colouring tasks. In the box task, the lowest quotient scores were observed in 3-5 year olds and adults, and the highest quotient scores were observed in 6-8 year olds.
- As in experiment one, there were no clear gaze patterns that could be linked to higher or lower quotient scores.

3.10.2. Developmental abilities

Theory of Mind

As expected, participants' performance on the false belief task increased with age until participants reached ceiling performance around the age of nine. This is consistent with previous findings, which show an improvement in this ability up to the age of 9 (Wimmer & Perner, 1983a). Despite the clear trajectory observed in ToM performance, there was no relationship observed between improvements in theory of mind and a shift in copying behaviour with reference to any of the tasks. These findings suggest that participants' ability to understand the demonstrator's intentions was not driving their subsequent task performance as expected.

Memory

Participants' performance on both the sequence and content memory tests across all three tasks rose with age, as expected, again reaching ceiling around the age of nine in most cases. This is with the exception of the content memory test that related to the box task, in which participants in all age groups performed at ceiling, indicating that it was too easy, as it would be expected that 3 and 4 year old participants should perform poorly on this due to their developing episodic memory abilities (Perner & Ruffman, 1995). One potential issue with the method used to test participants on their memory of the content of the demonstration video was that the false images used in this test were too different from the content of the demonstration video. As still images were used for this test, it was difficult to include any false images that were not very obviously different. Future research should investigate other ways of measuring this ability, perhaps by using more subtle variations of some of the actions used. Across all three tasks, participants with the lowest quotient scores in each task performed poorly in the sequence memory tests. This may suggest that participants did not emulate through choice, but instead because they were unable to remember what they had observed during the demonstration video, and had instead focused on the goal state. This is consistent with findings in younger children by Loucks & Meltzoff (2013), who determined that these participants had a tendency to prioritise (and therefore remember) information relating to the overall goal as opposed to arbitrary sequences of relevant or irrelevant actions. This may suggest that humans are naturally goal-focused, but may copy other information if they have the capacity to do so. Almost all adult participants, with the exception of one, chose to copy the demonstrator precisely during the colouring task, in terms of the irrelevant drawing and the colours used to fill the shapes. The adult participant who did not copy the demonstrator indicated to the experimenter that this was a conscious choice; she was aware of the colours used by the demonstrator but wished to complete hers in a different manner. It is also possible that participants' performance was affected by working memory, particularly with the hook task, which was complex and required completion of a number of steps. This is reflected in the results obtained, which indicate that the youngest participant group achieved the lowest sequence memory score across all three tasks. This age group also achieved significantly lower quotient scores, and were less likely to use the same order as the demonstrator when completing tasks, which suggests that their developing ability to remember actions observed and the order in which they occurred has impacted their performance, potentially suggesting that the development of working memory plays a role here. Indeed, Subiaul & Schilder (2014) suggest that working memory constraints in young children have an impact on both the type and amount of information copied. Since sequence memory appears to have had an impact on participants' performance across the tasks, it would therefore be beneficial to investigate participants' working memory ability in conjunction with the copying strategies used by participants in future research, to determine whether improvements in working memory have an impact on participants' copying fidelity.

Attention

Across all three tasks, the total mean time participants spent attending to the demonstration video as a whole increased with age, as expected. This is consistent with the suggestion that the ability to sustain attention to a task improves with age (Hagen & Hale, 1973; Roebers et al., 2010), and is also consistent with findings from experiment one. The ability to attend to the demonstration video is important, as in order to have the ability to copy, participants must be able to attend to the demonstration video for long enough in order to see the relevant and irrelevant actions taking place.

3.10.3. Behavioural performance on tasks

Quotient Scores and Task Type

As expected, 3-5 year old participants achieved the lowest quotient scores across all three tasks. As previously discussed, participants in this age group also achieved the lowest memory scores, and were the most likely to fail the theory of mind task, suggesting that these developmental features do have an impact on participant performance. This finding also contradicts the suggestion that the youngest participants copy actions because they do not understand their functions, as in this case we would expect to see 3-5 year olds achieving the highest task scores.

Adult performance was variable across the tasks, suggesting that adult participants were using flexible copying strategies depending on the task type. Adults achieved a low quotient score in comparison to participants aged 6-17 on the box task. This is inconsistent with previous findings, which suggest that adults should be highly overimitative, even at the expense of efficiency (Flynn & Smith, 2012; McGuigan et al., 2011b). However, adults were found to have the highest quotient score on both the hook and colouring tasks. This suggests that inherent differences in task types influence adults to use different copying strategies. This may relate to the type of actions shown in the demonstration video, and is discussed further below.

Participants in the remaining age groups (6-8, 9-11 and 12-17) showed at least moderate copying fidelity, achieving a score of .50 or above in all three of the tasks. Interestingly, participants aged 6-17 all achieved a score of between .50 and .60 on the box task, in comparison to the score of .27 achieved by the adult group. However, in the hook and colouring tasks, copying fidelity increased steadily with age, meaning that imitation precision did increase with age. According to Gardiner, Greif, & Bjorklund (2011) this would be expected, as younger children have been observed to copy more general aspects of a demonstration, whereas older participants tended to copy both broad and precise elements. This may be due to increased episodic memory ability; participants who achieved a higher sequence memory score were more likely to use the same pieces as the demonstrator, possibly because they were able to remember not only the basic sequence and shape used, but also individual elements of the task (e.g. specific pieces used in the hook task, or the colour used for each shape in the colouring task). These differences in task performance are in some way consistent with previous research (Keupp et al., 2015b), which suggests that children assess the rationality of actions when overimitating, and choose to imitate selectively if the outcome meets their intended goal (Over & Carpenter, 2012).

The arbitrary nature of the colouring task also indicates that overimitation is not purely goal-driven, which may add support to underlying theories relating to the social nature of this behaviour, such as that suggested by Over & Carpenter (2012), as opposed to purely causal or normative theories which suggest that overimitation is an inflexible process (Derek E Lyons et al., 2007). It could be argued that overimitation was not strictly measured within this task, as the “irrelevant” drawing was not necessarily irrelevant to task completion, depending on participants' interpretation of the goals. Therefore, future research could potentially include obviously irrelevant actions, such as hand movements, added to the process of completing the picture in order to further understand whether overimitative behaviour was indeed being observed in participants. This would also allow for better comparison with the previous tasks, which involved physical movements or actions as the irrelevant aspects of the task.

Action Type

Participants of all ages were less likely to copy actions that appeared to be obviously irrelevant (e.g. the knock action in the box task, and the tap action in the hook task). When actions appeared to be potentially functional (e.g. the handle in the hook task, and the top lever in the box task), then they were copied by the majority of participants in all age groups.

Few participants in any of the age groups copied either the knock or turn irrelevant actions in the box task. Additionally, the majority of participants in all age groups other than the adult age group were found to copy the irrelevant lever action. These findings are consistent with those from experiment one, in which adults in the control condition displayed more similar behaviour to those in the goal focused group, rather than the copying group.

In contrast to findings by Hernik & Csibra (2009), younger participants did not imitate the irrelevant tap action any more than participants in other age groups. However, the tendency to include the irrelevant handle did rise with age, with adult participants being most likely to include the handle when building the hook. This is consistent with findings by Flynn & Smith (2012) and Whiten et al. (2016), which suggest that during social learning tasks, adults overimitate more than children, even at the expense of efficiency. The imitation of the handle action may have occurred as the demonstrator was seen to successfully complete the task, and therefore participants misinterpreted the inclusion of the handle as an efficient method to solve the problem. This is consistent with the suggestion that people overimitate in order to avoid wasting time and effort in developing new solutions when a viable one is already available (Legare & Nielsen, 2015). This is linked to the concept of cumulative culture, in which information is transmitted socially that allows a ratcheting effect to occur in which each generation builds on the work of the previous one by increasing the complexity of an object or tool (Dean et al., 2014; Tennie et al., 2009). In this way, it is much more efficient to overimitate a minor detail rather than to reinvent the hook itself from scratch. Therefore in this case, it could be argued that although the inclusion of the handle itself is inefficient, adopting a previously successful strategy is actually a more efficient method overall than trying to develop a completely novel tool.

These findings suggest that many participants were displaying selective copying, although the reasoning behind their choice to copy some, but not all, irrelevant actions is not clear. As there is some suggestion that overimitation occurs when the function of an object, or the goal of a task is not clear (Hernik & Csibra, 2009; Keupp et al., 2015), then it may be beneficial to explore participants' understanding of the goal of the task, and the role that the irrelevant actions play in completion of the task. This would allow for further exploration of the motivations (either social or goal-driven) behind overimitation, and would also allow for differentiation between participants who understand that the top lever is unnecessary for task completion, as opposed to those who believe that it is required.

As previously discussed, the "irrelevant" action included in the colouring task was not obviously irrelevant, and may have appeared to be a relevant part of the task to participants. Indeed, as with other actions that appeared relevant, copying of the drawing increased with age.

Order of Actions

Participants were most likely to copy the order of actions from the demonstrator in the box task. However, it is important to note that this refers only to the relevant actions (the left and right levers), rather than all five actions demonstrated. In this task, 80% of all participants aged 6-17 copied the order of relevant actions from the demonstrator. However, the majority of 3-5 year old participants did not copy order. This may be linked to their sequence memory ability, as they also achieved very low sequence memory scores, as discussed previously. It may therefore be the case that whilst irrelevant actions were not deemed rational, and so were discarded, imitation of the order of relevant actions was appropriate as they were seen to be successful during the demonstration video, and so using an alternative strategy would have been less efficient (Keupp et al., 2016a; Southgate, Chevallier, & Csibra, 2009)

In the hook and colouring task, copying of the order of actions increased with age, also in line with sequence memory abilities. It is also of note that participants were most likely to use the same colours as the demonstrator for the first and last shapes, with a lower percentage of participants using the same colours as the demonstrator for the two middle shapes (the triangle and the drawing) , despite the colouring video being on screen for some time. Additionally, the oval shape (the final shape to be coloured) took a long time to complete, meaning that participants had time to mentally rehearse the colour of each of the three previous shapes if they had wished to. The pattern of results seen in this task appears to suggest evidence of both a recency effect, in which an individual is more likely to remember the last item in a sequence, and a primacy effect, in which an individual is better able to remember the first item in a sequence, as they have had time to transfer it to long term memory through the use of rehearsal strategies (Postman & Phillips, 1965).

3.10.4. Eye tracking tasks

The aim of using eye tracking in this research was to determine whether developmental changes could be observed in participants' gaze patterns, and whether there was any relationship between these changes in looking patterns and a shift in copying fidelity. This would help to determine at which stage influences in copying fidelity are apparent: in the "attention phase" (i.e. during the demonstration itself), or afterwards, as participants complete the task for themselves.

Differences in participants' mean looking time at the levers in the box task, and at the board in the colouring task, were found, in which adults looked longer at these AOIs in comparison to child participants. This is consistent with the suggestion that the ability to attend to relevant stimuli increases with age. However, as in experiment one, no differences in gaze patterns were observed between participants achieving each of the different quotient scores, suggesting that participants' tendency to copy was not influenced by the information they had attended to during the demonstration. This is in direct contrast to previous findings by Brubacher, Roberts & Obhi (2013), who determined that participants were more likely to imitate an action that was "gazed-at", in other words, viewed during a demonstration. However, some of the issues in identifying specific gaze patterns may have occurred as a result of the tasks themselves. In the colouring task, participants did not need to sustain their attention throughout the entire video in order to be able to successfully copy all aspects of the task. This was for two reasons: participants did not need to watch the demonstrator filling in the entire shape, they could simply look at any point during this time to observe what colour the demonstrator used, and in which order she was completing the picture, and once the shapes were coloured, they remained visible on-screen for the rest of the duration of the video. This was reflected in the results observed for participants who watched the demonstrator draw the stick figure, and the number of participants who did actually draw the figure themselves; the percentage of participants who drew the figure was higher than the percentage who watched the demonstrator draw it. For this reason, it may have been more appropriate to measure the order in which participants drew the figure: was it in the same order as the experimenter, or did they produce the same shape, but using a different order? It was not possible to discern this information from video recordings of participants, but future research could investigate this copying precision further to determine whether differences were present.

An additional limitation was the difficulty in identifying fixation locations during demonstration videos which involved very small components that frequently moved; this was particularly true for the hook task. This meant that the potential for human error was high. Future research should find alternate methods for identifying dynamic areas of interest within a demonstration video with multiple moving objects, in order to increase the precision of information captured.

It was predicted that participants who looked at the demonstrator's face prior to the start of the demonstration may be more likely to be classed as overimitators due to an increased feeling of affiliation with the demonstrator, and this was not found to be the case. There was no significant difference observed between groups in terms of the mean proportion of time spent looking at the demonstrator's face in any of the three tasks. This contradicts previous research, which has indicated that participants are more likely to imitate actions if they have displayed increased looking patterns at the demonstrator's face (Vivanti et al., 2017). This is due to the fact that participants are better able to understand the goals of the demonstrator (Carpenter & Call, 2007). One issue with the colouring task in particular may have been the familiarity of the task to participants. Since this task was likely to be already familiar to participants, they may have felt it unnecessary to focus on the demonstrator's face for further clues about how to correctly complete it (Hobson & Hobson, 2007). In all tasks, the demonstrator's face was directed towards the task as opposed to towards the participant, which may have also affected participants' tendency to look at it.

3.10.5. Task comparisons

The inclusion of three tasks within this experiment allows for comparisons to be made across task types in terms of copying strategies used. In contrast to previous studies that report high levels of overimitation in puzzle box tasks (Lyons et al., 2007), levels of overimitation were lowest for the box task for participants in all age groups. The order in which participants copied actions from the demonstrator increased with age in the colouring and hook tasks, but not in the box task. However, this measurement was based on the relevant actions from the box task, which meant that only two actions were involved, in comparison to the multi-step processes involved in both the colouring and hook tasks, which may have been impacted by working memory ability (Subiaul & Schilder, 2014). As all three tasks involved the same demonstrator, and therefore did not differ in terms of any social components, it is likely that the changes across tasks were due to developmental factors as opposed to social ones. However, it is interesting that the behavioural pattern seen in the box task is so markedly different to that seen in the hook and colouring tasks, particularly when puzzle boxes are one of the most commonly used tasks in the social learning literature. This indicates that the box is somehow different to the other two tasks, and so may not be the best measure of differences in copying strategies between groups. In this study, participants were not asked about their understanding of the intended goal of the task, which may have revealed where these differences lay. Future research should investigate this, to determine why the behaviour seen in the box task was so unexpected, and to determine whether this pattern of behaviour can be replicated. Additionally, the hook task was arguably the most complex of the three tasks, with the largest number of potential outcomes, allowing participants to be more flexible when completing the task. This is reflected in the task scores seen: although the quotient scores for the hook task rise with age, even adults only achieve a mean score of .74, in comparison to .94 on the colouring task. This could either suggest that adults were using a more efficient strategy than that demonstrated to them, therefore reducing their mean score, or that copying the task precisely was very complex, particularly when choosing the precise pieces to use. It would therefore be interesting to determine whether adult participants' scores change when they are provided with fewer pieces to create the hook with. This would still allow participants the opportunity to be creative, but would also provide a way to test whether the number of available pieces increased the complexity too much, therefore making the hook more difficult to copy. Another way to achieve this may be using colour coded pieces, to see whether participants match both colour and shape, or colour or shape only.

As in previous chapters, developmental differences in ability and task performance were observed as expected. However, once again, gaze patterns were not sufficiently strong enough to allow for the identification of participants of particular copying types based on eye tracking data alone. This suggests that the changes in copying fidelity observed across age groups are likely to be driven by developmental factors, such as the ability to remember sequences and actions from the demonstration video. Additionally, it is also possible that participants made a conscious choice to either copy or omit actions when completing tasks, based on their understanding of the functionality or purpose of the task. As it was not possible to determine whether this was the case from the data collected in this experiment, further research is required to investigate this. Therefore, experiment 3 will focus on participants' understanding of the goal of the task and the demonstrator's intentions, as well as developing working memory ability and its relation to copying fidelity.

Chapter 4: Experiment 3

4.1. Introduction

4.1.1. Previous findings

Experiments one and two made use of a puzzle box task, in which participants were required to operate levers to retrieve a reward. In experiment one, participants in the copying and goal-focused groups were provided with instructions prior to completing the puzzle box, but participants in the video control group were only told “now it’s your turn”. Of these video control participants, only 7% of adults were found to achieve a quotient score of .80 or above, which was inconsistent with previous research that suggested that adults should be naturally overimitative. However, in experiment two, in which participants were provided with the same instruction prior to completing the puzzle box task, 44% of adults achieved a quotient score of .80 or above. It is unclear as to why these differences occurred; although an extra irrelevant action was added to the task in experiment two, this should not have affected the percentage of overimitators so drastically; in fact, this should have made the task more difficult, as more actions had been added to the sequence. In both experiments one and two, adult participants were all undergraduate students studying psychology at Newcastle University, and were all given the same instruction, therefore it would be expected that their behaviour during this task should be consistent across both experiments. Additionally, adult participants were excluded from taking part in experiment two if they had participated in experiment one. This experiment was intended to further investigate both the inconsistencies in copying behaviours between experiments one and two, and to further understand participants' interpretation of the overall goal of the task, and how that subsequently affects their behaviour.

4.1.2. Box Task

The box task consists of a transparent box with functional levers on the left and right, and an irrelevant lever on the box lid. The functional levers have small platforms attached to the end inside the box. Prior to completion of the task, a small plastic ball is placed onto one of these platforms, meaning that the functional levers can then be operated to release the ball into a chute at the bottom of the box. In order to assess overimitative behaviour, irrelevant actions for task completion were also added to the demonstration video shown to participants before they were given the box task to complete. In experiment one, two irrelevant actions were used: knocking on the lid of the box, and operating the non-functional lever attached to the box lid. The number of relevant and irrelevant actions copied by the participant could then be measured. In experiment two, an additional irrelevant action, turning the box 360° clockwise, was added. Unlike in experiment one, all participants were given the instruction "now it's your turn", with no further instructions provided, meaning that it was expected that participant behaviour would match that of the video control groups in experiment one, who received the same instruction.

4.1.3. Interpretation of task goal

As previously described, no measure was used in either experiments one or two that allowed for assessment of participants' understanding of the goal. This was primarily due to time constraints required as a result of testing very young participants; it was felt that adding any additional measures, especially to experiment two, would be too much for the youngest participants. However, participants are likely to behave very differently if they think the goal is simply to retrieve the plastic ball, rather than if they understand the goal to be to copy all the actions in the correct sequence (i.e. a memory task). Alternatively, participants may wish to perform the task in the same way as the demonstrator, independently of the actions and goal involved (i.e. interpreting it as an affiliative task). One way to assess this may be to ask some debriefing questions after participants have completed the task, to determine what they understood about its goal and overall function.

Williamson & Markman (2006) investigated 3 year old participants' understanding of the goal of a task, using a series of actions performed on novel objects, some of which had an obvious goal, and some of which did not. It was determined that when participants did not understand the goal of a task, they were more likely to copy the demonstrator precisely. However, when participants did understand the task goal, they appeared to be more confident in using their own method or strategy to achieve it. To determine whether this is also the case with older participants, an additional manipulation will be added to this experiment, in which the demonstrator performs the same actions on the puzzle box as in experiment two, but with the reward removed, meaning that the final goal state is less obvious to participants. It is expected that participants in the "no goal" condition will overimitate more than those in the condition in which the reward is present, due to the fact that they may not understand the overall purpose of the task, in line with findings by Williamson & Markman (2006). This would also be consistent with findings from the colouring task in experiment two, in which participants were found to overimitate more in comparison to the box and hook tasks. This is potentially because the goal state was less obvious, and participants therefore interpreted the actions required to complete the task (e.g. colouring the shapes in order, with the same colours), to be the goal.

However, it is also important to consider that some participants, especially the youngest children, may believe that the irrelevant actions in the goal condition are necessary for task completion. This is known as the "automatic causal encoding" theory (Lyons, Young, & Keil, 2007), and may therefore result in children copying all actions because they believe that they are required in order to retrieve the ball. It is therefore important to question participants not only about whether they believe that they have copied all of the actions shown by the demonstrator, but also whether they believe that all of those actions are necessary. Allowing participants to attempt the puzzle multiple times may also provide further information about what is deemed to be absolutely necessary, and what can be omitted whilst still producing a successful outcome, through trial and error. For this reason, participants will be given the opportunity to complete the puzzle once, as in experiment two, and will then be asked to complete the puzzle a second time, under the guise of demonstrating it to another individual. It is expected that if a participant believes that the irrelevant actions truly are necessary, then they will perform it both times. However, some participants may determine that the irrelevant actions are not necessary after performing them in the first instance, and so may then choose to omit them during their second attempt.

4.1.4. Developmental factors

An additional factor that may have affected participants' performance is working memory. Although there was a relationship observed between content and sequence memory and task performance in experiment two, the delay between participants watching the demonstration and completing the puzzle task was extremely short. This meant that it could not be considered a test of episodic memory, which normally involves a substantial delay of anywhere from 30 minutes to 24 hours (Nakano & Kitazawa, 2017; Óturai, Kolling, & Knopf, 2013). However, a relationship was found between performance on the sequence memory tasks and copying fidelity. For this reason, it may be beneficial to additionally assess participants' working memory, due to the fact that the box has multiple task demands and is likely to particularly affect younger participants' performance. Since participants in this experiment will be presented with the puzzle box immediately after viewing the demonstration video for the third time, they should not be affected too heavily by working memory decay as a result of a time delay (e.g. Barrouillet et al., 2004). However, they may be affected by increased cognitive load if the number of actions involved in completing the puzzle exceeds their recorded digit span. Previous research has suggested that children are indeed more likely to emulate when working memory load is high (Subiaul & Schilder, 2014). In order to control for participants' working memory ability, a condition in which the demonstration continues to play on a loop as participants complete the task will be added, meaning that participants are able to refer back to the demonstration at any time. If working memory is a factor in overimitation, then the mean number of irrelevant actions copied should be higher in conditions where the demonstration video is looped in comparison to those in which participants only see the demonstration three times.

As in experiment two, a theory of mind measure will be administered, based on a false belief task used by Astington & Jenkins (1999). Although it is not expected that the use of this test will add to findings obtained in experiment two, participants' theory of mind ability may have an effect on their task performance, particularly if they interpret the task goal to be an affiliative one. Previous research suggests that social cues and affiliation are factors in participants' choice to copy a demonstrator (Hoehl et al., 2014; Nielsen & Blank, 2011), and individuals with Autism Spectrum Conditions have been determined to display reduced overimitation, potentially due to a deficit in theory of mind ability (Vivanti et al., 2017).

Due to the fact that aspects of this experiment will be quite heavily language-dependent (particularly the debrief questions required to assess participants' understanding of the goal, as well as the digit span task used to assess working memory), then 3 year old participants will not be included in this study. It was initially intended to include adolescent participants in this research, in order to allow for comparisons to be made with findings from experiment two, however, due to time constraints and recruitment difficulties, this age group was not included. Therefore, an attempt has been made to keep child age groups as close to those in experiment two as possible, and for this reason, participants will be divided into the following age groups: 4-5 year olds, 6-8 year olds, 9-11 year olds and adults. One potential limitation identified with regard to the adult age groups used in experiments one and two was that they were both entirely composed of undergraduate psychology students, aged 18-21. This means that results are unlikely to be comparable to a number of other social learning studies, which have used a wider age range of adult participants. For this reason, the age range of adult participants was expanded to 18-45 year olds, and an effort was made to recruit participants from the general population, alongside undergraduate students.

4.2. Method

4.2.1. Participants

107 typically developing children, aged 4-11, were recruited from the local area via social media and word of mouth. 42 adult participants, aged 18-45, were recruited via the Newcastle University participant pool and via word of mouth. Child participants received a certificate and either a small toy or a £5 voucher in return for their participation. Adult participants received either course credit (for those who were undergraduate students) or a £5 shopping voucher. Prior to completion of the research, written consent was obtained from parents (for children) or from participants themselves (for adults). For the purpose of analysis, participants were divided into the age groups indicated in table 4.1.

Table 4.1: Mean ages for each assigned participant age group

Age Group	N	Mean	Std. Deviation
4-5	44	4.48	0.505
6-8	33	6.82	0.882
9-11	30	9.43	0.626
Adults	42	26.60	7.651
Total	149	12.23	10.054

All adult data were analysed together. All participants were free to withdraw their participation at any time.

4.2.2. Materials and apparatus

Theory of Mind

Participants completed the same theory of mind task as used in experiment two. This was a false belief task, adapted from Astington & Jenkins (1999). In this task, participants were shown the image in figure 4.1:

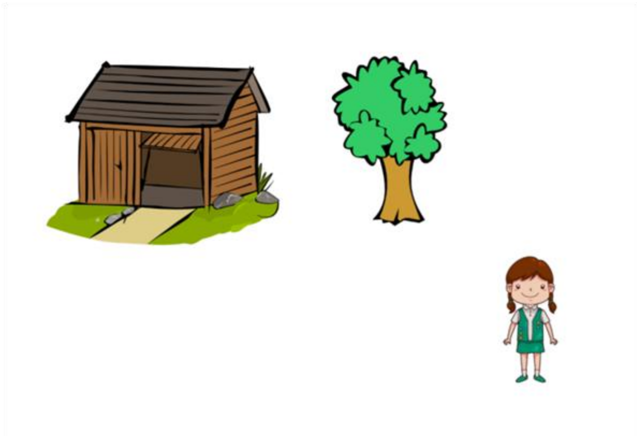


Figure 4.1: False belief task image shown to participants

The same procedure for the theory of mind task was used as in experiment two; see chapter 3 for further information on instructions provided to participants.

Digit Span

All participants completed a forward digit span task, designed to assess their working memory capability, adapted from the Wechsler Intelligence Scales for Children-Revised (WISC-R). The task used can be found in Appendix A. Participants were asked to repeat strings of numbers, which increased in length until the participant was no longer capable of repeating all of the numbers in the correct order. Their score was the length of the last sequence that they had been able to repeat in full (e.g. correctly repeating the sequence 8-4-2-3-9 would result in a score of five).

Box Task

The box task used was identical to that used in experiment two. The box itself measured 20cm tall, 30cm wide and 16cm deep (see Figure 4.2.2), and contained levers on the left and right hand sides which could be operated in any order to release a reward. Both the internal and external mechanisms of the task were visible to participants throughout the demonstration video. Three external levers operated the internal mechanisms of the box. On the left hand side, a wheel could be rotated either forwards or backwards to turn the platform containing the reward. On the lid of the box, a distractor lever could be moved up or down but this had no effect on the outcome of the task. On the right hand side of the box, a lever could be moved up or down, which moved the attached platform up and down. Upon moving the left and right levers, the reward would be tipped down a chute at the bottom of the box. This chute was designed to be small enough that participants could not use it to retrieve the reward without using the levers by using their hand to reach into the box. The reward consisted of a small plastic egg which could be swapped with the experimenter for a sticker. The lid of the box was held closed using a strong magnet, and a key was required to release the lid to reset the task. This key was not accessible by participants. The purpose of this was to prevent young participants from simply opening the box to retrieve the reward.

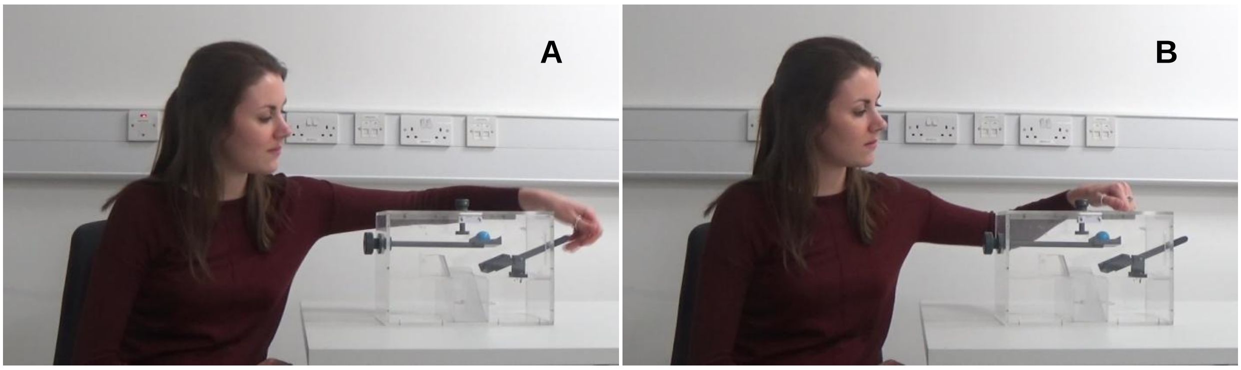
Figure 4.2: Box apparatus used in both demonstration video and task

Demonstration video

Unlike experiments one and two, participants did not undergo eye tracking during this study; only behavioural data was collected. Although the actions performed in the demonstration video remained the same as those in experiment two, a new demonstration video was recorded in order to increase the length of time in between the actions performed. It was felt that the previous demonstration video was too quick for younger participants, who may require increased processing time (Kail, 1993), and was therefore putting them at an unfair disadvantage in comparison to older participants. As before, a female demonstrator was used for this task (the same demonstrator as experiment two). An additional variation of the video was recorded, in which the task goal (the plastic egg) had been removed from the box; this is referred to as the no goal condition. These videos are described in more detail below.

Goal-focused demonstration video

This video can be viewed by clicking on the following link (password: demo). The demonstrator faces the camera at the beginning of the video, in order to make eye contact with the participant to signal that the demonstration is about to take place. The demonstrator first tips the lever on the right hand side of the box to move the lower platform into a downwards position. This does not alter the position of the reward. She then knocks on the lid of the box – an unnecessary action for task completion. A second unnecessary action involves the demonstrator pulling the lever attached to the lid of the box, which does not have any function in solving the task. The demonstrator then performs a third unnecessary action, turning the box around 360º clockwise.. Finally, the demonstrator rotates the wheel on the left hand side of the box, tipping the platform that releases the prize, which falls onto the lower lever, and then downwards through the chute at the bottom of the box. The demonstrator picks up the egg, and then faces the camera again and smiles, whilst holding the reward (see Figure 4.3).



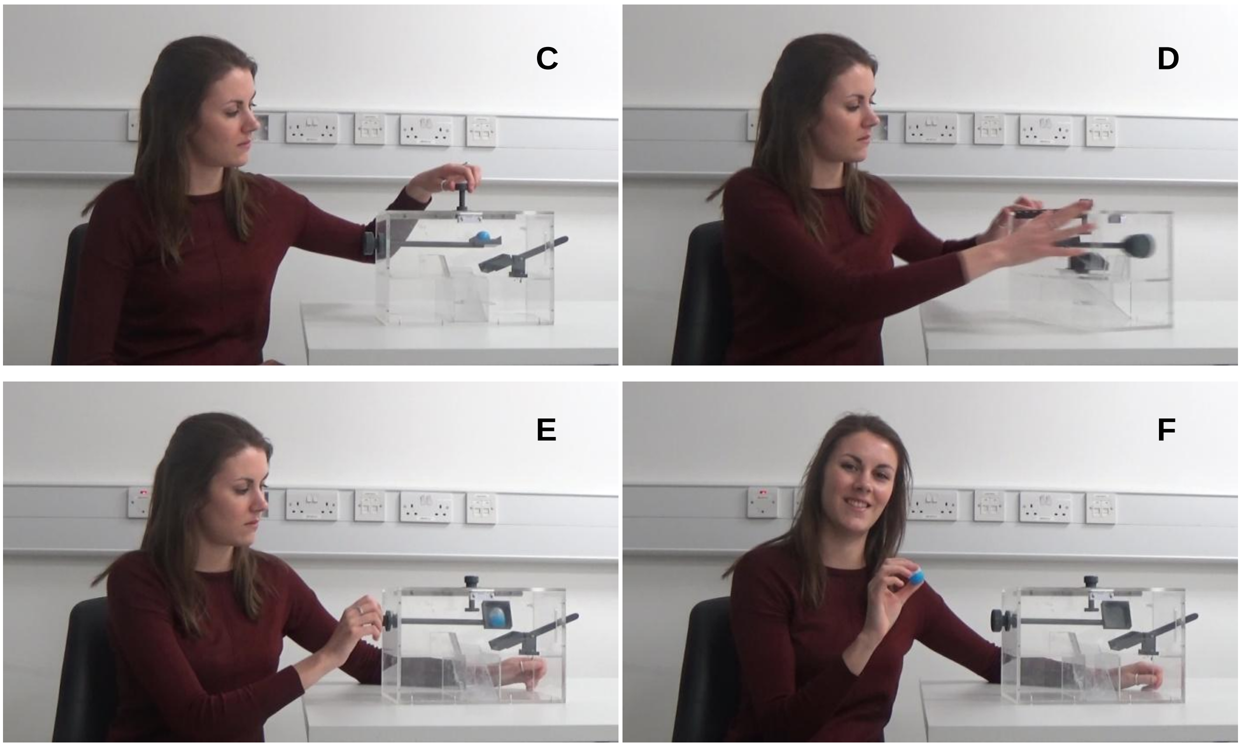


Figure 4.3: Actions included in goal demonstration video. A: The demonstrator tips the right lever into a downwards position. B: She knocks on the lid of the box, just above where the ball is situated. C: She lifts the irrelevant top lever, and lets go so that it drops back into place. D: She turns the box around 360 degrees clockwise, so that at the end of the action, the box is back in its original position. E: She turns the left lever clockwise, to tip the ball onto the platform of the right lever. The ball then falls down the chute. F: The demonstrator picks up the ball, turns to the camera and smiles.

No goal demonstration video

This video can be viewed by clicking on the following link (password: demo). The actions are performed in the same order as in the goal-focused demonstration, but no plastic egg is present, and therefore cannot be retrieved by the demonstrator after performing the actions. As with the goal-focused demonstration, the demonstrator faces the camera and smiles at both the start and the end of the video (see Figure 4.4.4).

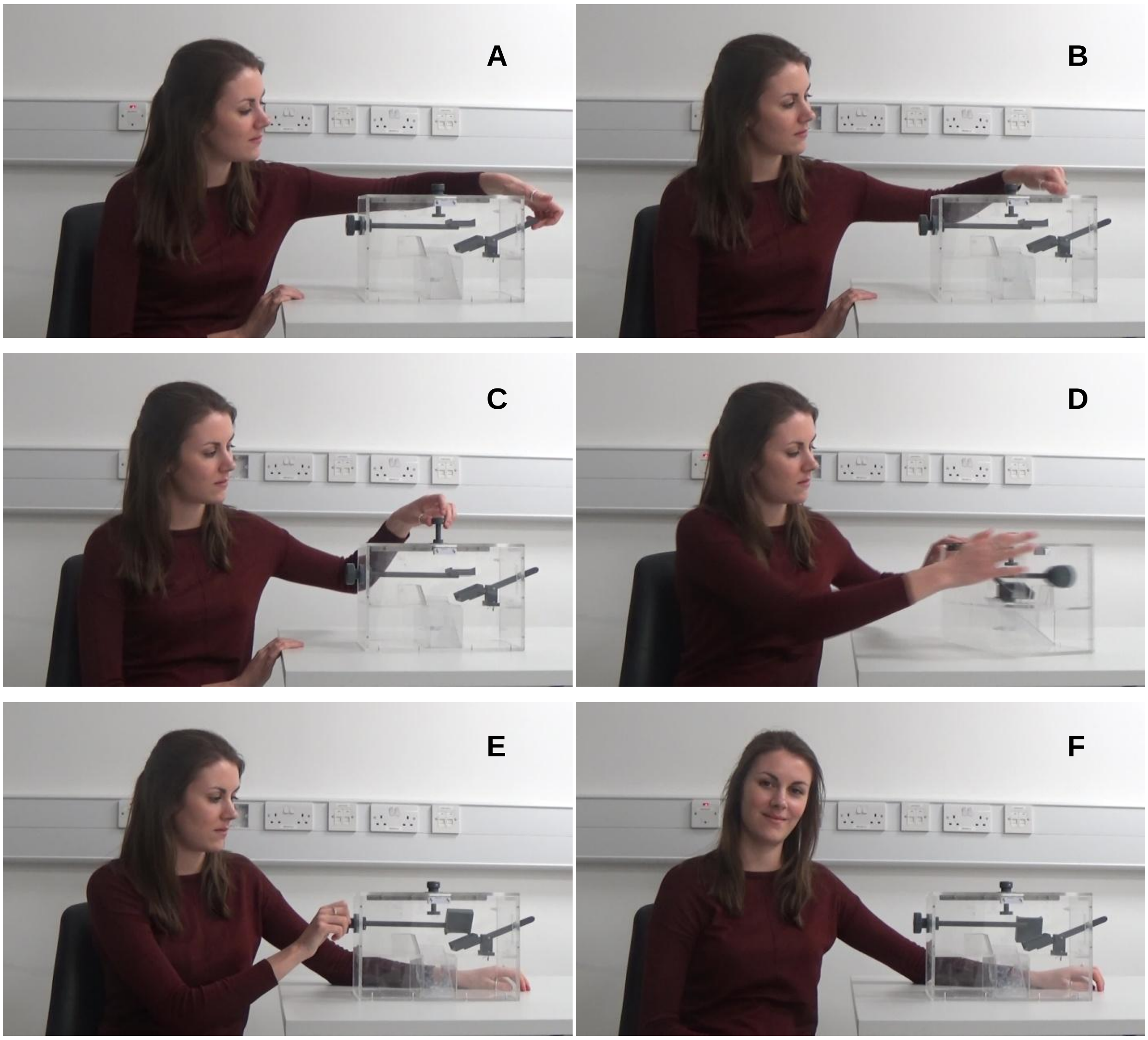


Figure 4.4: Actions included in no goal demonstration video. A: The demonstrator tips the right hand lever into a downwards position. B: She knocks on the lid of the box, above the platform of both levers. C: She lifts the irrelevant top lever and lets go so that it falls back into its original position. D: She turns the box 360 degrees clockwise, so that at the end of the action, the box is returned to its original position. E: She turns the left lever clockwise, so that the platform tips. She then turns the lever anticlockwise so that it tips the opposite way. F: She turns to the camera and smiles.

In order to more accurately assess the underlying features contributing to participants' copying performance, they were assigned to one of four conditions.

Goal-focused (x3)

Participants viewed the goal-focused demonstration three times, and were then given the puzzle to complete. This condition was designed to serve as an opportunity to replicate the results obtained in experiment two as the procedure was the same. It was therefore expected that children in this group would overimitate more than adults, with the highest levels of overimitation being observed in the 3-6 year old participants. In addition, it was expected that there would also be a relationship between a higher quotient score and the participants' working memory score, with those achieving a higher score on the forward digit span task being more likely to copy a higher number of irrelevant actions. Finally, it was expected that participants who achieved a higher quotient score would do so because they understood the irrelevant actions to be necessary for task completion.

No goal (x3)

Participants viewed the no goal demonstration three times, and were then given the puzzle to complete (minus the plastic egg). Since previous research suggests that when an obvious goal state is removed, the actions themselves become the goal, it was expected that all participants in this condition in would overimitate more than participants in the goal focused (x3) condition. However, since participants only had the opportunity to view the demonstration three times before completing the task, it was expected that the youngest participants would display reduced copying fidelity in comparison to older participants due to constraints on working memory. It was also expected that very few participants would be able to identify a goal relating to this puzzle.

Goal-focused (looped)

Participants viewed the goal-focused demonstration three times, and were then presented with the puzzle to complete. However, the demonstration video continued to play in the background on a loop, although participants' attention was not drawn to the fact that it was still playing at any time. It was expected that a similar pattern of behaviour would be observed as in participants in the goal x3 condition (high overimitation in children, reduced overimitation in adults). However, there was also potential for increased overimitation in the youngest participants in comparison to both the 3x goal condition and the 3x no goal condition, as participants would have the demonstration video to refer back to, which would reduce working memory load.

No goal (looped)

Participants viewed the no goal demonstration three times, and were then presented with the puzzle to complete, but with no plastic egg inside. As with the previous looped condition, the demonstration video continued to play in the background on a loop, and participants' attention was not drawn to the fact that it was still playing at any time. It was expected that all participants in this condition would overimitate the most in comparison to all other conditions, as no obvious goal was present, suggesting that the imitation of the actions themselves should become the goal. Additionally, participants were not required to remember any of the actions, as they could continually refer back to the demonstration video as them completed the task, meaning that working memory load was reduced.

4.2.3. Design

This was a 2x2 design, intended to explore the influence of the task and its intended goal on participants' copying behaviours. Participants in each age group were randomly assigned to one of the four conditions. Table 4.2 indicates the behaviours that were expected to be observed in each condition.

Table 4.2: Expected findings for each of the four assigned conditions

Goal-focused demonstration (x3)	No Goal demonstration video (x3)
This condition was designed to replicate the results obtained in experiment two.	Removing the ball, and therefore the goal state, should encourage participants to copy the sequence of actions. Younger participants will copy fewer actions than older participants due to memory constraints.
Goal-focused demonstration (looped)	**No Goal demonstration video (looped)**
This condition is designed to remove the working memory load involved in remembering the full action sequence, particularly for younger children, meaning that if they wish to overimitate, they should be able to do so more easily.	It is expected that the highest levels of overimitation will be observed in this condition, as participants may view completing the sequence of actions as the goal, and the working memory load is reduced as they may refer back to the video whilst completing the task if they wish.

4.2.4. Procedure

Participants were all tested in the child development lab at Newcastle University. Parents remained in the room with their children at all times, although they were encouraged not to help their children with any of the tasks. Prior to completing any of the tasks, verbal consent was sought from children, as well as written consent from parents or from adult participants themselves, and participants were given the opportunity to ask questions at any time during the testing process.

All testing sessions were video recorded. Participants were seated in front of a computer monitor. The theory of mind task was administered first, followed by the forward digit span task. Participants were then asked to look at the computer monitor in front of them, and were informed that they would see a video of someone completing a puzzle. Participants in all four conditions viewed the demonstration video three times before being presented with the puzzle to complete, and being told "when you're ready, it's your turn". If participants asked how to complete the puzzle, or whether they should copy the demonstrator, they were told "it's up to you". No further instructions were given. The number of relevant and irrelevant actions completed were recorded, as well as the order in which actions were performed. As the actions used for the box task in experiment three were identical to those used in experiment two, inter-rater reliability tests were not repeated for this study. As previously stated, a Cohen's Kappa test for inter-rater reliability with regard to scoring the actions in the box task indicated good agreement between raters: $K = .696$, $p<.001$. Once participants had completed the puzzle once, younger participants were asked to explain how to complete the puzzle to a teddy bear who had been "sleeping" on the other side of the room and had missed the demonstration. Older children and adults were asked to imagine that they were explaining the puzzle to a friend who had missed the previous demonstration. Participants were asked to show their friend (or the teddy) how to perform the puzzle, using the box itself, as well as talking their friend through the actions being performed. Again, the number of relevant and irrelevant actions, as well as order, were recorded, in order to determine whether there were any differences in the actions copied. Information about participants' understanding of the puzzle could be identified during participants' explanation to the teddy or their friend (e.g. "you turn the box around to see what's inside", or "on the video she pulled this [the lever] but you don't have to do that, it doesn't do anything"). Participants were then asked a number of follow up questions:

"What is the box for?"/"What was the lady on the video trying to do?/What was your job when you were doing this puzzle?"

These questions were designed to assess participants' understanding of the goal of the task.

"Did you copy everything the lady did?"
To assess whether participants had consciously omitted actions, or whether they had simply forgotten them.

"Did you need to do all of those things to get the puzzle to work?"
To determine whether the participants interpreted the irrelevant actions as necessary for task completion.

Participants' answers to these questions were video recorded and coded as described below.

4.2.5. Data Analysis

Behavioural scoring

Participants' quotient scores for experiment 3 were calculated in the same way as in experiment two. Participant data was scored depending on the number of irrelevant actions produced and the order in which the relevant actions were copied. Within this task, there were three possible irrelevant actions (knock, turn, lever) and two potential measures of the order of actions completed: the order in which participants completed the relevant actions (left lever, right lever) only, and the order in which they completed all actions. Participants were awarded a score of 1 for each of these elements that they completed, meaning that the maximum possible score for this task was 6. A quotient score was then calculated for each participant, by taking their actual score on the task and dividing it by the maximum possible score. A quotient score closer to 1 would indicate higher fidelity copying (overimitation), whereas a quotient score closer to 0 would indicate lower fidelity copying (emulation). As well as quotient score, participants' performance of individual aspects of the task, such as the irrelevant actions used, was also examined, in order to gain a better understanding of more nuanced behaviour during task completion.

Participants completed the box task twice, meaning that they each achieved two quotient scores: an imitation score, based on their task performance immediately following the demonstration video, and a demonstration score, based on which actions they performed when they were demonstrating the task to either the teddy bear or their “friend”. This was designed to determine whether participants’ performance on the task changed on their second attempt. Allowing participants to become the demonstrator also meant that they could verbalise what they were doing as they completed the actions, providing a greater insight into their understanding of the task. Based on participants’ answers to the debriefing questions, their behaviours were coded as in Table 4.3.3.

Table 4.3: Examples of coded behaviours based on participants’ answers to debriefing questions

Question	Score
Did participants identify a goal relating to the puzzle?	Participants were awarded a score of 1 if they identified any goal (e.g. Participant 21 (aged 9): “Trying to get the ball out of the box.”; Participant 91 (aged 8): “Probably do the same thing.”), and a score of 0 if they were unable to do so (e.g. “I don’t know”).
What goal was identified?	Participants were grouped based on whether they identified the goal as retrieving the ball, copying the demonstrator, or whether they were unable to identify a goal at all.
Did participants intentionally attempt to copy the irrelevant actions?	Participants were awarded a score of 1 if they stated that they had tried to copy exactly what the demonstrator had done (whether they had unintentionally omitted actions or not), and a score of 0 if they indicated that they purposefully chose to omit irrelevant actions. This score could also be based on participants’ responses when demonstrating the task either to the teddy bear or their imaginary friend (e.g. “on the video she knocked but I don’t need to do that”).
Did participants need to complete these actions to get the puzzle to work correctly?	Participants were given a score of 1 if they stated that the irrelevant actions were required for the puzzle to work correctly, and a score of 0 if they indicated that the puzzle would still work when actions were omitted. A score of 1 was still given if participants indicated that any of the irrelevant actions performed were required (e.g. if they only performed the lever action, but indicated that it was functional).

Working memory

Participants received a score on the forward digit span task based on the length of the longest sequence of numbers that they were able to correctly repeat.

Theory of Mind

As in experiment two, participants received a score of 1 if they passed the false belief task, and 0 if they failed. In order to allow for comparison with the results from experiment two, similar participant age groups (4-5, 6-8, 9-12, adults) were used for analysis of the theory of mind data collected in this experiment.

4.3. Results

4.3.1. Theory of Mind

Performance on the theory of mind task was investigated for each of the previously identified age groups, and is shown in figure 4.5. Using a chi squared test, a significant difference between age groups was found: $x2(3)=27.08$, $p<.001$. As in experiment two, participants appeared to be able to pass this test consistently by the age of nine. 30% of four to five year old participants (10 out of 23 four year olds, and 3 out of 21 five year olds) failed this task, in comparison to only 6% (two participants, both aged six) of 6-8 year old participants. All participants aged seven and above passed the test, which is consistent with findings from experiment two, and from previous research, which suggests that the ability to pass this type of false belief task should develop around the age of seven. This suggests that this task worked as intended.

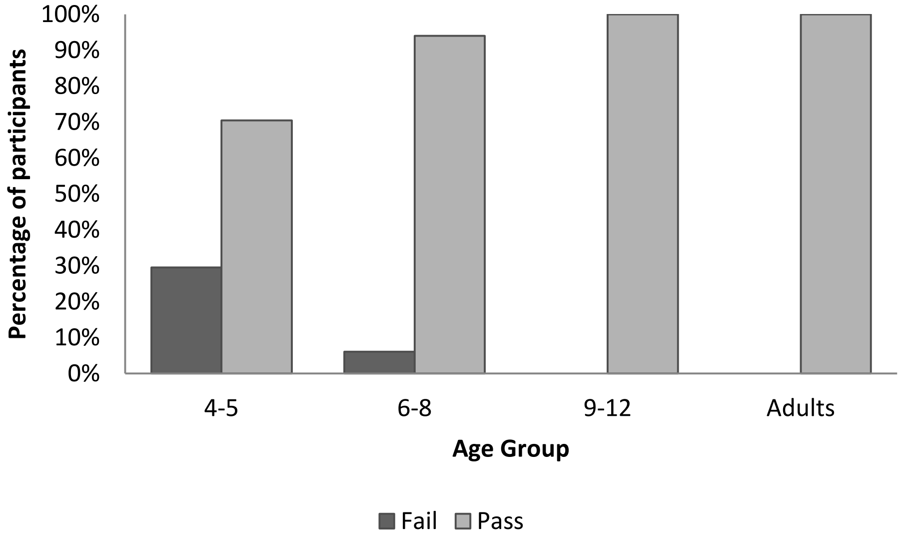


Figure 4.5: Percentage of participants in each age group that passed or failed the theory of mind task

However, contrary to expectations, performance on the theory of mind task did not appear to influence participants' copying strategies. There was no significant difference in participants' quotient score of actions imitated between those who passed the theory of mind task (N=133) and those who failed it (N=15): F(1,148)=.147, p=.702. Similarly, there was no significant difference in participants' quotient score of actions demonstrated between those who passed or failed the theory of mind task: F(1,48)=.455, p=.501. The mean imitation and demonstration quotient scores for participants who passed or failed the theory of mind task are indicated in table 4.4.

Table 0: Mean imitation and demonstration quotient scores for participants who passed and failed the false belief task

Theory of Mind Score	N	Mean Imitation Quotient Score	Std. Deviation	Mean Demonstration Quotient Score	Std. Deviation
Fail	15	0.61	0.316	0.57	0.345
Pass	134	0.58	0.313	0.51	0.328

In order to determine whether participants' theory of mind performance influenced their tendency to copy the demonstrator, a binary logistic regression was conducted, using participant age group and theory of mind score as predictors, and whether participants stated that they tried to copy or not as the dependent variable. The model was found to be significant overall: $x^2(2)=42.70$, p<.001, and accounted for 34.6% of the variance. Participant age was found to be a significant predictor: $x^2(1)=.26.67$, p<.001. However, theory of mind score was not found to be a significant predictor of participants' understanding of the functionality of the irrelevant actions: $x^2(1)=.130$, p=.719. Since age was a significant predictor, it is important to consider that this difference might be due to participants' understanding of the question. As indicated in figure 4.6, 93% of the participants who failed the theory of mind task stated that they had tried to copy the demonstrator precisely. In comparison, only 62% of participants who passed the theory of mind task stated that they had attempted to copy the demonstrator. However, the number of participants that failed the theory of mind task (N=15) was low as a percentage of the total number of participants (N=148), making up around only 10% of the total, and so perhaps there were simply too few participants who failed the task to allow for a fair comparison. These theory of mind results are plausible, however, given the age range of participants within this study.

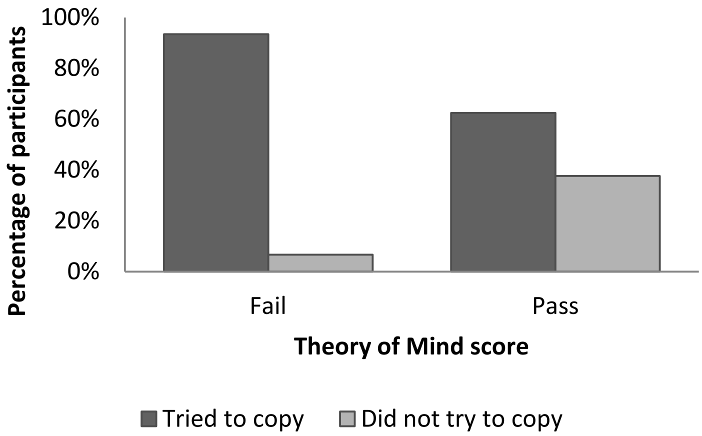


Figure 4.6: Percentage of participants that passed or failed the theory of mind task and attempted to copy the actions of the demonstrator

4.3.2. Working Memory

Age comparisons

The mean digit span score for each participant age group is shown in table 4.5.

Table 4.5: Mean digit span score for participants in each assigned age group

Age Group	N	Mean	Std. Deviation
4-5	44	4.20	1.133
6-8	33	5.24	1.091
9-12	30	5.90	1.155
Adults	42	6.88	0.942
Total	149	5.53	1.491

As indicated in table 4.5, participants’ performance on the task rose with age as expected. An ANOVA indicated a significant difference in digit span score between age groups: F(3,148)=46.17, p<.001. Bonferroni post-hoc tests indicated a significant difference in digit span score between 4-5 year olds (M=4.20) and 6-8 year olds (M=5.24, p<.001), 9-12 year olds (M=5.90, p<.001) and adults (M=6.88, p<.001). There was also a significant difference between 6-8 year olds and adults (p<.001) and 9-12 year olds and adults (p=.001).

Number of irrelevant actions copied

In order to determine whether there was any relationship between working memory and participants' task performance, an ANCOVA was used, with age group as a covariate. Table 4.6 indicates the mean digit span for participants achieving each of the possible imitation quotient scores.

Table 4.6: Mean digit span scores for participants (all age groups) achieving each of the possible imitation quotient scores (.00, .20, .40, .60, .80, 1.00)

Imitation Quotient	N	Mean	Std. Deviation
.00	3	7.33	0.577
.20	32	6.00	1.244
.40	33	5.15	1.822
.60	27	5.07	1.357
.80	14	5.00	1.414
1.00	40	5.83	1.299
Total	149	5.53	1.491

Perhaps surprisingly, participants that achieved an imitation quotient score of 0 also achieved the highest mean digit span score of 7.33. This suggests that participants' ability to remember the sequence was not the main influence on their subsequent behaviour when completing the task. No significant difference in mean digit span score was observed between participants belonging to the different quotient score groups: F(5,148)=1.31, p=.264. As would be expected, given the results in table 4.6, there was a significant relationship between the covariate, age, and mean digit span score: F(1,148)=116.63, p<.001. In order to further investigate the differences between groups, the range of responses, and the minimum and maximum digit span scores for participants achieving each of the imitation quotient scores was examined (table 4.7).

Table 4.7: Minimum and maximum responses on the digit span task for participants achieving each of the possible imitation quotient scores (.00, .20, .40, .60, .80, 1.00)

Imitation Quotient	N	Minimum	Maximum	Range
.00	3	7	8	1
.20	32	4	8	4
.40	33	2	9	7
.60	27	3	9	6
.80	14	2	7	5
1.00	40	3	8	5

Interestingly, participants with high quotient scores had a high range of responses, with some participants achieving a score of .80 and yet only achieving a digit span score of 2, suggesting that they could only remember a sequence of two numbers.

The mean digit span score for participants in each of the demonstration quotient score groups was also examined, again using age group as a covariate. Results are shown in table 4.8.

Table 4.8: Mean digit span score for participants (all age groups) who achieved each of the possible demonstration quotient scores (.00, .20, .40, .80, 1.00)

Demonstration Quotient	N	Mean	Std. Deviation
.00	8	5.88	1.458
.20	42	6.19	1.401
.40	32	5.09	1.634
.60	21	4.71	1.102
.80	10	5.30	1.160
1.00	35	5.71	1.341
Total	148	5.55	1.467

As before, there was no significant difference observed between groups: $F(5,147)=1.35$, $p=.247$, suggesting that participants' task performance was not influenced by their working memory. Again, participants' range of responses, as well as their minimum and maximum responses were investigated (table 4.9).

Table 4.9: Minimum and maximum responses on the digit span task for participants achieving each of the possible demonstration quotient scores (.00, .20, .40, .60, .80, 1.00)

Demonstration Quotient	N	Minimum	Maximum	Range
.00	8	4	8	4
.20	42	2	9	7
.40	32	3	9	6
.60	21	3	7	4
.80	10	4	7	3
1.00	35	3	8	5

As before, some participants who achieved the maximum demonstration quotient score only achieved a digit span score of 3, indicating that working memory was not the only influence on task performance. However, this may also indicate that either the working memory test used was not appropriate to allow for a comparison with participants' task performance.

Finally, the mean working memory score for participants in each of the four conditions was examined. Since the working memory task was conducted at the start of the trial, before participants received any instructions or information regarding the task, the condition to which they were assigned should not have had any effect on their working memory score. The mean working memory score for each condition is shown in table 4.10.

Table 4.10: Mean digit span score for participants (all ages) in each of the assigned conditions

Assigned Condition	N	Mean	Std. Deviation
Goal (3x)	39	5.38	1.227
No Goal (3x)	33	5.45	1.872

Goal (looped)	41	5.63	1.356
No Goal (looped)	36	5.64	1.552
Total	149	5.53	1.491

Levene's test indicated a significant result (p=.013), and therefore a Kruskal-Wallis test was used for analysis instead of an ANOVA. As expected, given the mean scores in table 4.10, there was no significant difference in mean working memory score between the assigned conditions: $x^2(3)=.971$, p=.808.

Order of relevant actions copied

It was anticipated that participants' digit span performance would be related to the order in which they copied actions from the demonstrator, since the digit span task tests memory for sequences. An ANCOVA was used to compare the mean digit span score for participants who imitated all of the actions from the demonstration video, in comparison to those who did not, using age as a covariate. A significant relationship between the order of actions performed and mean digit span score was found: F(1,148)=6.29, p=.013, with participants who completed the actions in the same order as the demonstrator achieving a higher digit span score (N=103, M=5.80, SD=1.26) than those who did not (N=46, M=5.41, SD=1.58). A significant relationship between age and mean digit span score was observed as expected: F(1,148)=145.77, p<.001.

4.3.3. Comparisons between results obtained in experiments two and three

As the goal (3x) condition was intended to replicate the findings obtained in experiment two, comparisons were made between the percentage of participants in this condition achieving each of the possible imitation quotient scores, and the percentage of participants achieving each of the possible imitation quotient scores in experiment two. Figure 4.7 indicates the differences in task performance across the two experiments. Although task performance was fairly consistent for the adult age group, there was much more variation between the younger age groups. In particular, 6-8 year old participants in experiment three were found to achieve higher mean quotient scores than those in experiment two, with all 6-8 year olds in experiment three achieving a score of at least .60. Participants in both experiments were provided with the same puzzle and instructions, and the demonstration video was identical, aside from slightly longer gaps in between actions in the demonstration video used in experiment three. It may be the case that the additional time between actions allowed participants more time to rehearse them, meaning that they were better able to remember them and imitate them by the time they came to complete the puzzle themselves.

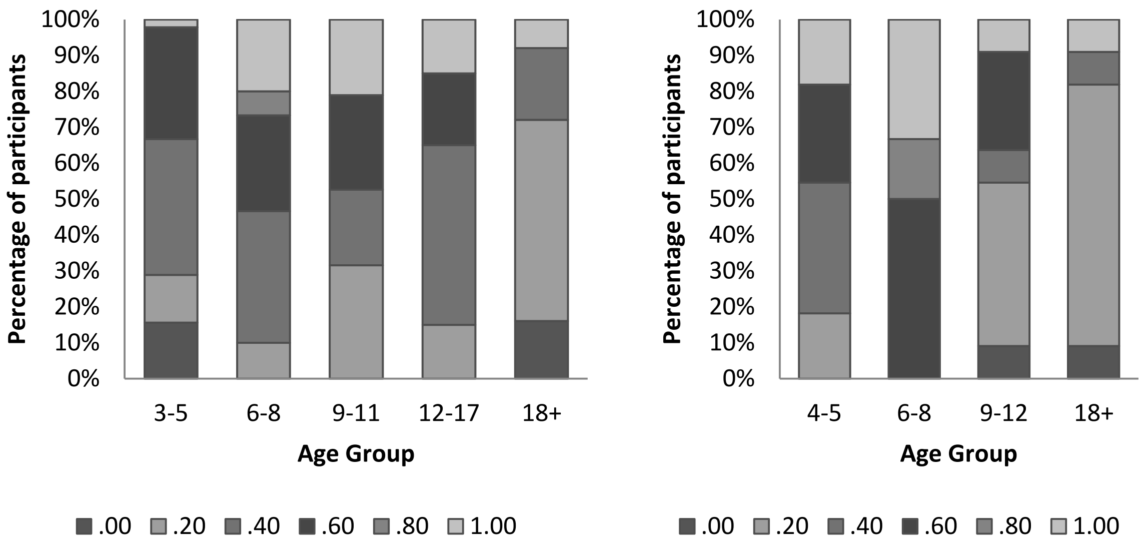


Figure 4.7: Comparisons between box task performance for participants in experiment two (L) and experiment three (R).

Mean quotient scores for each task across the four age groups used in both experiments are indicated in table 4.11. Interestingly, only adult scores remained consistent across both experiments: adults achieved identical mean scores in the box task in experiment 2, and in experiment 3 (3x goal condition only). The mean quotient score increased for the youngest two age groups, potentially due to the increased time between actions in the demonstration video in experiment 3, which would have provided additional time for participants to process the actions, and may have also allowed participants more time to mentally rehearse the actions. Additionally, the mean quotient score decreased for the 9-11 age group. This may again be due to the increase in time between actions performed, which may have allowed participants to more easily interpret the irrelevant actions as non-functional and unnecessary for task completion.

Table 4.11: Mean quotient scores for participants in each age group in experiments 2 and 3

	Experiment 2		Experiment 3	
Age Group	**Mean**	**Std. Deviation**	**Mean**	**Std. Deviation**
4-5	0.41	0.248	0.53	0.272
6-8	0.58	0.259	0.77	0.197
9-11	0.52	0.300	0.38	0.289
Adults	0.27	0.251	0.27	0.257

4.3.4. Imitation quotient score

The imitation quotient score refers to the quotient score obtained by participants during their first attempt at the task, immediately following the demonstration video.

A 2x2 ANOVA was used to investigate the relationship between participant age groups, the condition to which they had been assigned, and their mean imitation quotient score. There was a significant main effect of age: F(3,148)=6.72, p<.001, with Bonferroni post-hoc tests indicating that 4-5 year olds achieved a significantly lower imitation quotient score overall (M=.58, SD=.27) in comparison to 6-8 year olds (M=.77, SD=.24, p=.013). There was also a significant difference between 6-8 year olds (M=.77, SD=.24) and adults (M=.48, SD=.35). These results indicate that 6-8 year olds were the most likely of all the age groups to display high fidelity copying.

There was also a significant main effect of assigned condition: F(3,148)=5.44, p=.001. Bonferroni post-hoc tests indicated that participants in the looped (no goal) condition achieved a significantly higher imitation quotient score (M= .76, SE=.05) than those in the goal (3x) condition (M=.49, SE=.04, p=.001). No other significant differences between groups were observed.

Figure 4.8 displays the mean imitation quotient score for each assigned condition and age group. A significant interaction was found between age group and copying group: F(9,148)=2.92, p=.003. Whilst the mean imitation quotient score remained fairly stable across all copying types in the 4-5 and 6-8 age groups, 9-12 year olds and adults in the looped (no goal) condition had a higher imitation score than participants in the remaining three conditions. It may have been the case that, in the absence of instructions, older participants interpreted the repetition of the video as a signal that they should copy it precisely. As in experiment two, these results indicate that adult participants may be more sensitive to the features of a task and its associated goals.

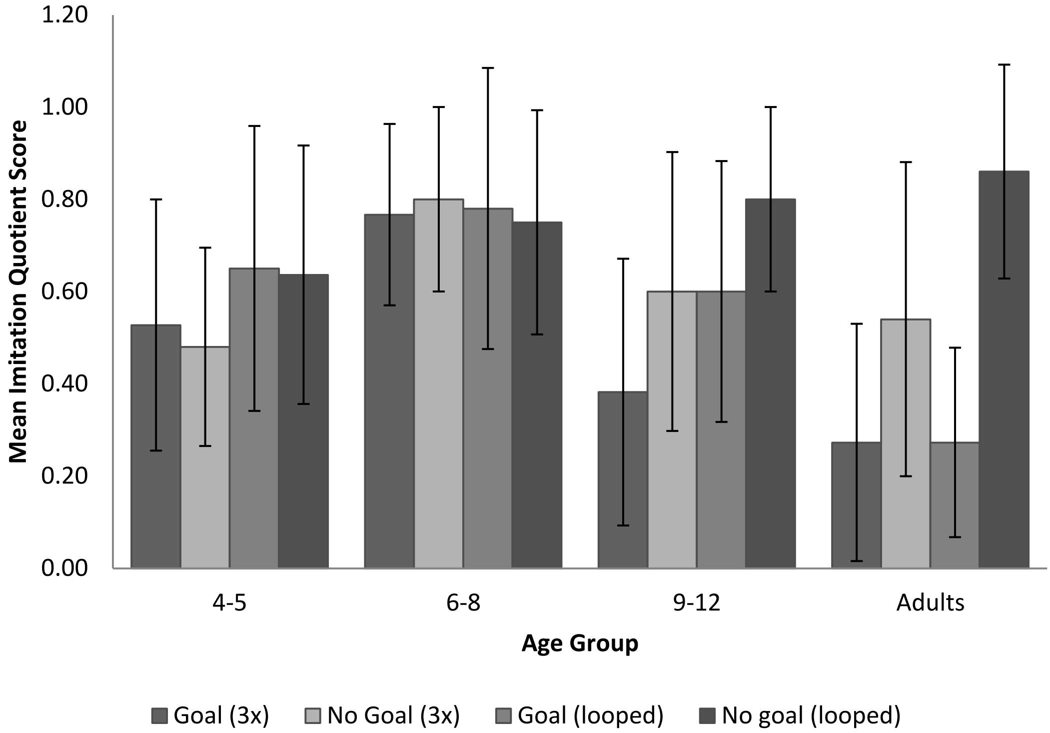


Figure 4.8: Mean imitation score for participants in each of the four assigned conditions (goal x3, no goal x3, looped goal, looped no goal) split across each of the age groups. Error bars represent standard deviation.

4.3.5. Demonstration quotient score

The demonstration quotient score refers to the quotient score obtained by participants on their second attempt at the task, when they were asked to demonstrate how to complete the task to another individual (either a teddy bear, or an imaginary peer, depending on their age). As previous research indicates that participants often change their strategy following multiple attempts at a task, it was expected that there would be a difference between participants' imitation and demonstration quotient scores. This may be because they may have learned more about the function of each of the actions during their first task attempt, and therefore may have chosen to omit these during their second attempt. Alternatively, although participants may have chosen to copy certain actions themselves, they may not have felt that it was necessary to pass on this information to others in order to allow them to complete the task for themselves. Additionally, for younger participants, memory was likely to be an issue here: they may simply have forgotten some of the actions between their first and second attempts.

A two way ANOVA was used to investigate the relationship between participant age groups, the condition to which they were assigned, and their mean demonstration quotient score. There was a significant main effect of age: $F(3,147)=8.02$, $p<.001$, with 6-8 year olds having the highest demonstration quotient score ($M=.70$, $SD=.30$). 6-8 year olds were found to have a significantly higher score than 9-12 year olds ($M=.43$, $SD=.28$, $p=.014$) and adults ($M=.40$, $SD=.35$, $p<.001$).

There was also a significant main effect of assigned condition: $F(3,147)=4.54$, $p=.005$. Participants in the looped (no goal) condition ($M=.69$, $SE=.05$) were found to have a significantly higher score than those in the goal (3x) condition ($M=.44$, $SE=.05$, $p=.003$), as well as those in the looped (goal) condition ($M=.50$, $SE=.04$, $p=.040$).

Finally, there was a significant interaction between age and condition: $F(9,147)=3.99$, $p<.001$. Figure 4.9 displays the mean quotient score for each copying group and age group. As with the imitation quotient score, 9-12 year olds and adults in the looped (no goal) group have a much higher score than in all other conditions for those age groups. As some participants reported interpreting the goal of the task as a “memory test” (see section 4.3.7), it may be that this is reflected specifically in the results for this condition.

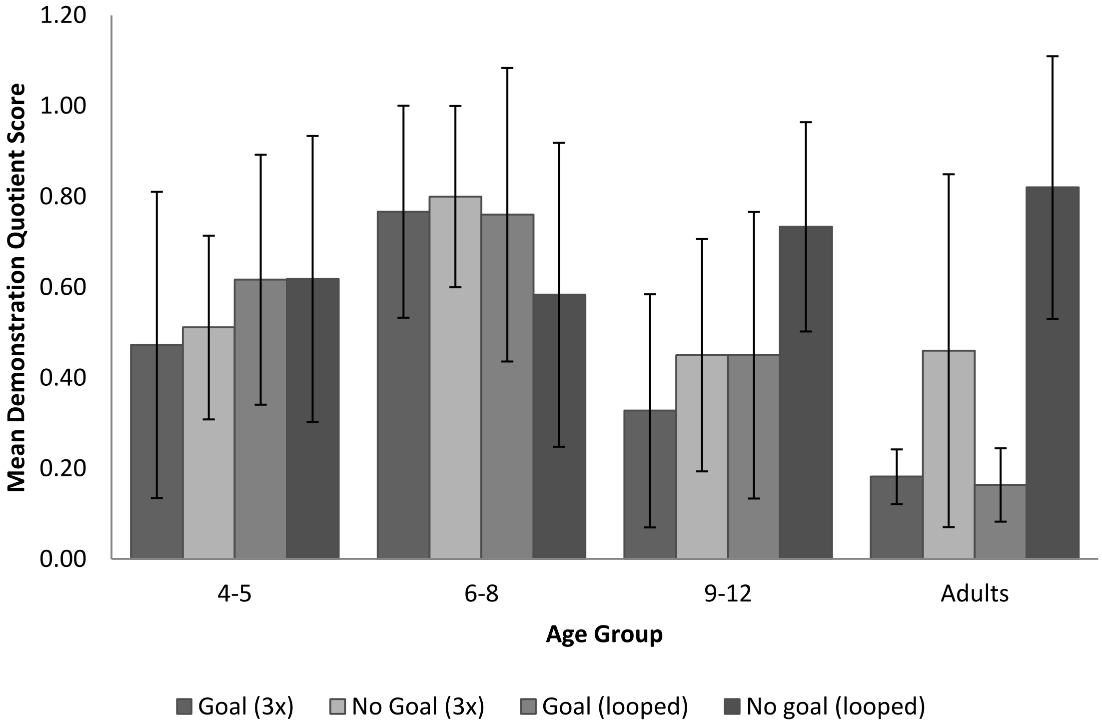


Figure 4.9: Mean demonstration quotient score for participants in each of the four assigned conditions (goal x3, no goal x3, looped goal, looped no goal) split across each of the age groups. Error bars represent standard deviation.

4.3.6. Differences in actions copied between attempts

In order to determine whether participants made an attempt to modify the copying strategy used between attempts at the task, the difference between participants' imitation quotient scores and demonstration quotient scores were calculated, by subtracting the demonstration quotient score from the imitation quotient score. The mean differences between attempts for participants in each age group are shown in table 4.12.

Table 4.12: Mean difference between imitation and demonstration quotient scores for participants in each age group

Age Group	N	Mean	Std. Deviation
4-5	43	0.023	0.165
6-8	33	0.067	0.238
9-12	30	0.107	0.250
Adults	42	0.081	0.207
Total	148	0.066	0.213

An ANCOVA, using age group as a covariate, was used to determine whether there were any differences in score between the two attempts. No significant effect of age was found: $F(1,148)=2.09$, $p=.150$.

There was also no significant effect of group: $F(3,148)=.148$, $p=.931$. The mean difference between attempts is for each group is shown in table 4.13.

Table 4.13 Mean difference between imitation and demonstration quotient scores for participants in each assigned condition

Group	N	Mean	Std. Deviation
Goal (3x)	39	0.056	0.177
No Goal (3x)	32	0.056	0.170
Goal (looped)	41	0.073	0.194
No Goal (looped)	36	0.078	0.296
Total	148	0.066	0.213

These results suggest that participants' strategies between their first attempt at imitating the actions from the demonstration video and then subsequently demonstrating them did not change significantly, and instead, their behaviour remained consistent.

To further investigate the relationship between participants' attempts at the task, a Spearman's rho correlation was run for each age group. The results of these correlations are in table 4.14.

Table 4.14: Correlation results comparing differences in quotient scores between attempts at the task across age groups. ** indicates significance at the .01 level.

Age Group	N	r	p
3-5	44	0.790	<.001**
6-8	33	0.658	<.001**
9-12	30	0.677	<.001**
Adults	42	0.790	<.001**

There were strong, significant positive correlations between imitation and demonstration quotient scores for participants in all age groups, indicating that performance stayed consistent between attempts at the task; if a participant achieved a high imitation score, then they were subsequently likely to achieve a high demonstration score.

The differences in score between participants in each assigned conditions were compared (table 4.15).

Table 4.15: Correlation results comparing differences in quotient scores between attempts at the task across assigned conditions. ** indicates significance at the .01 level.

Condition	N	r	p
Goal (3x)	39	0.808	<.001**
No goal (3x)	32	0.719	<.001**
Goal (looped)	41	0.828	<.001**
No goal (looped)	36	0.509	.002**

Again, there were strong, significant, positive correlations between imitation and demonstration quotient scores for participants in all assigned conditions. These findings are contrary to previous research, which suggested that participants may refine their strategy between attempts at the task.

Figure 4.10 indicates the percentage of participants that performed each of the irrelevant actions (knock, lever, turn) in the imitation and demonstration attempts.

Figure 4.10 shows that although participants in each of the age groups were less likely to imitate actions in the demonstration condition in comparison to the imitation condition, the difference in the percentage of participants that imitated the actions was low.

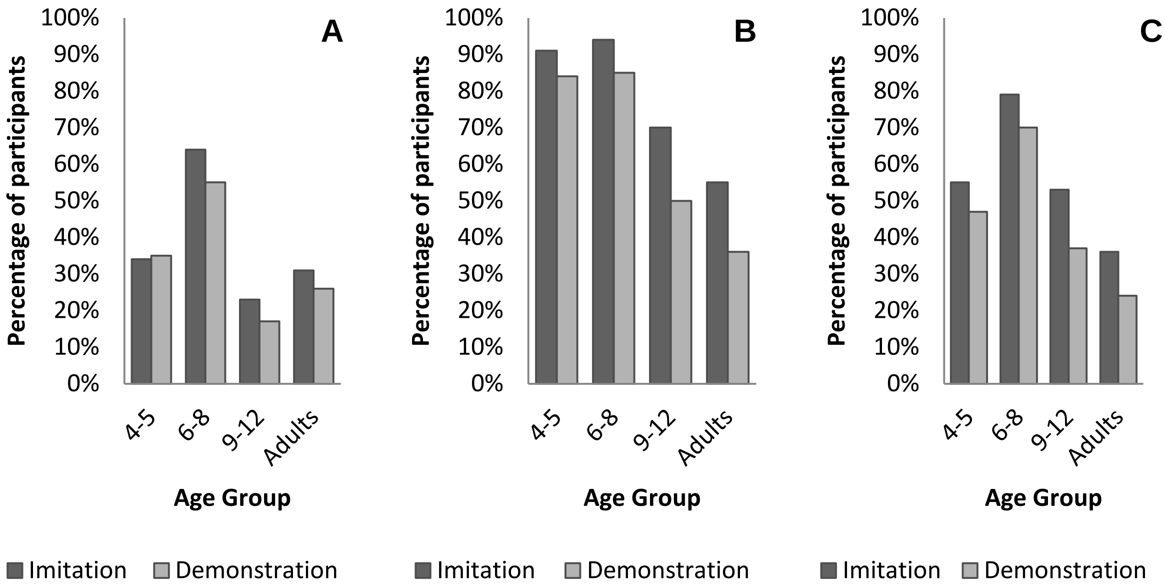


Figure 4.10: Percentage of participants in each age group that copied each of the irrelevant actions (A: Knock, B: Lever, C: Turn) in both the imitation and demonstration attempts.

Figure 4.11 indicates the percentage of participants that performed each of the irrelevant actions (knock, lever, turn) in the imitation and demonstration attempts. Again, participants in each of the conditions were less likely to copy each of the irrelevant actions in the demonstration attempt in comparison to the imitation attempt.

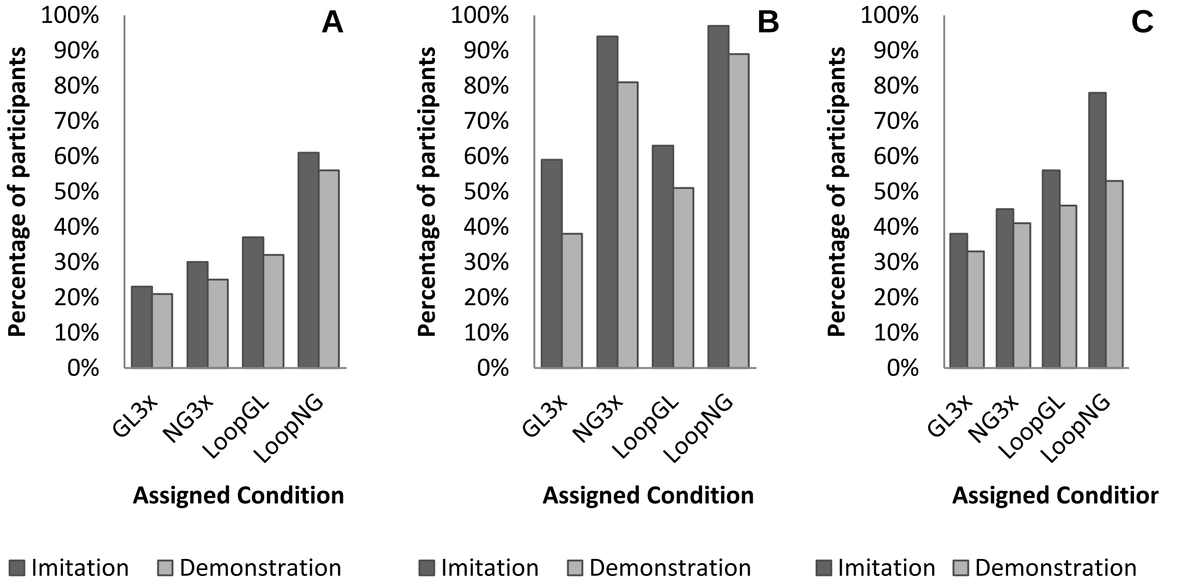


Figure 4.11: Percentage of participants in each assigned condition that copied each of the irrelevant actions (A: Knock, B: Lever, C: Turn) in both the imitation and demonstration attempts.

4.3.7. Understanding of task and function of actions

Identifying a goal relating to the task

Comparisons were made between age groups to investigate whether participants' understanding of the task, as well as the intended goal state, changed with age (figure 4.12). A significant difference was found in the number of participants across all conditions who were able to identify a specific goal relating to the task (either to release the ball, or to copy the demonstrator); x^2 (3)=18.81, p<.001. Whilst the vast majority of 9-12 year olds and adults were able to identify some sort of goal related to the task, almost half of 4-5 year olds and 38% of 6-8 year olds were not. However, it is important to consider that this may be due to the fact that these participants are still developing their language skills, and they may not have understood the questions being asked of them.

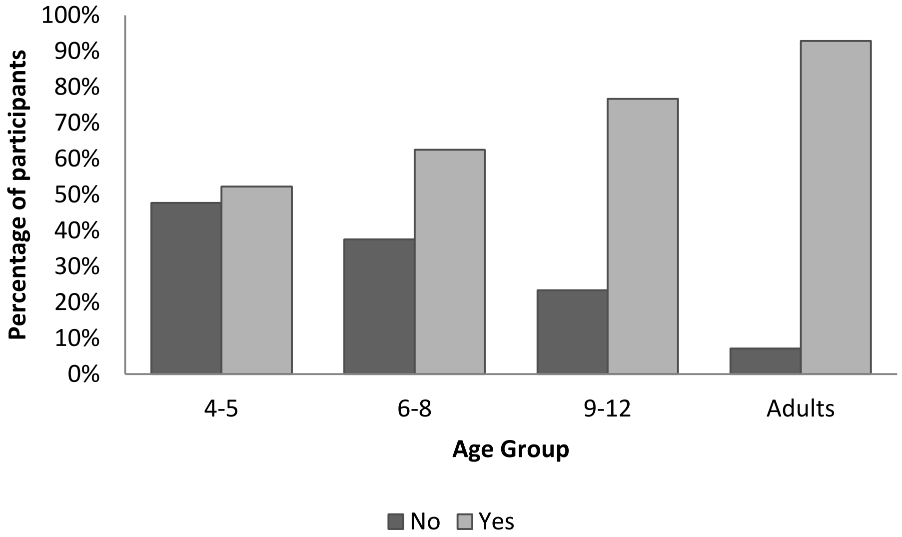


Figure 4.12: Percentage of participants in each age group able to identify a goal

Figure 4.13 indicates the percentage of participants in each condition that were able to identify a specific goal relating to the task (either to retrieve the ball, or to copy the demonstrator).

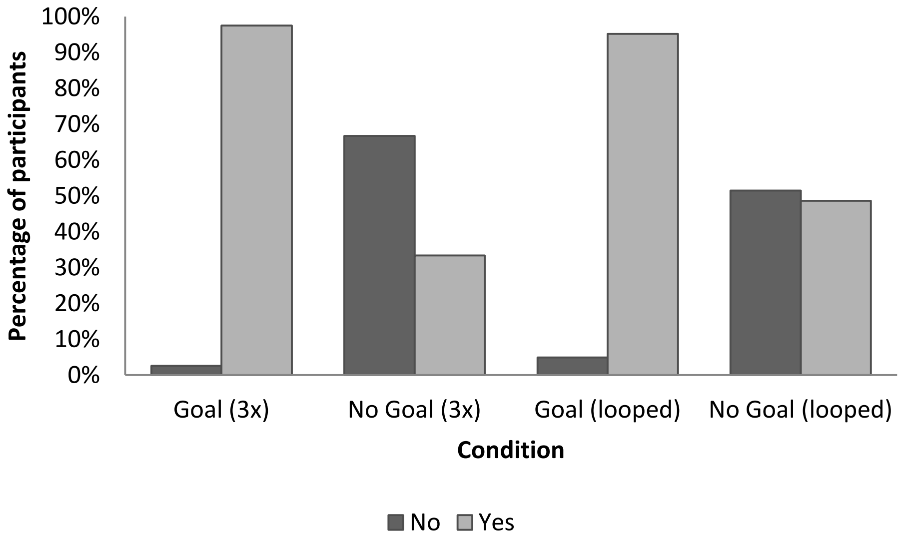


Figure 4.13: Percentage of participants in each assigned condition able to identify a goal

A significant difference between conditions was observed: $x^2(3)=56.05$, $p<.001$. The majority of all participants in the goal (3x) and goal (looped) conditions were able to identify a goal relating to the task. However, 67% of participants in the no goal (3x) were unable to identify a goal. Similarly, 51% of participants in the no goal (looped) condition were also unable to identify a goal.

Goal types identified by participants

All participants identified one of three broad goals: retrieving the ball from the box, copying the demonstrator, or no goal. The goals identified by participants in each condition are indicated in figure 4.14.

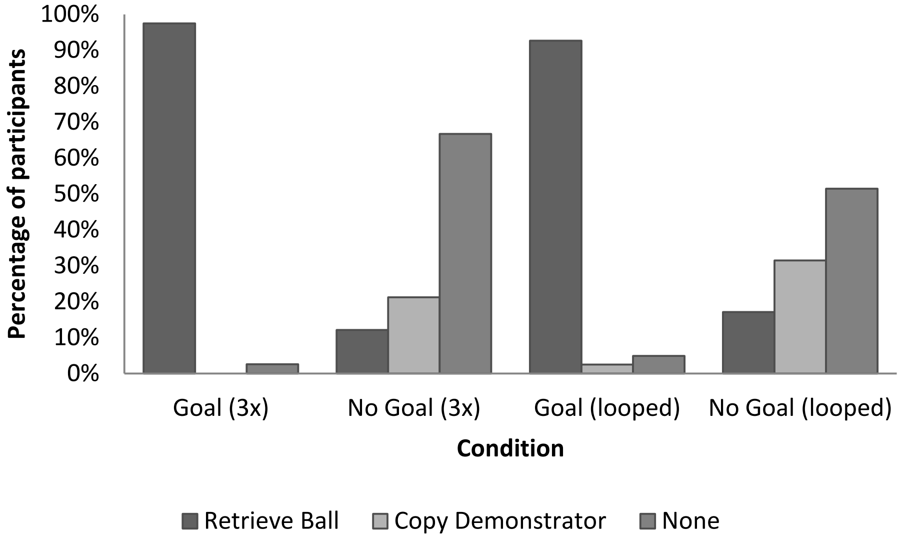


Figure 4.14: Goals identified by participants in each of the conditions (goal x3, no goal x3, looped goal, no goal looped)

A significant difference was observed between groups: $x^2(6)=100.39$, $p<.001$. The majority of all participants in the goal (3x and looped) conditions stated that the goal was to retrieve the ball. However, in the no goal conditions, where the aim of the task was less obvious, the majority of participants were unable to identify a goal. 12% of participants in the no goal (3x) condition and 17% of participants in the no goal (looped) condition identified the goal as retrieving the ball even though there was no ball visible in the demonstration video. These participants were identified as having this goal if they stated that they thought that they would probably put something (e.g. a ball or marble) inside the box and use the levers to move it around. These participants were coded as “retrieve ball”, even though a ball was not present, since that was the goal they had identified.

The goals identified by participants in each age group were also investigated. These are displayed in figure 4.15.

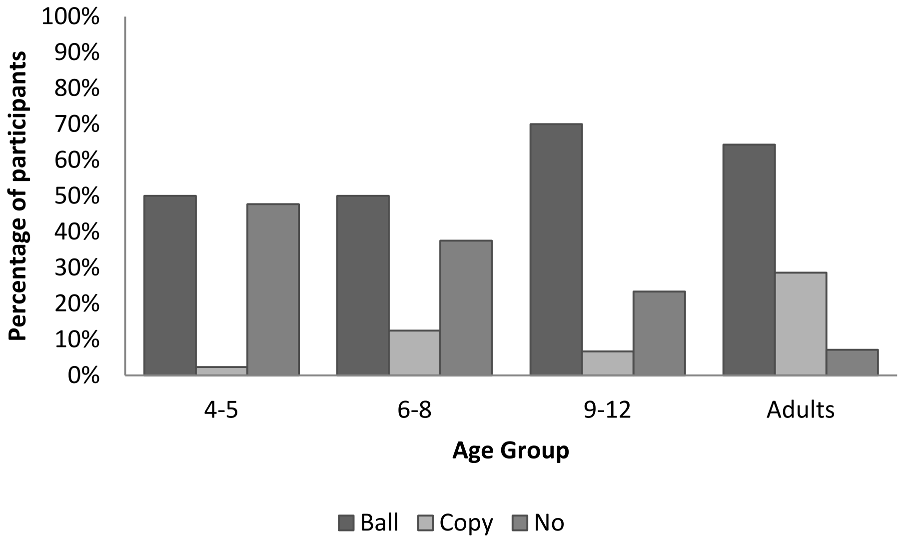


Figure 4.15: Goals identified by participants in each age group

A chi-squared test indicated a significant difference between groups: x^2 (6)=28.03, p<.001. The majority of participants in all age groups identified the goal as being to retrieve the ball from the box. However, almost half (48%) of 4-5 year old participants were unable to identify a goal, as were 38% of 6-8 year old participants. Adults were also the most likely age group to identify the goal as being to copy the demonstrator.

Deciding whether or not to copy irrelevant actions performed

Participants were asked if they had attempted to perform the puzzle in the same way as the demonstrator (“Did you do everything the lady in the video did?”). If participants indicated that they had attempted to do this, they were given a score of 1, whether they actually did copy all of the actions demonstrated or not, as the purpose of this question was to measure intent to copy rather than the ability to remember and subsequently perform each of the actions shown. Participants who answered “I don’t know”, or stated that they had not tried to copy the demonstrator, were given a score of 0.

The percentage of participants in each age group that chose to copy, or chose not to copy the demonstrator is shown in figure 4.16.

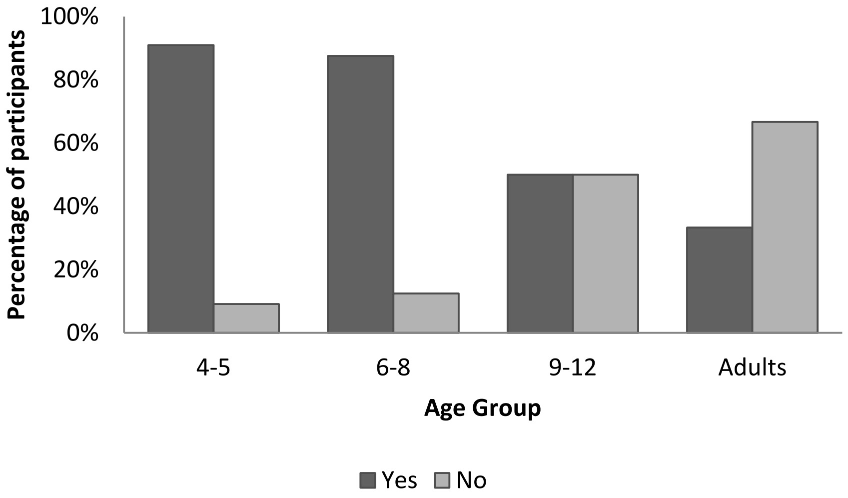


Figure 4.16: Percentage of participants in each age group that reported that they attempted to copy the actions performed by the demonstrator

A significant difference between age groups was found: x^2 (3)=41.87, p<.001. Participants’ intention to copy the demonstrator appears to decrease with age. Almost all participants in the 4-5 and 6-8 age groups stated that they had tried to copy what the demonstrator had done in the video. Half of participants in the 9-12 age group stated that they had tried to copy the demonstrator, in comparison to only 33% of adults.

Figure 4.17 shows the percentage of participants in each assigned condition that reported that they attempted to copy the actions performed by the demonstrator. A significant difference between assigned conditions was observed: x^2 (3)=13.50, p=.003. Removing the goal from the demonstration video appears to have increased participants’ intention to copy the demonstrator, as 67% of participants in the no goal (3x) condition and 89% of participants in the no goal (looped) condition reported that they intentionally copied the demonstrator. Additionally, looping the video also appears to have increased participants’ intention to copy, as 61% of participants in the goal (looped) condition reported that they intentionally copied the demonstrator.

This is further supported by the fact that participants in the no goal (looped) condition were the most likely to report that they intended to copy the demonstrator.

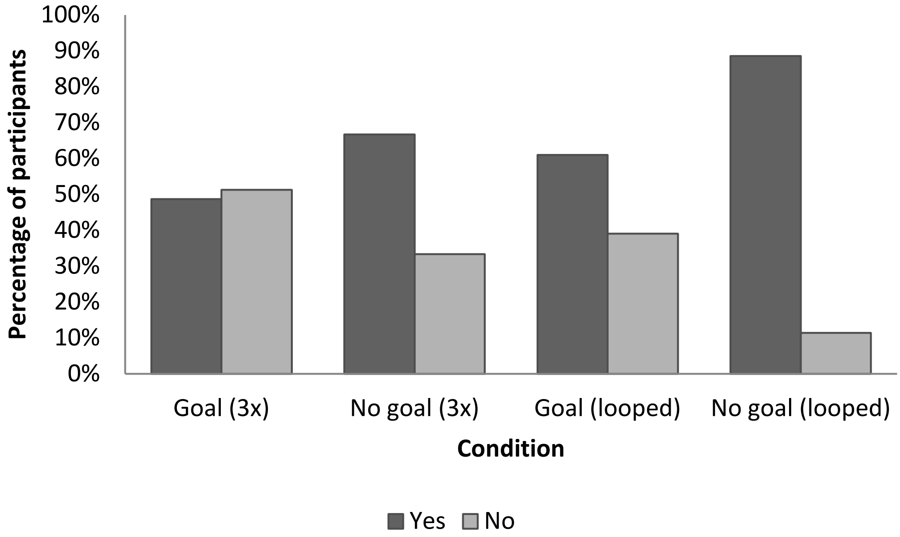


Figure 4.17: Percentage of participants in each age group that reported that they attempted to copy the actions performed by the demonstrator

Interpreting the functionality of actions performed

Participants were asked whether they believed that the actions performed (specifically referring to the irrelevant actions) were necessary in order to complete the task (figure 4.18).

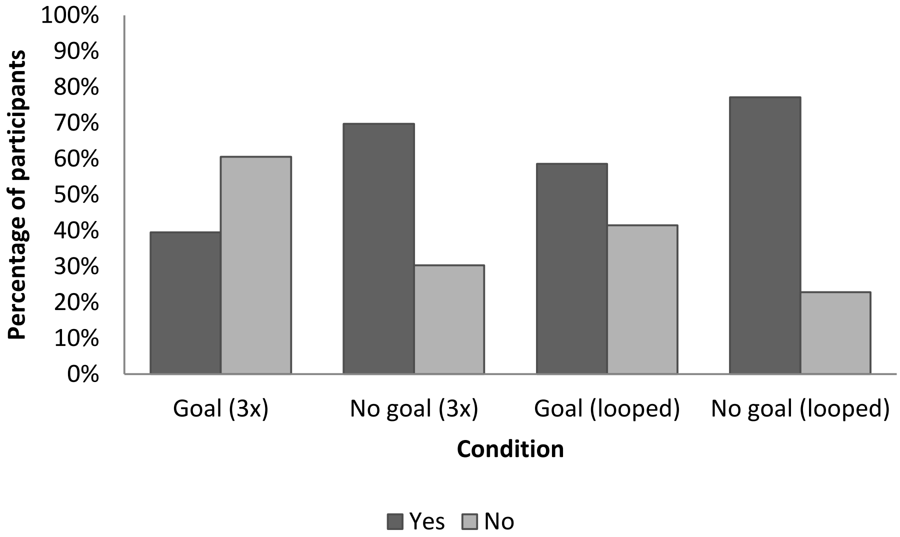


Figure 4.18: Percentage of participants in each assigned condition that believed that the irrelevant actions were functional

There was a significant difference between conditions in terms of the number of participants who interpreted the irrelevant actions as being functional, using a chi squared test: x^2 (3)=12.34, p=.006. The majority of participants in the goal (3x) condition did not interpret the irrelevant actions as being necessary for task completion. However, the majority of participants in all other conditions did interpret at least one of the irrelevant actions as being functional.

The percentage of participants in each of the age groups that interpreted the actions as functional is shown in figure 4.19.

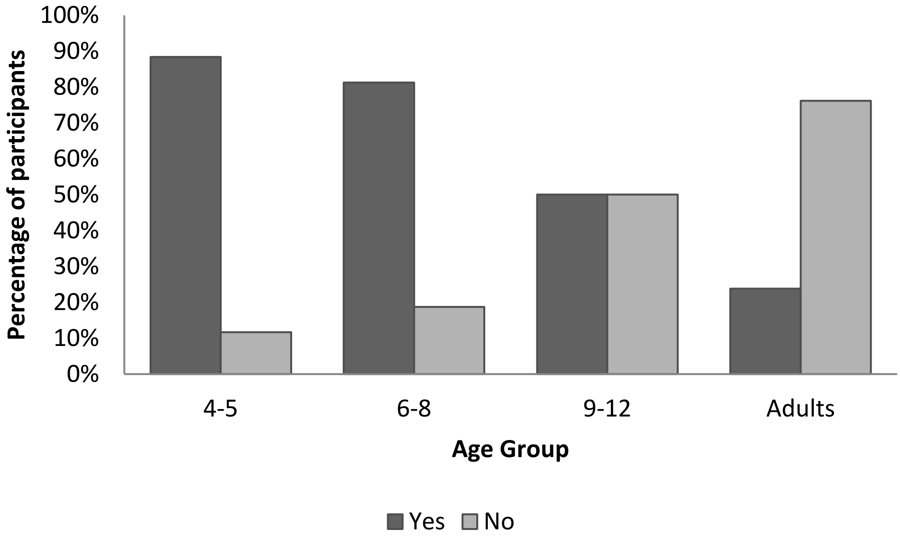


Figure 4.19: Percentage of participants in each age group that interpreted the irrelevant actions as functional

A significant difference between age groups was observed: x^2 (3)=44.81, p<.001. Participants' tendency to identify the irrelevant actions as functional decreased with age. Whereas the majority of participants in the 4-5 and 6-8 age groups identified the irrelevant actions as being necessary for task completion, only half of 9-12 year olds suggested that they were required. Additionally, only 24% of adults reported that the irrelevant action were necessary for task completion.

4.4. Discussion

4.4.1. Summary of findings

- As in experiment two, theory of mind scores increased with age, with all children above the age of nine passing the task.

- Digit span scores also increased with age. However, higher digit span scores were associated with lower quotient scores. There was no significant difference in digit span score observed between the assigned conditions.

- Higher quotient scores were seen in experiment three for both adults and children in comparison to box scores from experiment two.

- As in experiment two, the highest quotient scores were seen in the 6-8 year old age group (both for imitation and demonstration phases). There were no significant differences in quotient score between the imitation and demonstration attempts for any of the age groups.

- The no goal (looped) condition participants achieved the highest quotient scores, but the difference in quotient score was only significant between no goal (looped) and goal (3x) participants in the imitation attempt, and no goal (looped) and goal (3x and looped) for the demonstration attempt.

- Participants' ability to identify a goal rose with age. Participants in the goal conditions (3x and looped) were better able to identify a goal than those in the no goal conditions. The majority of participants in the goal conditions stated that the goal was to retrieve the ball from the box.

- The majority of participants in the 4-5 and 6-8 year old age groups stated that they had tried to copy the demonstrator, whereas the majority of participants in the 9-12 and adult age groups stated that they had not tried to copy.

4.4.2. Differences in irrelevant actions copied between age groups and conditions

In both goal-related groups, the mean quotient score decreased with age, whereas in the no goal (3x) condition, there was no difference in the number of irrelevant actions copied observed between age groups. Additionally, in the no goal (looped) condition, the mean quotient score increased with age. The majority of participants in the goal-related conditions identified the goal of the task as being to retrieve the ball, indicating that any participants in these groups who overimitated did understand the end goal of the puzzle. However, participants in the no goal conditions found it more difficult to identify a clear goal, as expected. When no obvious goal was present in the demonstration video, the youngest participants struggled to suggest a possible purpose for the puzzle. However, 70% of adults in the no goal (3x) condition and 100% of adults in the no goal (looped) condition were able to suggest some sort of intended goal. Additionally, the tendency to copy irrelevant actions from the demonstrator decreased with age as expected.

4.4.3. Interpretation of actions

The belief that the irrelevant actions were functional decreased with age. This provides some support for the suggestion that children overimitate because they believe the irrelevant actions to be necessary for task completion (Lyons, Young, & Keil, 2007). However, even 24% of all adults believed that some of the irrelevant actions were functional in some way, although when these differences were examined further, it was determined that these participants were all in the no goal conditions, in which it was not as easy to determine whether actions were required or not. This may provide support for the unspecified purpose hypothesis, in which participants copy irrelevant actions because their purpose is unclear (Horner & Whiten, 2005). In fact, adult participants were more likely to attempt to explain why the demonstrator might be performing the irrelevant actions, even if the participant themselves believed the action to be non-functional. Example participant responses were:

Participant 32: "She knocked to show you that the lid of the box was solid."

Participant 130: "I think she was just showing you all of the things that you can do with the box."

Participant 131: "She was pretending she didn't know how to do it."

Some 9-11 year olds also attempted to do this:

Participant 21 (aged nine): "She turned it to check everything was okay."

Participant 125 (aged nine): "She knocked to try and move the ball but it didn't work."

Additionally, the interpretations provided differed depending on which assigned condition participants were in. 9-11 year old participants provided the following descriptions of the task:

Goal conditions

Participant 7 (aged ten): "You would flick this part [right lever] up, to make that go down, then spin this part [left lever], which means that it can come into that part and that can slide down there, and it [the ball] can just slide out. [...] I think she was trying to add a few extra steps in to confuse you."

Participant 9 (aged ten): "She did some random things that you didn't really need to do. [...] She was trying to get this [the ball] out and she didn't know how."

No Goal conditions

Participant 37 (aged nine): "First off you pull that up [right lever], then you can turn it around if you want to, like that, to see all of it. then you pull this [top lever] up and back down, and this turning thing [left lever] here, you turn that there, and then turn it back like that."

Participant 74 (aged nine): "If you put a little bouncy ball in, I think then if you turn this [left lever], it will roll down, and bounce onto there [top of chute], and then come through here [the chute]. If you put it on there [top lever] and put it into there [right lever], and then drop it down into there [the chute]. [...] She knocks on here, does that [top lever] and then turns it around, and then puts this [right lever] down and puts this back up. ["Why do you think she knocks"] Does it hold the ball? It's on here [left lever], and then you pull this [top lever] up and then it pushes it into there [the chute]."

In both cases in the goal condition, participants identified that the irrelevant actions were not necessary for task completion, and provided some explanation as to why the demonstrator had chosen to include them. However, in the no goal responses, participants included a description of the irrelevant actions, as well as the irrelevant actions when explaining to "their friend" how to complete the puzzle. Participant 74 also makes some attempt to explain what the purpose of these actions might be. It may be the case that in the absence of a clear goal, the action themselves become encoded as goals (Hernik & Csibra, 2009), which would result in participants displaying higher levels of overimitation in the no goal conditions.

4.4.4. Intention to copy actions

Although no adults in the goal (3x) and goal (looped) conditions copied any of the irrelevant actions, participants in both groups subsequently reported that they had, after being asked the question "did you do everything she [the demonstrator] did?" This may indicate that the question used did not successfully capture the intended behaviour. However, it may also indicate that adults reported that they copied, even when they did not, as they did not want to appear as if they had somehow done the wrong thing. There is some suggestion that the function of overimitation is to strengthen social bonds (McGuigan et al., 2011b), and therefore participants may have stated that they copied in order to display affiliative behaviour.

4.4.5. Differences in irrelevant actions copied between attempts

The mean number of irrelevant actions imitated decreased from participants' first attempt at completing the puzzle in comparison to their second attempt. In fact, some participants made it clear during their second attempt at completing the puzzle that they had initially considered one of the initial actions to be functional, but now realised that it was not necessary for task completion. For this reason, they had omitted it the second time around (e.g. "I'm not sure if this actually does anything or not" – participant 14, aged 10, referring to the lever). These findings are consistent with previous research, which suggests that when participants are able to attempt a task multiple times, irrelevant actions tend to be omitted (Caldwell & Millen, 2009; McGuigan, 2012). However, a significant decrease in the mean number of irrelevant actions demonstrated by participants in the youngest age group was not observed between attempts one and two. This is partially inconsistent with findings by McGuigan & Graham (2010), who determined that in a transparent task, five year olds were able to omit irrelevant actions as they passed down a diffusion chain, but in an opaque task, irrelevant actions were maintained. It may be the case that as some of the youngest participants in this study were unable to identify a goal (particularly in the no goal conditions), then this task can be viewed as opaque (in terms of difficulty), rather than obvious.

Overall, the results of this experiment appear to indicate that, rather than underlying developmental abilities being responsible for overimitation, it is in fact participants' understanding of the goal of the task and the functionality of the demonstrator's actions that are responsible for eliciting high fidelity copying behaviours. These results provide support for both the automatic causal encoding hypothesis (Lyons et al., 2007), and the unspecified purpose hypothesis (Horner & Whiten, 2005). However, as with the previous two experiments, participants only viewed a demonstration video, and so were never in the presence of the demonstrator themselves. This is likely to have reduced many of the social cues that would take place in a more natural, everyday social learning situation, such as eye contact, and so these explanations for overimitative behaviour cannot be discounted. A number of studies have suggested that participants' behaviour changes when in the presence of the individual who demonstrated the task, for example, only performing irrelevant actions when the demonstrator who also performed these actions was present, and omitting the actions in the presence of another adult (Nielsen & Blank, 2011). Additionally, all participants were tested in an environment obviously intended for small children, and all child participants had their parents present. This may have further changed the way that participants might behave in comparison to a more natural setting.

In order to further investigate what might result in differences in behaviour between age groups, developmental factors were investigated.

4.4.6. Theory of Mind

As in experiment two, performance on the theory of mind task increased with age, with all eight to eleven year old participants and adults passing the task. This is consistent with previous research that suggests that false belief task performance increases until approximately age nine (Wimmer & Perner, 1983a). However, unlike in experiment two, no significant relationship was found between theory of mind performance and the number of irrelevant actions copied. However, participants who failed the false belief task were significantly more likely to attempt to copy the demonstrator. This is in direct contrast with previous findings which suggest that individuals wish to overimitate because they understand the motivations of the demonstrator due to the development of theory of mind (Spengler et al., 2010). This result may have occurred due to the fact that only participants in the youngest group failed the theory of mind task, and they may be less likely to understand how the task works.

4.4.7. Working Memory

Participants' performance on the working memory task increased with age as expected. However, contrary to expectations, participants who achieved a higher working memory score were those who were least likely to copy any of the irrelevant actions. However, participants who copied one irrelevant action also achieved the lowest memory score. This is partially consistent with findings by Subiaul & Schilder

(2014), who suggest that children are more likely to use an emulative strategy when working memory load is high. If participants achieve a low working memory score, then they are likely to only copy the action they could remember, even if they wished to copy all actions. Additionally, adults were less likely overall to copy irrelevant actions, but more likely to achieve a higher working memory score, which may have affected the results obtained. It is important to consider that memory is unlikely to be the only factor involved in overimitation, and therefore many participants may have remembered all three actions, but made a conscious choice not to copy them. The looped video conditions were included in order to scaffold younger participants' working memory, allowing them to refer back to the demonstration video if they wished to imitate all actions. This does not appear to have been effective, as no significant differences in the number of irrelevant actions copied were observed between either participants in the goal (3x) and goal (looped) conditions, or between those in the no goal (3x) and no goal (looped) conditions. It is unclear if this is because participants did not use the video to support their memory, or whether they made a conscious choice not to imitate some of the actions. Although there has been some suggestion that children may be more likely to copy when viewing a repeated demonstration in comparison to a limited number of demonstrations (Buchsbaum, Gopnik, Griffiths, & Shafto, 2010; Byrne & Russon, 1998; Nielsen, 2013), this does not appear to have occurred here. Interestingly, the addition of the looped videos did change adults' identification of the goal of the task, with 80% of adults in the no goal (looped) condition suggesting that the purpose of the task was to copy the demonstrator, in comparison to 40% of adults in the no goal (3x) condition.

As in experiments one and two, developmental factors such as working memory and theory of mind appear to have influenced participants' copying behaviours. However, this experiment also provided further information into participants' understanding of the task and the demonstrator's intentions. Additionally, allowing participants to complete the same task multiple times also made their performance more efficient, and often resulted in the omission of irrelevant actions. Further research should consider whether the debriefing questions were truly capturing the information that they were intended to, particularly with the youngest participants.

Chapter 5. Discussion

5.1. Introduction

The overarching aim of this thesis was to determine the underlying factors that drive high and low fidelity copying behaviours in children and adults. Although there is a wealth of research into social learning in these age groups (particularly in younger children) to date, it has a tendency to focus on disparate features (e.g. one task, or one aspect of a demonstration) in isolation, with few attempts made to link them. Whilst this does provide further understanding into what might influence copying strategies used, it remains difficult to determine how factors such as a participants' age, attention and memory ability, and understanding of the goal, interact with the task type used, and instructions provided. It is of particular importance to understand these relationships, as all of these aspects are present in everyday learning situations that do not occur within the constraints of the lab, where some variables may be tightly controlled, or tasks may be more carefully scaffolded. Therefore, the intention was to include a range of tasks, age groups, and additional measurements of participants' developing abilities (e.g. memory, theory of mind, and attention) in order to try to gain an overarching view into what might be subsequently driving their everyday learning, and therefore influencing them to perform as an "overimitator" or an "emulator". It was hoped that this would allow for a better understanding of the end-to-end process of the completion of a social learning task for participants throughout childhood and into adulthood, including how individuals attend to the demonstration, what additional factors might be influencing their ability to complete the task, and finally, how they actually perform the task themselves. To date, little research has included an investigation of the full process, and no research has used such a wide age range of participants to try to understand the development of social learning mechanisms as we age. Four broad aims were identified.

Participants' attention during initial demonstrations of a task was measured using eye tracking, in order to determine whether the "attention phase" of a task subsequently influenced their behaviour when completing the task for themselves. If no differences in gaze patterns between high and low fidelity copiers were present, then this may suggest that additional factors are driving task performance instead. However, if specific gaze patterns for high and low fidelity copiers were identified, what did each of these participant groups focus on? Previous research in this area was limited and suggested conflicting findings.

A range of age groups was used, to attempt to determine whether a developmental trajectory of copying behaviours exists, and if so, what direction this trajectory moves in. Additionally, the role of other developing abilities, such as memory and theory of mind, were investigated, in order to further investigate the differences in high and low fidelity copying behaviours seen in previous research.

Multiple task types were used with the same participant group, in order to attempt to determine whether performance changed across tasks, and whether task features, such as difficulty, influenced copying fidelity. It was hoped that this would indicate whether participants were able to display flexible copying behaviours between tasks. If this was the case, then it would contradict the suggestion that individuals automatically interpret irrelevant actions as being causally functional or necessary for task completion (Lyons et al., 2007).

Participants' understanding of the goal of the task, and of the intentions of the demonstrator was investigated, in order to determine how this subsequently affected task performance. A number of theories explaining overimitative behaviour appear to suggest that individuals copy when they do not understand the purpose of actions, or of the task itself (Horner & Whiten, 2005; Lyons et al., 2007).

Although completion of a social learning task may initially seem like a straightforward process, there are numerous contributing influences involved at each step, as indicated in figure 5.1 below.

Contributing factors during attention phase (demonstration)

Participant must be able to attend to relevant areas of the demonstration.

Participant needs to sustain attention to see the full demonstration.

Contributing developmental factors

Participant needs to:

- Remember content of demonstration

- Remember full sequence of actions and the order in which they took place

- Understand intentions of demonstrator

- Understand the purpose/goal of the task

Contributing social factors

- Demonstrator characteristics may influence behaviour

- Environmental factors may make the process more difficult

- Presence of other individuals may affect task performance

Contributing task factors

Factors that may affect task performance:

- Difficulty of task

- Familiarity with task/novelty of task

- Task type (e.g. heavily goal/reward driven)

- Flexibility of task (e.g. whether there are multiple potential solutions or a fixed number of possible actions)

Figure 5.1: Contributing factors to task performance

Findings relating to each of these contributing factors will be discussed in more detail in this chapter.

5.1.1. Summary of findings

Table 5.1 describes the findings from each of the chapters with regard to features relating to the task, developmental abilities, and assigned conditions.

Table 5.1: Summary of findings for experiments one, two and three

	Experiment 1	Experiment 2	Experiment 3
Differences between assigned conditions	No significant differences in quotient score between conditions for child participants. Adults in copying condition achieved higher quotient scores than those in goal focused or video control conditions.	N/A	Participants in the looped no goal condition achieved higher quotient scores than those in the other conditions. Participants in the goal conditions (3x and looped) were better able to identify a goal than those in the no goal conditions.
Gaze Patterns	No gaze patterns were identified that could differentiate a high fidelity copier from a low fidelity copier.	No gaze patterns were identified that could differentiate a high fidelity copier from a low fidelity copier.	N/A
Theory of Mind	N/A	Scores increased with age. All participants above the age of nine passed.	Scores increased with age. All participants above the age of nine passed.
Memory	N/A	Sequence memory scores increased with age. In the hook and colouring tasks, sequence memory scores also increased with quotient score. Content memory tasks were found to be too easy for most participants.	Digit span scores increased with age. Higher digit span scores were associated with lower quotient scores.
Task Differences	Both adult and child participants achieved higher quotient scores for the box task than the board task.	Participants aged 6-8 achieved the highest score in the box task, and those aged 3-5 and 18+ achieved the lowest scores. Hook and colouring quotient scores increased with age. Participants achieved the highest quotient scores on the colouring task overall.	N/A

5.1.2. Contributing factors during the attention phase

As indicated in figure 5.1, there are two main considerations during the "attention phase" of a social learning task. When viewing the demonstration, the participant must be able to attend to relevant features of the task, and must be able to sustain their attention for the duration of the demonstration in order to see the full sequence of actions and end stage of the task. However, as described previously in chapter 1, very little research has investigated what typically developing participants of various ages actually attend to during a social learning demonstration. For this reason, one of the main aims of the experiments within this thesis was to determine whether participants' copying strategies are influenced at the attention phase, by using eye tracking to identify whether participants who copied more faithfully attended to different areas of interest in comparison to those who used their own strategy to complete the task.

Across all demonstrations in experiments one and two, an increase in ability to sustain attention to the demonstration video was noted when comparing younger participants to older participants. This was in line with expected findings, as the ability to sustain attention overall has been found to increase with age (Hagen & Hale, 1973; Plude & Doussard-Roosevelt, 1989). However, this ability to sustain attention alone does not indicate what participants were attending to during the demonstration. This is arguably more important, as if participants were not attending to actions or movements when they were demonstrated, then they are unlikely to have been able to replicate them if they wished to. Although limited, some previous research has attempted to investigate potential differences in gaze patterns during a demonstration. Taylor & Herbert (2014) determined that infant participants who had imitated actions shown in the demonstration had been more likely to attend to the demonstrator than to irrelevant areas of the video. However, in contrast, Vivanti, Nadig, Ozonoff, & Rogers (2008) found no significant difference in gaze patterns at either the task or the demonstrator's face between typically developing 8-15 year old children who imitated more precisely and those who did not.

Within this thesis, two methods were used in order to assess the importance of the attention phase. In experiment one, an attempt was made to induce particular copying behaviours (e.g. high and low fidelity copying) through the use of specific instructions. It was expected that if attention to the demonstration drove these behaviours, then marked differences in gaze patterns should be observed between those in the copying group and those in the goal-focused group. Although minimal differences were observed between participants achieving higher and lower quotient scores, the findings appear to have been in line with Vivanti, Nadig, Ozonoff, & Rogers (2008), particularly for adult participants, who showed no significant differences between copying groups or fidelity scores across both tasks. These results suggest that attention to the demonstration was not driving participants' subsequent copying behaviours. It is of note, however, that adults in all groups were significantly more likely to attend to relevant areas of the demonstration video (particularly the tasks) in comparison to child participants. These findings are consistent with those found by Taylor & Herbert (2014).

Eye tracking was also used in experiment two, in which a wider age range of participants was used. Unlike experiment one, these participants were not provided with specific instructions on how to complete the task, and so it would be expected that participants' performance here would be closer to the way they would naturally behave in a social learning situation. As in experiment one, the total time attending to relevant areas of the demonstration video increased with age. However, as before, there were no significant differences found between participants who copied faithfully and those who did not, across all three tasks. Even participants in the youngest age group (three to five-year-olds) spent the majority of their time attending to the tasks, albeit not as much time as participants in older groups. This, combined with the findings from experiment one, appears to strongly indicate that it is not attention to specific aspects of a demonstration that drives higher fidelity copying behaviours. Nonetheless, the ability to attend to a demonstration remains important, but this ability appears to have developed enough by the age of three so that participants are able to attend to relevant task features and subsequently attempt the task for themselves.

It is important to note that there were some limitations relating to the eye tracking data collected in experiment one: it was of poorer quality than that collected in experiment two for a number of reasons. Although the Eye Tribe tracker used in this experiment was able to collect usable data for adult participants, it was not appropriate for use with the youngest child participants. This tracker was initially used as it was portable, and could therefore increase potential options for data collection as it could be used outside of the eye tracking lab. However, calibration rates for experiment one were extremely low, and often even when participants completed calibration, too much movement resulted in poor quality data. One way to resolve this would have been to use a chin rest, but it was considered that this might be too distressing or uncomfortable for the youngest participants. In addition, when testing was conducted either in a classroom or in local museums, the environment proved to be too distracting for younger participants, and therefore further reduced the quality of the data. As a result of these issues, data collection for this study took significantly longer than expected. Due to these limitations, the method used to collect data in experiment two was adapted. A more powerful eye tracker, the EyeLink 1000 tracker, was used. This required participants to visit the eye tracking laboratory at Newcastle University with their parents in order to complete the tasks. The rate of calibration and data collection, as well as the quality of data collected improved significantly when using this tracker (from around 40% in experiment one, to around 80% in experiment two). The tracker required that participants wear a target sticker on their foreheads, which allows for a much greater range of movement. This was of particular importance with the youngest participants.

As with experiment one, there were limitations relating to the methodology used in experiment two. OpenSesame software was used to program and display the task to participants, with the intention of analysing data using the DataViewer software provided with the EyeLink tracker. However, this was not possible due to conflicts between the pieces of software used. Reprogramming the task with a different piece of software would have resulted in the loss of previously collected data, so an alternative method of analysis was used, in which individual fixations were displayed frame by frame using MATLAB and were then hand-coded. This method was extremely time consuming, particularly for the longest demonstration videos. This method should be avoided in future by ensuring that analysis can be performed accurately using an appropriate piece of software that is able to work with video stimuli, thereby avoiding the need for frame by frame analysis.

One additional issue with experiment two was the way in which testing was conducted. Participants watched the demonstration video for one task, then looked away from the screen to perform the task, potentially resulting in a fair amount of movement before watching the second demonstration video. Although drift correction was performed between tasks, this may still have resulted in poor calibration and messy data. The use of an alternative eye tracker, such as a glasses-based tracker, that allows for full movement of a participant’s head, may have been a better solution here. Additionally, a glasses-based tracker could potentially be used to determine what aspects of the task a participant focuses on as they actually complete the puzzle, and could then be compared to the areas of interest attended to during the demonstration. This would be particularly useful with younger participants, who struggled to explain the way in which they were performing the task in experiment three, when they were asked to demonstrate the puzzle to a teddy bear. Additionally, this would allow for a comparison between live demonstrations and video demonstrations, as previous research has suggested that the presence of the demonstrator heavily influences participants' subsequent behaviour (Nielsen & Blank, 2011; Wood, Kendal, & Flynn, 2012; Wood, Kendal, & Flynn, 2013). Furthermore, the use of a live demonstrator more closely replicates natural learning situations that participants would encounter in their daily lives, for example, learning from a teacher or parent.

Although there were a number of methodological limitations that may have influenced the findings within this thesis, it is still important to consider the fact that very few studies have previously used this technology with such a wide age range of typically developing participants. The findings from both experiments one and two suggest that both high and low fidelity copiers were able to successfully attend to the demonstration video, indicating that it is not just simply the case that low fidelity copiers are less able to attend to relevant stimuli than high fidelity copiers. Whilst previous research has highlighted increased attention to the demonstrator's face in children with ASD who copy more faithfully (Vivanti et al., 2008), this does not appear to occur in typically developing children, potentially suggesting that participants' copying strategies in experiments one and two were not driven by any kind of affiliation to the demonstrator. However, this is to be expected, since the demonstrator was never present and could therefore not see the way in which participants performed the task. Research by Nielsen & Blank (2011) suggests that individuals modulate their copying strategies depending on whether an efficient or inefficient demonstrator is present, in order to avoid offending the inefficient demonstrator. Therefore, the results obtained in experiments one and two may have been different if a live demonstrator had been present, as discussed above. For this reason, eye tracking comparisons between live and video demonstrations may also be beneficial.

5.1.3. Contributing developmental factors

Much of the research into copying behaviours involving typically developing participants focuses on very young children and adults only (e.g. Flynn & Whiten, 2008; Lyons, Young, & Keil, 2007; Marsh, Ropar, Hamilton, Erdohegyi, & Csibra, 2014; McGuigan, Makinson, & Whiten, 2011; Simpson & Riggs, 2011; Watson-Jones, Legare, Whitehouse, & Clegg, n.d.; Whiten et al., 2016b). This means that there is potentially a gap in our knowledge about how these behaviours develop. This is particularly important, since the behaviours observed in adults and children are often markedly different; it would therefore be beneficial to determine when and why these changes occur. Therefore, one aim of this thesis was to establish whether there was a developmental trajectory of copying behaviours that changed from early to middle childhood, and through adolescence into adulthood. For this reason, experiment two was novel: no other study has looked at the performance of such a wide age range (3-21) across multiple tasks. As indicated in figure 5.1, a number of developmental factors may influence participants' ability to complete a task, including their ability to remember the content of the demonstration, the full sequence of actions, and the order in which they occurred, as well as being able to interpret a goal of the task, and understand the intentions of the demonstrator. These abilities are reliant on the development of memory and theory of mind, and although these abilities do become apparent around the age of four or five (Scarf et al., 2013; Wimmer & Perner, 1983), they continue to develop into middle childhood and adulthood (Barrouillet et al., 2004; Vetter, Altgassen, et al., 2013). Therefore, it may be expected that these developmental changes are driving some of the behavioural changes seen in social learning strategies. Assessments of participants' memory and theory of mind abilities were taken in experiments two and three, but not in experiment one.

In the box task in experiment two, participants in middle childhood and adolescence appeared to overimitate the most, even more than adults, potentially indicating that the driving factor here was not simply memory ability, since adult participants performed just as well on the memory tasks as those in the age groups who copied more faithfully. Instead, the results of this task may provide further support for the unspecified purpose hypothesis (Horner & Whiten, 2005), which suggests that individuals copy actions when their purpose is unclear, in case the demonstrator intentionally performed it for a specific reason. These results were replicated in experiment three, in which participants in middle childhood were also found to display much higher levels of overimitation in comparison to adult participants. These results alone do not indicate a clear developmental trajectory from early childhood to adulthood.

However, in the hook task in experiment two, adults were found to be the most overimitative group, potentially indicating a clearer developmental trajectory than seen in the box task. This increase in overimitative behaviour is to be expected, due to the fact that the hook task was more complex, and therefore copying it precisely was more difficult in comparison to the box task. Due to the increased complexity, it is likely that younger participants would have found it more difficult to remember the precise pieces and order used, and this was reflected in their performance on the sequence and content memory tasks. It is therefore likely that adults' improved memory performance allowed them to better remember what they had seen on the demonstration video. Indeed, previous research has suggested that when the cognitive load involved in a task is too high, younger participants turn to more emulative solutions (Subiaul & Schilder, 2014). Although the colouring task was arguably of similar complexity to the box task in terms of the number of actions in the sequence, adults were again found to be the most overimitative group. Since cognitive load is unlikely to be the driving factor here due to the lower number of actions in the sequence, participants' performance may be better explained by the fact that the goal of this task was less clear. Previous research suggests that when a clear goal is not present, imitation of the steps taken to complete the task becomes the goal instead (Watson-Jones et al., 2014). However, findings from both the hook and colouring tasks are consistent with previous research, which found that adults were "super-copiers" (McGuigan, Makinson, & Whiten, 2011) – this may simply have been driven by their increased memory ability in comparison to child participants, along with social influences.

In experiment two, a relationship between task performance and content memory score was also found, with participants that achieved a lower quotient score also achieving a lower content memory score in comparison to those who copied faithfully in most cases. This may suggest that emulation, or lower copying fidelity generally, sometimes occurs simply because participants cannot remember what they have seen during a demonstration.

However, in experiment three, the relationship between developing memory abilities and task performance were not as clear as in experiment two. There was no significant relationship between participants' performance on the working memory (digit span) task designed to further assess memory and task performance. These conflicting results potentially suggest that the digit span task was not the most appropriate test of sequencing ability to serve as a comparison with participants' copying strategies. The digit span task is reliant on verbal memory, whereas the memory test in experiment two was reliant on visual memory, and since the demonstration is reliant on visual information, it may be the case that a test of visual working memory would have been more appropriate in this case. Differences in visual and verbal memory have been widely studied in children, and it has been suggested that children rely more heavily on visual than verbal memory (Hitch, Woodin, & Baker, 1989). For this reason, a Corsi blocks test, in which children watch a demonstrator touch blocks in a particular order, and are required to repeat this sequence themselves, may have been a more appropriate measure of working memory in this case (Isaacs & Vargha-Khadem, 1989; Vandierendonck, Kemps, Fastame, & Szmalec, 2004). Future research should investigate the relationship between participants' score on this task, and the copying strategy they subsequently use. Alternatively, it may be the case that it is participants' sequencing ability that causes a difference in task performance, rather than their ability to remember all of the actions involved, meaning that unless a task is extremely complex, the participant is likely to remember the actions involved, but perhaps not the order in which they were performed.

Interestingly, no significant relationship was found between the copying strategy used by participants and theory of mind in either experiments two or three, potentially indicating that participants' choice to emulate was not influenced by a lack of understanding of the demonstrator's intentions as previously suggested (Burns, Russell, & Russell, 2015; Huber et al., 2009; Marsh, Pearson, Ropar, & Hamilton, 2014). In experiment three, participants' theory of mind ability was also not significantly related to their ability to identify a goal. These findings appear to provide further support for the rational normative action interpretation hypothesis (Gergely & Csibra, 2003), which suggests that when an individual has not yet developed theory of mind, they consider the situational constraints of a task, in line with their own intended goals, in order to use a rational approach to solve the task. The results obtained in both experiments two and three appear to indicate that basic theory of mind ability is therefore not a driving factor in influencing participants' copying strategy, as they are able to use alternative methods (such as those described by Gergely & Csibra (2003)) to interpret how to complete the task instead.

In experiment three, participants around the age of nine were able to identify different motivations that they had attributed to the demonstrator when explaining why they had chosen to copy or not. However, it is important to note that previous research suggests that theory of mind is continuing to develop at this age and well into adolescence (Devine & Hughes, 2013; Dumontheil, Apperly, & Blakemore, 2010; Schwanenflugel, Henderson, & Fabricius, 1998; Vetter, Altgassen, et al., 2013). It would therefore be of interest to explore explanations behind the demonstrator's actions more fully in these age groups, to determine whether these explanations change with age. This may help to further explain some of the age differences observed.

Findings from experiment three also indicated that participants' ability to identify a goal increased with age, with the youngest participants being most likely to be unable to describe what the goal of the puzzle box task was. This was particularly true in conditions where the reward was not present. Additionally, the belief that the irrelevant actions were functional decreased with age, although adult participants often made attempts to explain the demonstrator's behaviour even when they had chosen to use a different strategy themselves. This can be explained by the fact that participants in the adult group will have fully developed theory of mind. As previously stated, theory of mind continues to develop during adolescence, which would explain the youngest participants' inability to explain the demonstrator's actions, and the fact that some of the nine and ten year old participants offered explanations, whilst others did not. However, only a very simple measure of theory of mind was used both here and in experiment two, as an initial attempt to investigate the influence of this ability on task performance. It may be the case that a more nuanced theory of mind task might capture more subtle differences between participants of different ages and abilities. Future research should consider using more complex theory of mind tasks with older participants (middle childhood and older) to determine whether more nuanced differences in theory of mind do subsequently affect task performance.

5.1.4. Contributing social factors

As indicated in figure 5.1, participants may face multiple social pressures or influences when completing social learning tasks. Due to the fact that even very subtle characteristics of the demonstrator can influence participants' behaviour, and the aim of the experiments included within this thesis was not to investigate the role of these demonstrator characteristics, then an attempt was made to ensure as much consistency as possible between participants by using video demonstrations. However, it is likely that in the absence of a live demonstrator, some important social cues, such as immediate feedback from the demonstrator, were lost. Previous research has indicated that participants alter their behaviour in the presence of different demonstrators (e.g. more or less efficient demonstrators), as well as in the presence of a live demonstrator in comparison to a video demonstration (Marsh et al., 2014; McGuigan, Gladstone, Cook, Macris, & Keil, 2012).

As previously described, participants were assessed on a false belief task in experiment three, and it was found that the ability to pass or fail this task did not influence participants' ability to identify a goal, nor did it affect their copying strategy. Additionally, participants in experiment three who failed the theory of mind task were actually more likely to report that they had attempted to copy the demonstrator than those who passed. However, this may be more of a measure of their wish to affiliate with the demonstrator, rather than their ability to be able to understand the demonstrator's actions or intentions. It would have been beneficial to include some sort of measurement of participants' views on the demonstrator, in order to determine whether they felt that they wished to affiliate with her or not, and how that subsequently affected their performance on the task.

Interestingly, when adults in experiment three were asked whether they had made an attempt to copy irrelevant actions from the demonstrator ("did you do the puzzle in the same way as the person on the video?"), participants in some groups stated that they had attempted to copy when in fact they had not reproduced any of the irrelevant actions. The delay between completing the task and being asked the debriefing questions was very short, suggesting that participants would not have had time to forget what they had done. This, along with adults' higher levels of performance on the memory task, suggests that participants' answers were perhaps driven by conformity in this case (Claidière & Whiten, 2012), potentially indicating that adults' performance is more socially driven than that of children. In experiments two and three, a number of participants in the middle childhood and adolescent groups, as well as adult participants in all experiments, reported concern that they would be unable to do the puzzle prior to seeing the demonstration video. They stated that they would somehow perform the puzzles "wrong", or were worried that they might look "silly" in front of the experimenter if they were unable to perform the task. This is consistent with research that highlights the extent of social pressure felt by teenagers and those in middle childhood (Lashbrook, 2000), and would benefit from further investigation in future research, particularly since there are conflicting findings around imitation in adolescence. Whilst social relationships are considered to be particularly important during adolescence, individuals are still developing their understanding of the perspective of others (Choudhury, Blakemore, & Charman, 2006). Whilst social influences appear to modulate imitative behaviour in adults, this is not the case with adolescent participants (Cook & Bird, 2011). Imitation has been described as a particularly useful process that individuals use in adolescence in order to learn more about themselves and others (Adams & Marshall, 1996), and so further investigation of this age group in particular would be beneficial, particularly since their behaviour is less likely to be driven by developmental factors, as with the younger participant groups.

From the results obtained within this thesis, it appears that social factors do influence participants' task performance. However, as indicated in section 5.1.2, social influences are not the only contributing factors. This is in contrast to previous research that suggests that overimitation is purely socially driven (Marsh, Ropar, Hamilton, Erdohegyi, & Csibra, 2014).

5.1.5. Contributing task factors

Multiple task types were used across experiments one and two: a puzzle box, a tangram puzzle, a hook-building task, and a colouring in task. Although some criticisms of puzzle boxes have been outlined previously (see chapter 1), a puzzle box task was used in each chapter to allow for comparisons to previous research, as well as to achieve a better understanding of the full process, from attention stage to completion stage, that a participant might go through when completing this task.

In experiment one, a magnetic tangram puzzle, similar to that used by Dickerson et al (2013), was used as a comparison against the puzzle box. The puzzle box was likely to be more novel to participants, whilst many participants were more familiar with traditional puzzles. For this reason, it would be expected that overimitation would be higher in the novel puzzle box task, as participants may not have realised that the irrelevant actions (particularly the lever attached to the lid of the box) were not required for task completion. This was found to be the case; the majority of participants were found to overimitate when completing the puzzle box, but not when completing the puzzle board. Additionally, the types of irrelevant actions performed by the demonstrator were different between tasks: while the irrelevant actions in the puzzle box appeared to be functional, the irrelevant actions in the puzzle board task were gesture-based. Previous research has suggested that children are less likely to copy gesture-based actions when they are not directly related to a goal (Bekkering et al., 2000). This was also found to be the case in this experiment: participants appeared to be able to determine that moving the puzzle piece in a particular way before placing it onto the board did not help to achieve the goal state, as the majority of participants omitted these actions. This supports previous suggestions that participants are more likely to use familiar strategies or actions to complete a task when they are able to do so (Pfeifer & Elsner, 2013). However, these results are in direct contrast to those found by Marsh, Ropar & Hamilton (2014), who found that children copied irrational actions even on familiar objects. Nonetheless, the results of experiment one indicated that even the youngest participants were able to alter their copying strategies between tasks, despite the fact that they had been provided with specific instructions on how to complete the tasks. This gives some preliminary indication that overimitation is not automatic, and directly contradicts the automatic causal encoding hypothesis (Lyons et al., 2007), which suggests that children automatically interpret irrelevant actions as necessary for task completion, and therefore overimitate in all tasks.

In order to investigate differences in task types further, experiment two also included multiple task types, including a different familiar (yet arbitrary) task. In this case, participants were not provided with specific instructions on how to perform the task as they were in experiment one. The inclusion of alternative comparison tasks (both novel and familiar) meant that further conclusions could be drawn about what influences participants to copy particular strategies, movements or actions. In contrast to the board task in experiment one, it was expected that overimitation would be highest in the colouring task, as previous research had suggested that in the absence of a goal, the steps taken to complete the task become the goal themselves (Schachner & Carey, 2013). This is different to the board task, in which there was arguably a more obvious goal: creating the fish shape. Additionally, the range of options for task completion in the board task was much more limited, with only four puzzle pieces being provided. Participants in all age groups were found to be most overimitative in the colouring in task, as expected, due to the lack of obvious goal. This was in direct contrast to the behaviour seen in the board task, in which very few participants were found to copy the irrelevant gestures. However, as previously stated, this may be due to the fact that the irrelevant aspects of the task were gesture-based. A similar pattern of behaviour to that seen in the colouring task was observed in the "no goal" conditions in experiment three, in which the reward had been removed from the puzzle box, further supporting the suggestion that when an obvious goal is not present, participants view imitation of the sequence of actions as the goal instead. This provides further support for the unspecified purpose hypothesis (Horner & Whiten, 2005), in which participants copy actions when their purpose is unclear, in case they are necessary for task completion.

In contrast to previous research (McGuigan, Makinson, & Whiten, 2011; Whiten et al., 2016), and findings in experiment one, participants in all age groups were found to overimitate the least in the box task. However, it is important to note that the puzzle box used in this thesis was much easier to solve than those used in previous research, meaning that participants may have been able to understand the inner mechanisms of the box more easily. This would reduce the likelihood of overimitation due to either automatic causal encoding of actions as necessary for task completion (Lyons et al., 2007), and might explain why lower overall levels of overimitation were seen in the puzzle box tasks used in this thesis in comparison to findings from previous research.

Despite this, it is of interest that the findings from the puzzle box tasks in experiments one and two were so different, even for participants in the video control group in experiment one, who were provided with the same prompt ("now it's your turn") as those in experiment two. This further supports the suggestion that puzzle box tasks may not be reliable measures of social learning behaviours used in isolation. Future research should aim to further investigate differences across tasks, using the same participant group, in order to determine whether these findings can be replicated to further support the suggestion that overimitation is flexible and task-dependent (Keupp, Bancken, Schillmöller, Rakoczy, & Behne, 2016b). This research should also include follow up questions that aim to elicit participants' understanding of the rationality of the actions involved in the task, and the goal of the task. It is important that these are investigated in conjunction, as in experiment three, some participants identified that the irrelevant actions were unlikely to be necessary to complete the task, and yet they chose to copy them anyway, "just in case", as described in the unspecified purpose hypothesis (Horner & Whiten, 2005). Without this knowledge, this overimitation could have been misinterpreted as automatic, or as participants interpreting the actions as absolutely necessary for task completion.

Results from the hook task were of particular interest, as it was arguably the most open-ended of the three tasks. Participants were provided with the pieces required to precisely copy the hook from the demonstrator, but also with a number of additional pieces, increasing the range of possible outcomes. Whilst the box task could eventually be solved by trial and error without ever having viewed the demonstration video, this was much more difficult with the hook task. The number of potential actions involved in building the hook was much larger, meaning that this task is "ill-structured", in that the steps involved in achieving the outcome are not as well-defined or limited as the other tasks (Cutting et al., 2011). Additionally, previous research has indicated that children are very poor at tool innovation (Beck et al., 2011; Cutting et al., 2014). However, there is some suggestion that this is due to childrens' perceived "ownership" of the materials required to build tools, see Sheridan, Konopasky, Kirkwood, & Defeyter (2016). The difficulty experienced by participants when trying to manufacture tools in this way suggests that they would need to rely heavily on a demonstration to achieve a successful outcome. Indeed, previous research has suggested that learning about the functions of tools from another individual is much more efficient than attempting to produce a tool without any help (Gardiner, Bjorklund, Greif, & Gray, 2012). It would therefore be expected that levels of overimitation on this task would be very high, in order to increase efficiency and reduce the need for innovation. In fact, overimitation in the hook task increased with age, with older participants showing increased copying precision, even down to the level of individual pieces used by the demonstrator. The youngest participants (aged three) showed high levels of perseveration, using all available materials to make the tool, even at the cost of efficiency and time, and despite the fact that they were able to identify the goal of the task when prompted. Unfortunately, as language skills are still developing in three year olds, it was difficult to determine why they behaved in this way. However, Sharon & DeLoache (2003) suggest that young children use perseveration as a method to increase their symbolic understanding relating to objects. If this is the case, then the youngest children may still have fully understood what they were attempting to do (use the pieces to build a hook to retrieve the bucket and get a sticker), but not that they were able to leave some of the pieces to one side and still produce a functional hook in order to complete the task. Although the majority of work into childrens' understanding of tool building is related to the innovation of tools, rather than tool manufacture following a demonstration, more information regarding the processes that children go through in order to build a tool for a particular task will allow for a better understanding into why children are so poor at tool innovation.

It is particularly interesting that only the adult participant group was able to achieve the highest levels of precision when building the hook, in comparison to all other tasks, across all three experiments, in which younger participant groups were better able to copy precisely. This is especially true when copying the hook increases efficiency significantly, and reduces the need for innovative behaviour, especially in comparison to the board, box and colouring tasks. These tasks could be solved easily using trial and error (in the case of the box and board tasks) or innovative behaviour (in the case of the colouring task, by using alternate colours if the goal is perceived to be to colour the shapes). This may suggest that the level of complexity involved in this task, as well as the cognitive load required to remember the pieces and sequence used, might have caused this marked difference in performance. Future research should compare this tool manufacture task with an equally complex non-tool task, perhaps involving a puzzle box, to determine whether the driving factor in this behaviour truly is the complexity of the task, or whether there is something different about tool manufacture tasks that makes them more difficult to copy precisely.

5.2. Directions for future research

Future research should address two main concerns:

5.2.1. Task development

Not all of the tasks used were entirely successful in fully investigating potential copying behaviours. The board task in experiment one and the colouring task in experiment two were both very limited in terms of the number of potential actions, as well as likely being familiar to participants. It is likely that this reduced the potential spectrum of copying behaviours that could be observed. Future research should aim to include tasks that allow for the investigation of more subtle differences in participants' copying behaviours. Whilst the box task was novel to participants and therefore potentially more successful, the range of actions demonstrated were very limited, and only two functional actions were used. The addition of more functional aspects to make the task more complex and therefore more causally opaque would potentially allow for more nuanced copying behaviours to be observed, in both children and adults. Additionally, the complexity of the box task should be increased in order to match that of the hook task, so that better comparisons between tasks can be made. In this way, if a participant performs differently on the box task to the hook task despite equal complexity, then it can be suggested that some other factor is driving these changes in task performance. This complexity should consider both the number of steps involved, but also the manual dexterity required by participants to be able to solve the task, as this was markedly different between the hook and box tasks used in experiment two, and is likely to have influenced performance by younger participants. The inclusion of a purely gesture-based task would also be beneficial as a comparison task, replacing the colouring and board tasks. Gesture based tasks often remove the goal state entirely, and would allow for better inferences to be drawn about the differences between highly goal-driven tasks, and goal-less tasks.

5.2.2. Participant age ranges

As previously stated, there has been minimal research focusing on the influences on high and low fidelity copying behaviours throughout middle childhood and adolescence. Imitation is considered to be an important mechanism by which adolescent individuals learn about the world around them, and about themselves (Adams & Marshall, 1996), and it would therefore be beneficial to have a fuller understanding of the processes that occur when an adolescent individual is learning a new skill from someone else. Additionally, research has determined that adolescents use more flexible learning strategies in comparison to those in middle childhood and adults (Gopnik et al., 2017), and this merits further investigation with relation to overimitation. Furthermore, imitation is considered to be a highly social process which may prevent ostracism from social groups (Over & Carpenter, 2012; Watson-Jones et al., 2014), and since adolescents are particularly susceptible to peer pressure and conformity (Lashbrook, 2000), it would be interesting to further determine the effects of this on participants' behaviour in social learning tasks.

5.3. Methodological contributions of this thesis, limitations, and considerations for future research

5.3.1. Eye Tracking

The experiments in this thesis reflect one of the first attempts to use eye tracking across a wide range of age groups, using typically developing participants, to attempt to determine whether specific eye tracking patterns relating to high or low fidelity copiers could be identified. As indicated in both experiments one and two, no clear significant results were found that suggest that particular copying behaviours are influenced during the "attention phase", or the demonstration of a task. This is not to say that eye tracking may not be beneficial within social learning research, but instead that attention to particular aspects of a task do not appear to be influencing subsequent performance. As previously discussed, there were a number of limitations faced both when collecting the data, and when analysing it. Whilst a portable tracker allowed for data to be collected in a wider range of environments, this was at times a limitation in itself, due to distractions in either classrooms or local museums that resulted in a loss of data. Using a more sophisticated tracker resulted in better data collection, but required participants to visit the eye tracking lab with their parents. This made testing of adolescents more difficult, as they were less likely to attend specific appointments with their parents, but could not provide their own consent. Analysis of the eye tracking data was extremely time consuming, requiring frame by frame analysis of hundreds of videos. This is avoidable when using appropriate software to program a task and collect the data: these factors should be considered carefully.

5.3.2. Content memory

The memory tasks used in experiment two were designed to investigate two aspects of what might be considered episodic memory: what and when. Although the sequence memory tasks (when) used in experiment two provided some interesting findings, the content memory tasks (what) were found to be too easy. It was difficult to find a balance between images that contained content that was very obviously different from that shown in the demonstration video, and images that were very subtly different, making the task too difficult for all but the oldest age groups. One solution may have been to test a range of images with a number of participant age groups prior to beginning any testing, however, time constraints due to delays from experiment one did not allow for this. It may also have been possible to use different sets of content memory test images for older and younger participant age groups, but this would have made comparisons between all age groups more difficult.

5.3.3. Use of questions

Experiment three relied heavily on the use of language, unlike experiments one and two, since participants were asked a number of questions at the end of the experiment to determine what they understood about the task. It is not clear as to how much the youngest participant groups actually understood what they were being asked. Although other research does use simple debriefing questions (e.g. see Whiten et al.,(2016)), the use of an additional language ability test would have helped to indicate whether the youngest participants were capable of understanding what they had been asked.

5.3.4. Tasks used

Some of the tasks used within this thesis were more successful in achieving their aims than others. Whilst the puzzle box task was designed to be most like those commonly used within the literature, it was much more simplistic and did not allow for such a nuanced investigation into high and low copying behaviours that a more complex task might have done. Similarly, the board task did not yield the expected results, possibly due to its familiarity. In the majority of tasks, the irrelevant actions were too obviously irrelevant: when designing tasks for future research, care should be taken to try to include more subtle irrelevant actions.

5.3.5. Quotient Score

As described previously, a quotient score was used across all experiments, which allowed for more effective comparisons between tasks. Additionally, investigations into the order of actions performed, and whether specific irrelevant actions were performed or not were conducted. However, the use of a quotient score does not allow a participant to be identified specifically as an emulator, an imitator or an overimitator. Initially, the intention was to classify participants in this way, using specific criteria for each task in order to categorize participants. However, this made comparisons between tasks difficult, and also resulted in extremely uneven sample sizes which subsequently made statistical analysis difficult. It is acknowledged that the use of a quotient score may not be as effective in identifying specific social learning behaviours, making it potentially less meaningful than the use of strict categories. Additionally, this may make comparisons with other research difficult. Despite this, it was felt that quotient scores were the best method to use in order to allow for comparisons across tasks and across experiments, which was one of the key aims in this thesis. As stated before, when categorising participants based on their performance in the tasks used in experiments one and two, sample sizes were found to be very uneven in some groups, particularly the emulation group. Using tasks with more nuanced actions, or even simply more actions, might have avoided this problem and allowed for this categorisation method to be used.

5.4. Links to theories of social learning

A number of current theories of social learning were introduced in chapter 1: these are discussed with respect to findings from experiments one, two and three below.

5.4.1. Rational Normative Action Interpretation

The rational normative action interpretation hypothesis suggests that since children have not yet developed theory of mind, they cannot use this ability to attempt to understand the intentions of a demonstrator, and so have to interpret situational constraints of a task when completing it as rational or not, to determine what to copy (Gergely & Csibra, 2003). For this reason, children may be more likely to perceive all of the actions performed by the demonstrator to be a conventional part of the activity. If this is the case, then it would be expected that there would be a marked difference in task performance between participants who could pass the theory of mind task, and those who could not, with the younger participants imitating more irrelevant actions than the older participants. However, in the hook and colouring tasks, this was not the case: copying fidelity actually increased with age. In the box task, copying fidelity peaked at the age of 6-8 in experiments two and three, with the lowest quotient scores in the 3-5 year old and adult age groups. These results do not appear to be consistent with the rational normative action interpretation hypothesis, suggesting that theory of mind is not necessarily the only driving factor behind changes in copying fidelity.

5.4.2. Social and affiliative theories

Over & Carpenter (2012) suggest that overimitation is driven by affiliation with the demonstrator. In this case, affiliation may have been measured by the amount of time that the participant spend attending to the demonstrator themselves during the demonstration. In experiments one and two, no distinct gaze patterns were found to suggest that those who looked more at the demonstrator displayed different copying behaviours in comparison to those who did not. However, this may not necessarily contradict affiliative theories, since no live demonstrator was present to provide true ostensive cues, or real-time feedback. Since the demonstrator was not present, the participant may have not only felt less affiliation, but also may have felt less social pressure to perform the task in the same way as them for fear of ostracism. Research using real-time eye tracking (for example, with the use of portable eye tracking glasses) might allow for further exploration of affiliative and social theories, by allowing for comparisons between participants who spend more time attending to the demonstrator in real time, and those who spend more time attending to the task or the rest of their environment.

5.4.3. Unspecified Purpose Hypothesis

The unspecified purpose hypothesis states that individuals imitate when the purpose of actions are unclear (Horner & Whiten, 2005). This means that whilst participants understand that the actions may be unnecessary for task completion, they choose to perform them “just in case” they are required. This was found to be the case in experiment 3, when some participants performed actions in the imitation phase, and then chose to omit them in the demonstration phase when they realised that they were no longer required. However, if this was the only explanation for overimitation, then it would be expected that participants would copy all irrelevant actions in all tasks, but this was not the case in the board or hook tasks, where most participants chose to omit the movement or gesture-based actions shown in the demonstration.

5.4.4. Automatic Causal Encoding

The automatic causal encoding hypothesis states that children automatically interpret all actions as necessary for task completion the first time that they see a task being performed. If this was the case, then it would be expected that children would overimitate across all tasks, but differences in task performance in experiments one and two suggest that this is not the case: child participants achieved lower overall mean quotient scores in the board task than in the box task in experiment one. This suggests that participants were able to use rational interpretation when deciding which actions to perform.

As previously discussed, it is unlikely that one hypothesis on its own could explain the variation in behaviour observed between age groups, and even between individuals. However, the experiments within this thesis are subject to a number of limitations, as discussed in this chapter, and so may not have fully explored the role of these theories in overimitation.

5.5. Conclusions

The experiments within this thesis add to the existing social learning literature, by using an expanded age range, multiple task types, and eye tracking to further understanding into emulative, imitative and overimitative behaviour in children, teenagers and adults. Prior to this, very few studies were conducted using multiple task types or eye tracking methodology, and to my knowledge, no previous studies have included such a wide age range of participants as in that tested in experiment two. Whilst much of the existing social learning research focuses on individual mechanisms in order to explain the variation observed in copying behaviours in young children and adults, the experiments within this thesis aimed to incorporate a wide range of developmental, social and task-based characteristics that have allowed for a novel preliminary investigation into the social learning process from beginning to end. The results of this research indicate that there are likely to be multiple underlying factors that drive high fidelity copying, including developmental abilities such as memory, social factors, and an understanding of the task and its goals, as well as the type of task being completed. Future research would benefit from the continued exploration of a range of tasks, preferably tested using the same participant group, to find a true assessment of participants’ copying flexibility in a range of situations, both simple and complex. This will avoid the current over-reliance on puzzle box tasks, and will allow for a more natural range of behaviours to be explored across a wide age span, particularly when some of the apparent driving factors behind these behaviours, such as the ability to remember a sequence of actions, continue to develop throughout childhood and into adulthood. Overall, it appears that the factors influencing participants’ task performance are complex, and are not simply driven by one underlying ability, theory or aspect. In fact, it is likely that these influences may change over the lifespan, from developing memory abilities in early childhood, to social pressures in adolescence and adulthood, and differences in experience and understanding of the tasks involved, obtained with age, which would explain the conflicting results obtained across the social learning literature. For this reason, it is vital that future research continues to include a wide range of participants, rather than the traditional focus on very young children and adults only.

Appendices

Appendix A: Examples of fixations coded for each of the areas of interest in the box task

Fixations are indicated by the blue cross, which has been enlarged in order to be seen more easily.

Box

The box area of interest refers to any “empty” area of the box.

Levers

The levers area of interest refers to any of the three levers present on the box: the right lever, left lever and non-functional top lever. Fixations located on any part of the lever (i.e. inside or outside the box) were coded under the lever AOI.

Face

The face area of interest refers to any fixations directed towards the demonstrator's face.

Hands

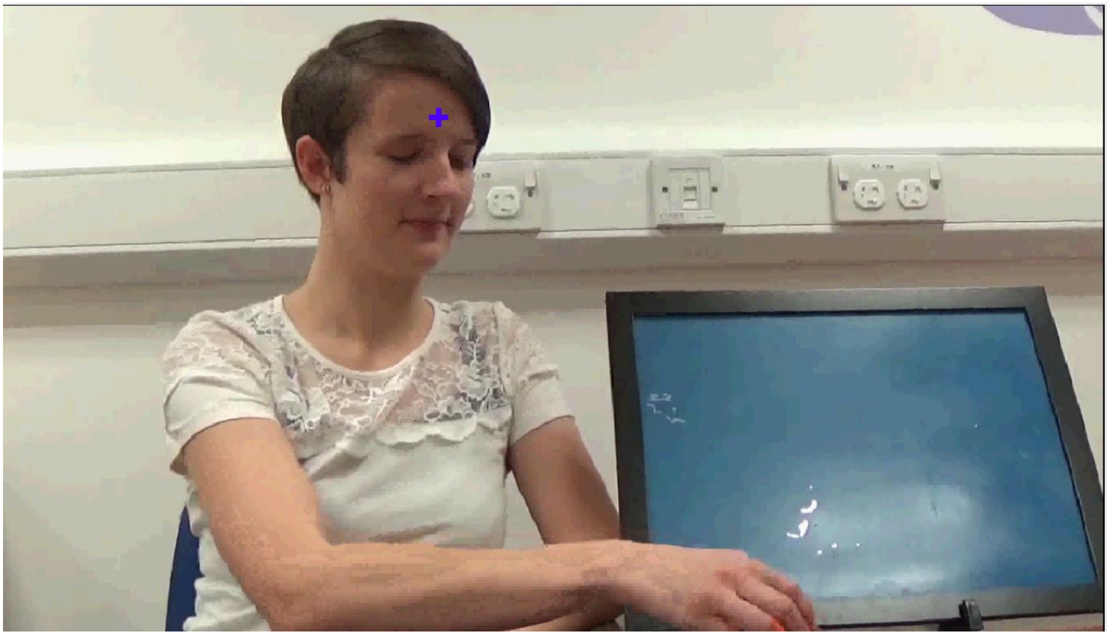
The hands area of interest refers to any fixation directed towards the demonstrator's hands.

Appendix B: Examples of fixations coded for each of the areas of interest in the board task

Fixations are indicated by the blue cross, which has been enlarged in order to be seen more easily.

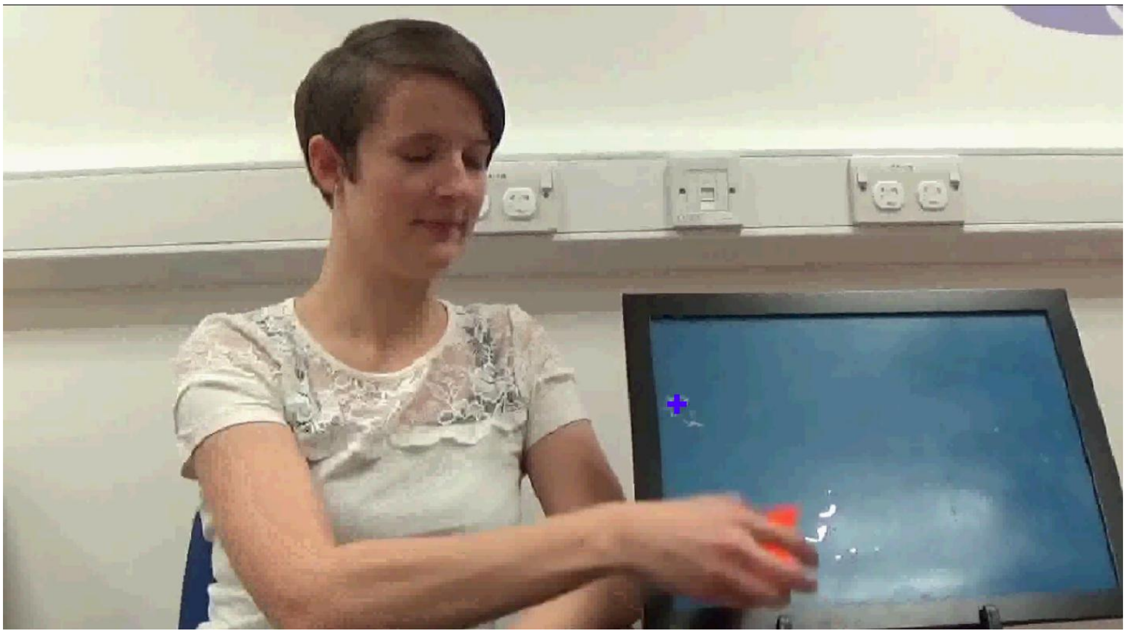
Face

The face area of interest refers to any fixations directed towards the demonstrator's face.



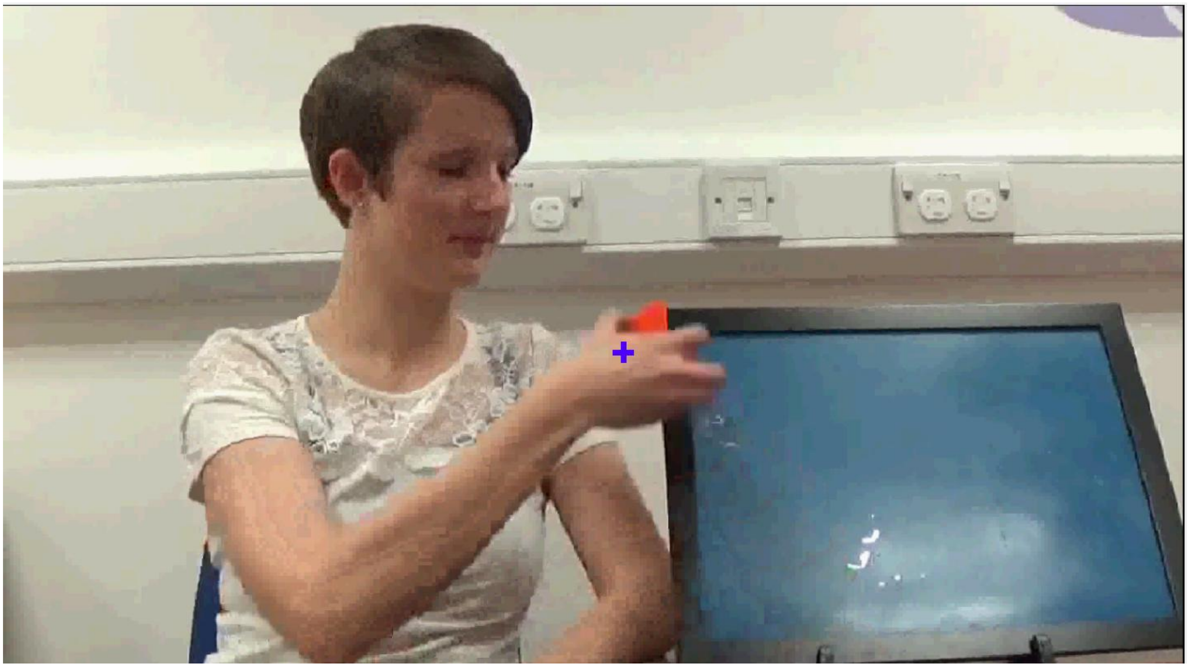
Board

The board area of interest refers to any fixations on the board itself, that do not cover either the puzzle pieces or the demonstrator's hands.



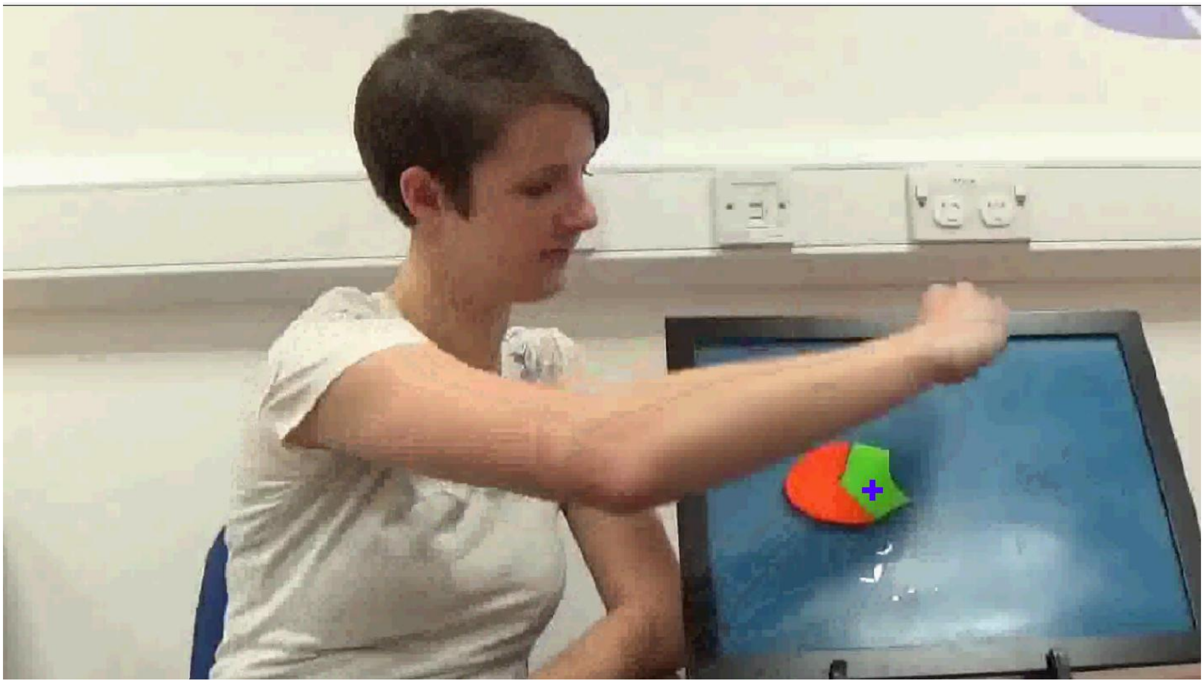
Hands

The hands area of interest refers to any fixation directed towards the demonstrator's hands.



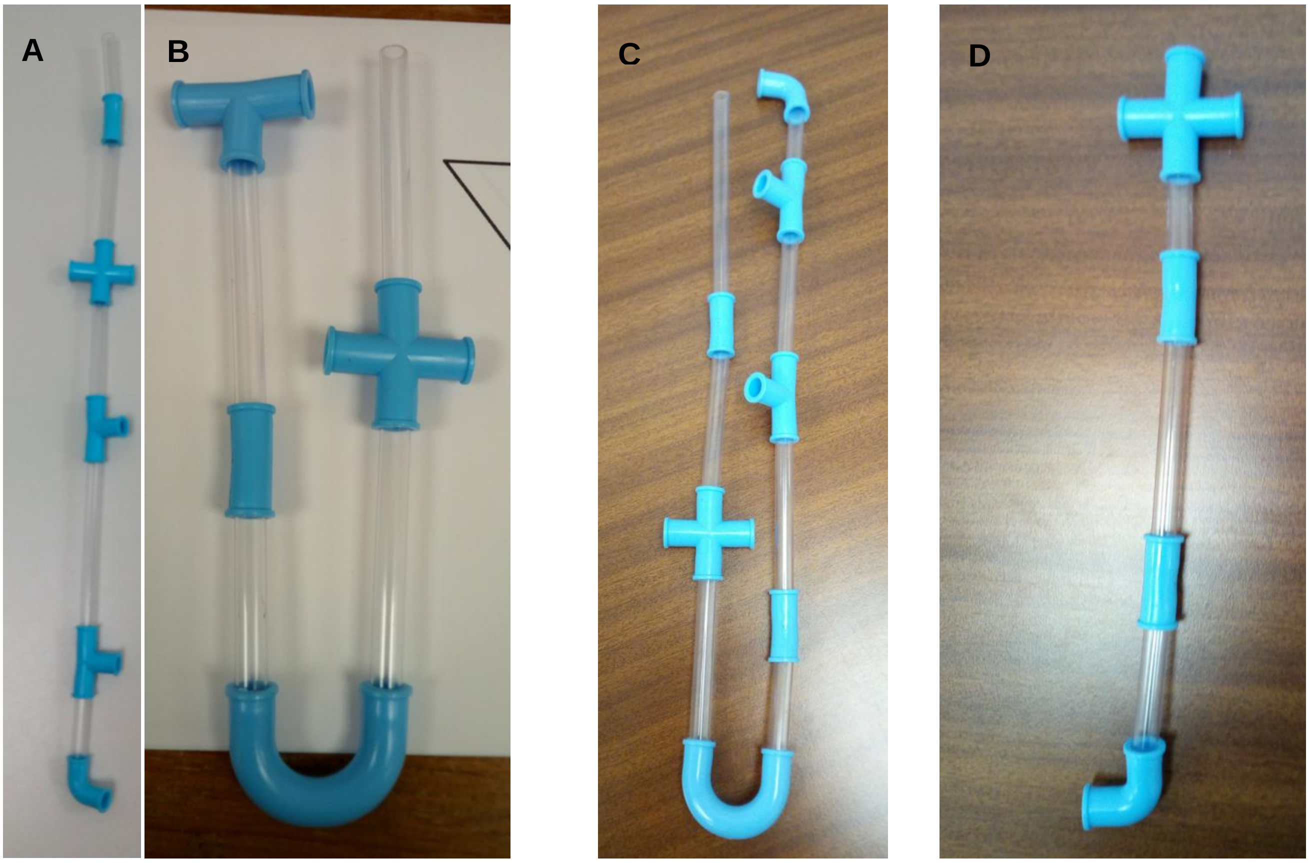
Pieces

The pieces area of interest refers to any fixation directed towards the puzzle pieces.



Appendix C: Example hooks produced by participants

Below are examples of hooks produced by participants that differed to the one shown in the demonstration. In each of these examples, participants were able to use the hook they had constructed to retrieve the bucket from the tube.



Appendix D: Forward Digit Span task used with all participants

Digit Span

I am going to say some numbers. Listen carefully, and when I am finished, say them after me. Ready?

Practice: 6-1-2	

Ready? **3-8-6**	
Ready? **3-4-1-7**	
Ready? **8-4-2-3-9**	
Ready? **3-8-9-1-7-4**	
Ready? **5-1-7-4-2-3-8**	
Ready? **1-6-4-5-9-7-6-3**	
Ready? **5-3-8-7-1-2-4-6-9**	

References

ADAMS, G. R., & MARSHALL, S. K. (1996). A developmental social psychology of identity: understanding the person-in-context. *Journal of Adolescence*, *19*(5), 429–442. https://doi.org/10.1006/JADO.1996.0041

Akers, R. L., & Lee, G. (1996). A Longitudinal Test of Social Learning Theory: Adolescent Smoking. *Journal of Drug Issues*, *26*(2), 317–343. https://doi.org/10.1177/002204269602600203

Anderson, D. R. (2005a). Television and Very Young Children. *American Behavioral Scientist*, *48*(5), 505–522. https://doi.org/10.1177/0002764204271506

Anderson, D. R. (2005b). Television and Very Young Children. *American Behavioral Scientist*, *48*(5), 505–522. https://doi.org/10.1177/0002764204271506

Anderson, D. R., & Levin, S. R. (1976). Young Children's Attention to "Sesame Street" *Child Development*, *47*(3), 806. https://doi.org/10.2307/1128198

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, *70*(9), 1–70. https://doi.org/10.1037/h0093718

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, *35*(5), 1311–1320. https://doi.org/10.1037/0012-1649.35.5.1311

Badger, J. R., & Shapiro, L. R. (2012). Evidence of a transition from perceptual to category induction in 3- to 9-year-old children. *Journal of Experimental Child Psychology*, *113*(1), 131–146. https://doi.org/10.1016/j.jecp.2012.03.004

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*(3), 110–118. https://doi.org/10.1016/j.tics.2009.12.006

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of Aggression Through Imitation of Aggressive Models. *Journal of Abnormal and Social Psychology*, *63*, 575–582. Retrieved from www.all-about-psychology.com

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind" ? *Cognition*, *21*(1), 37–46. https://doi.org/10.1016/0010-

0277(85)90022-8

Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6- to 24-month-old infants. *Infant Behavior and Development*, *19*(2), 159–170. https://doi.org/10.1016/S0163-6383(96)90015-6

Barr, R., & Hayne, H. (1999). Developmental Changes in Imitation from Television during Infancy. *Child Development*, *70*(5), 1067–1081. https://doi.org/10.1111/1467-8624.00079

Barrouillet, P., Barrouillet, P., Bernardin, S., De Bourgogne, U., Camos, V., & V, U. R. D. P. (2004). Time constraints and resource sharing in adults' working memory spans. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: GENERAL*, 83--100. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.379.9208

Bauer, P. J., & Kleinknecht, E. E. (2002). To " ape " or to emulate ? Young children ' s use of both strategies in a single study, (1996), 18–20.

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

Bekkering, H., Wohlschlager, A., & Gattis, M. (2000). Imitation of Gestures in Children is Goal-directed. *The Quarterly Journal of Experimental Psychology Section A*, *53*(1), 153–164. https://doi.org/10.1080/713755872

Berl, R. E. W., Hewlett, B. S., Murray, K., Lu, H., Nédélec, Y., & Pacis, A. (2015). Cultural Variation in the Use of Overimitation by the Aka and Ngandu of the Congo Basin. *PLOS ONE*, *10*(3), e0120180. https://doi.org/10.1371/journal.pone.0120180

Blairy, S., Herrera, P., & Hess, U. (1999). Mimicry and the Judgment of Emotional Facial Expressions. *Journal of Nonverbal Behavior*, *23*(1), 5–41. https://doi.org/10.1023/A:1021370825283

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2010). Childrenâ€™s imitation of causal action sequences is influenced by statistical and pedagogical evidence. *COGNITION*. https://doi.org/10.1016/j.cognition.2010.12.001

Burns, P., Russell, C., & Russell, J. (2014). Preschool children's proto-episodic memory assessed by deferred imitation. *Memory (Hove, England)*, (November), 1–21. https://doi.org/10.1080/09658211.2014.963625

Burns, P., Russell, C., & Russell, J. (2015). Preschool children's proto-episodic memory assessed by deferred imitation. *Memory*, *23*(8), 1172–1192. https://doi.org/10.1080/09658211.2014.963625

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective Imitation of In-Group Over Out-Group Members in 14-Month-Old Infants. *Child Development*, *84*(2), 422–428. https://doi.org/10.1111/j.1467-8624.2012.01860.x

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, *21*(5). https://doi.org/10.1017/S0140525X98001745

Cacioppo, J. T., & Patrick, W. (2008). *Loneliness : human nature and the need for social connection*. Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=w8pWZ2AGI4MC&oi=fnd&pg=PR9&dq=affiliation+beneficial+humans+social&ots=lF8K8yr81X&sig=i7TkiD2UtZ6DO1FV1XJayqwGMkI#v=onepage&q&f=false

Caldwell, C. A., & Millen, A. E. (2009). Social Learning Mechanisms and Cumulative Cultural Evolution. *Psychological Science*, *20*(12), 1478–1483. https://doi.org/10.1111/j.1467-9280.2009.02469.x

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. https://doi.org/10.1037/a0026828

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 (Washington, D.C. : 1983)*, *126*(2), 161–9. https://doi.org/10.1037/a0026828

Carmo, J. C., Gonçalves, F., Souza, C., Pinho, S., Filipe, C. N., & Rumiati, R. I. (2017). Over-imitation in autism spectrum disorder: causally opaque and transparent actions. *Journal of Cultural Cognitive Science*, *1*(2), 77–87.

https://doi.org/10.1007/s41809-017-0010-6

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*(2), 315–330. https://doi.org/10.1016/S0163-6383(98)90009-1

Carpenter, M., & Call, J. (2002). The chemistry of social learning. *Developmental Science*, *5*(1), 22–24. https://doi.org/10.1111/1467-7687.00199

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science*, *8*(1), F13–F20. https://doi.org/10.1111/j.1467-7687.2004.00385.x

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i–vi, 1-143. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9835078

Carr, K., Kendal, R. L., & Flynn, E. G. (2015). Imitate or innovate? Children's innovation is influenced by the efficacy of observed behaviour. *Cognition*, *142*, 322–332. https://doi.org/10.1016/J.COGNITION.2015.05.005

Chappell, J., Cutting, N., Apperly, I. A., Beck, S. R., & B, P. T. R. S. (2013). The development of tool manufacture in humans : what helps young children make innovative tools ? The development of tool manufacture in humans : what helps young children make innovative tools ? Author for correspondence :, (October).

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development*, *15*(4), 481–498. https://doi.org/10.1016/S0885-2014(01)00037-5

Chartrand, T. L., & van Baaren, R. (2009). Chapter 5 Human Mimicry. *Advances in Experimental Social Psychology*, *41*, 219–274. https://doi.org/10.1016/S0065-2601(08)00405-X

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–

239. https://doi.org/10.1016/j.tics.2012.02.007

Choudhury, S., Blakemore, S.-J., & Charman, T. (2006). Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience*, *1*(3), 165–174. https://doi.org/10.1093/scan/nsl024

Claidière, N., & Whiten, A. (2012a). Integrating the study of conformity and culture in humans and nonhuman animals. *Psychological Bulletin*, *138*(1), 126–145. https://doi.org/10.1037/a0025868

Claidière, N., & Whiten, A. (2012b). Integrating the study of conformity and culture in humans and nonhuman animals. *Psychological Bulletin*, *138*(1), 126–145. https://doi.org/10.1037/a0025868

Clay, Z., & Tennie, C. (2017). Is Overimitation a Uniquely Human Phenomenon? Insights From Human Children as Compared to Bonobos. *Child Development*. https://doi.org/10.1111/cdev.12857

Coates, T. J., Petersen, A. C., & Perry, C. (1982). *Promoting Adolescent Health : a Dialog on Research and Practice.* Elsevier Science.

Cook, J., & Bird, G. (2011). Social attitudes differentially modulate imitation in adolescents and adults. *Experimental Brain Research*, *211*(3–4), 601–612. https://doi.org/10.1007/s00221-011-2584-4

Corriveau, K., & Harris, P. L. (2009). Choosing your informant: weighing familiarity and recent accuracy. *Developmental Science*, *12*(3), 426–437. https://doi.org/10.1111/j.1467-7687.2008.00792.x

Costanzo, P. R., & Shaw, M. E. (1966). Conformity as a Function of Age Level. *Child Development*, *37*(4), 967. https://doi.org/10.2307/1126618

Courage, M. L., Reynolds, G. D., & Richards, J. E. (2006). Infants' Attention to Patterned Stimuli: Developmental Change From 3 to 12 Months of Age. *Child Development*, *77*(3), 680–695. https://doi.org/10.1111/j.1467-8624.2006.00897.x

Custance, D., Prato-Previde, E., Spiezio, C., Rigamonti, M. M., & Poli, M. (2006). Social learning in pig-tailed macaques (Macaca nemestrina) and adult humans (Homo sapiens) on a two-action artificial fruit. *Journal of Comparative Psychology*, *120*(3), 303–313. https://doi.org/10.1037/0735-7036.120.3.303

Custance, D., Whiten, A., & Fredman, T. (1999). Social learning of an artificial fruit task in capuchin monkeys (Cebus apella). *Journal of Comparative Psychology*, *113*(1), 13–23. https://doi.org/10.1037/0735-7036.113.1.13

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

Cutting, N., Apperly, I. a, Chappell, J., & Beck, S. R. (2014). The puzzling difficulty of tool innovation: why can't children piece their knowledge together? *Journal of Experimental Child Psychology*, *125*, 110–7. https://doi.org/10.1016/j.jecp.2013.11.010

Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., & Laland, K. N. (2012). Identification of the social and cognitive processes underlying human cumulative culture. *Science (New York, N.Y.)*, *335*(6072), 1114–8. https://doi.org/10.1126/science.1213969

Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E., & Kendal, R. L. (2014). Human cumulative culture: a comparative perspective. *Biological Reviews of the Cambridge Philosophical Society*, *89*(2), 284–301. https://doi.org/10.1111/brv.12053

Devine, R. T., & Hughes, C. (2013). Silent Films and Strange Stories: Theory of Mind, Gender, and Social Experiences in Middle Childhood. *Child Development*, *84*(3), 989–1003. https://doi.org/10.1111/cdev.12017

Dickerson, K., Gerhardstein, P., Zack, E., & Barr, R. (2008). Age-Related Changes in Learning Across Early Childhood : A New Imitation Task. https://doi.org/10.1002/dev.21068

DiYanni, C., & Kelemen, D. (2008). Using a bad tool with good intention: Young children's imitation of adults' questionable choices. *Journal of Experimental Child Psychology*, *101*(4), 241–261. https://doi.org/10.1016/j.jecp.2008.05.002

Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, *13*(2), 331–338. https://doi.org/10.1111/j.1467-7687.2009.00888.x

Eisenberg, N., & Strayer, J. (1990). *Empathy and its development*. Cambridge University Press. Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=PVQ4AAAAIAAJ&oi=fnd&pg=PA317&dq=mimicry+eye+contact&ots=Kl5T8ulmBz&sig=9KbrpohxoxELPOgFgUa98LGKPPE#v=onepage&q=mimicry eye contact&f=false

Elsner, B., Hauf, P., & Aschersleben, G. (2007). Imitating step by step: A detailed analysis of 9- to 15-month-olds' reproduction of a three-step action sequence. *Infant Behavior and Development*, *30*(2), 325–335. https://doi.org/10.1016/j.infbeh.2006.10.001

Esseily, R., Nadel, J., & Fagard, J. (2010). Object retrieval through observational learning in 8- to 18-month-old infants. *Infant Behavior and Development*, *33*(4), 695–699. https://doi.org/10.1016/j.infbeh.2010.07.017

Feng, G. (2011). Eye Tracking: A Brief Guide for Developmental Researchers. *Journal of Cognition and Development*, *12*(February 2015), 1–11. https://doi.org/10.1080/15248372.2011.547447

Flynn, E., & Smith, K. (2012). Investigating the Mechanisms of Cultural Acquisition How Pervasive is Overimitation in Adults ?, *43*(4), 185–195. https://doi.org/10.1027/1864-9335/a000119

Flynn, E., & Whiten, A. (2008a). Cultural transmission of tool use in young children: A diffusion chain study. *Social Development*, *17*(3), 699–718. https://doi.org/10.1111/j.1467-9507.2007.00453.x

Flynn, E., & Whiten, A. (2008b). 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. https://doi.org/10.1016/j.jecp.2008.05.009

Fogarty, L., Rendell, L., & Laland, K. N. (2012). Mental time travel, memory and the social learning strategies tournament. *Learning and Motivation*, *43*(4), 241–246. https://doi.org/10.1016/j.lmot.2012.05.009

Frith, C. D., & Frith, U. (2012). Mechanisms of Social Cognition. *Annual Review of Psychology*, *63*(1), 287–313. https://doi.org/10.1146/annurev-psych-120710-100449

Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology*, *15*(17), R644–R645. https://doi.org/10.1016/j.cub.2005.08.041

Gardiner, A. K., Bjorklund, D. F., Greif, M. L., & Gray, S. K. (2012). Choosing and using tools: Prior experience and task difficulty influence preschoolers' tool-use strategies. *Cognitive Development*, *27*(3), 240–254. https://doi.org/10.1016/j.cogdev.2012.05.001

Gardiner, A. K., Greif, M. L., & Bjorklund, D. F. (2011). Guided by Intention: Preschoolers' Imitation Reflects Inferences of Causation. *Journal of Cognition and Development*, *12*(3), 355–373. https://doi.org/10.1080/15248372.2010.542216

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 to 15 Years of Age. *Developmental Psychology*, *40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Gergely, G. (2006). Sylvia ' s recipe : The role of imitation and pedagogy in the transmission of cultural knowledge, 1–15.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: the naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292. https://doi.org/10.1016/S1364-6613(03)00128-1

Gergely, G., & Csibra, G. (2006). Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge Social transmission of behaviors in non-human primates "Simple" (goal-driven) teleology and tool use in primates. *Byrne and Russon Boesch and Boesch Goodall McGrew*, 229–255. Retrieved from https://pdfs.semanticscholar.org/aadb/968f7a109d3f8d566bd5a4e300de414ed350.pdf

Ghetti, S., & Bunge, S. A. (2012). Neural changes underlying the development of episodic memory during middle childhood. *Developmental Cognitive Neuroscience*, *2*(4), 381–395. https://doi.org/10.1016/j.dcn.2012.05.002

Ghetti, S., DeMaster, D. M., Yonelinas, A. P., & Bunge, S. A. (2010). Developmental differences in medial temporal lobe function during memory encoding. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(28), 9548–56. https://doi.org/10.1523/JNEUROSCI.3500-09.2010

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., … Dahl, R. E. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, 201700811. https://doi.org/10.1073/pnas.1700811114

Gowen, E. (2012). Imitation in autism: why action kinematics matter. *Frontiers in Integrative Neuroscience*, *6*, 117. https://doi.org/10.3389/fnint.2012.00117

Gredebäck, G., Johnson, S., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology*, *35*(February 2015), 1–19. https://doi.org/10.1080/87565640903325758

Gruber, T., Deschenaux, A., Frick, A., & Clément, F. (2017). Group Membership Influences More Social Identification Than Social Learning or Overimitation in Children. *Child Development*. https://doi.org/10.1111/cdev.12931

Hagen, J. W., & Hale, G. A. (1973). THE DEVELOPMENT OF ATTENTION IN CHILDREN1. *ETS Research Bulletin Series*, *1973*(1), i-37. https://doi.org/10.1002/j.2333-8504.1973.tb00453.x

Haun, D. B. M., & Tomasello, M. (2011). Conformity to Peer Pressure in Preschool Children. *Child Development*, *82*(6), 1759–1767. https://doi.org/10.1111/j.1467-8624.2011.01666.x

Hayne, H., Herbert, J., & Simcock, G. (2003). Imitation from television by 24- and 30-month-olds. *Developmental Science*, *6*(3), 254–261. https://doi.org/10.1111/1467-7687.00281

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

Herrmann, P. A., Legare, C. H., Harris, P. L., & Whitehouse, H. (2013). Stick to the script: The effect of witnessing multiple actors on children's imitation. *Cognition*, *129*(3), 536–543. https://doi.org/10.1016/j.cognition.2013.08.010

Hess, U., Philippot, P., & Blairy, S. (1999). Mimicry: Facts and Fiction, 213–241.

Retrieved from https://www.researchgate.net/profile/Ursula_Hess/publication/246829273_Mimicry_fact_and_fiction/links/5847d82b08ae61f75de28337.pdf

Hetherington, E. M., & Mavis, E. (1965). A developmental study of the effects of sex of the dominant parent on sex-role preference, identification, and imitation in children. *Journal of Personality and Social Psychology*, *2*(2), 188–194. https://doi.org/10.1037/h0022374

Heyes, C., & Cecilia. (2011). Automatic imitation. *Psychological Bulletin*, *137*(3), 463–483. https://doi.org/10.1037/a0022288

Hilbrink, E. E., Sakkalou, E., Ellis-Davies, K., Fowler, N. C., & Gattis, M. (2013). Selective and faithful imitation at 12 and 15?months. *Developmental Science*, *16*(6), n/a-n/a. https://doi.org/10.1111/desc.12070

Hitch, G. J., Woodin, M. E., & Baker, S. (1989). Visual and phonological components of working memory in children. *Memory & Cognition*, *17*(2), 175–185. https://doi.org/10.3758/BF03197067

Hobson, J. A., & Hobson, R. P. (n.d.). Identification: The missing link between joint attention and imitation? Retrieved from http://search.proquest.com/openview/4888550b5e97a8a864fc77dd116cc2a8/1?pq-origsite=gscholar&cbl=30453

Hoehl, S., Zettersten, M., Schleihauf, H., Grätz, S., & Pauen, S. (2014a). The role of social interaction and pedagogical cues for eliciting and reducing overimitation in preschoolers. *Journal of Experimental Child Psychology*, *122*, 122–133. https://doi.org/10.1016/j.jecp.2013.12.012

Hoehl, S., Zettersten, M., Schleihauf, H., Grätz, S., & Pauen, S. (2014b). The role of social interaction and pedagogical cues for eliciting and reducing overimitation in preschoolers. *Journal of Experimental Child Psychology*, *122*, 122–133. https://doi.org/10.1016/j.jecp.2013.12.012

Hopper, L. M., Lambeth, S. P., Schapiro, S. J., & Whiten, A. (2008). Observational learning in chimpanzees and children studied through “ghost” conditions. *Proceedings of the Royal Society of London B: Biological Sciences*, *275*(1636). Retrieved from http://rspb.royalsocietypublishing.org/content/275/1636/835.short

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

Horowitz, A. C., Whiten, A., Custance, D. M., Gomez, J.-C., Teixidor, P., & Bard, K. A. (n.d.). Do Humans Ape? Or Do Apes Human? Imitation and Intention in Humans (Homo sapiens) and Other Animals. https://doi.org/10.1037/0735-7036.117.3.325

Howard, L. H., Henderson, A. M. E., Carrazza, C., & Woodward, A. L. (2015). Infants' and Young Children's Imitation of Linguistic In-Group and Out-Group Informants. *Child Development*, *86*(1), 259–275. https://doi.org/10.1111/cdev.12299

Huang, C.-T., & Charman, T. (2005). Gradations of emulation learning in infants' imitation of actions on objects. *Journal of Experimental Child Psychology*, *92*(3), 276–302. https://doi.org/10.1016/j.jecp.2005.06.003

Huber, L., Range, F., Voelkl, B., Szucsich, A., Virányi, Z., & Miklosi, A. (2009). The evolution of imitation: what do the capacities of non-human animals tell us about the mechanisms of imitation? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1528), 2299–309. https://doi.org/10.1098/rstb.2009.0060

Isaacs, E. B., & Vargha-Khadem, F. (1989). Differential course of development of spatial and verbal memory span: A normative study. *British Journal of Developmental Psychology*, *7*(4), 377–380. https://doi.org/10.1111/j.2044-835X.1989.tb00814.x

Isquith, P. K., Gioia, G. A., & Espy, K. A. (2004). Executive Function in Preschool Children: Examination Through Everyday Behavior. *Developmental Neuropsychology*, *26*(1), 403–422. https://doi.org/10.1207/s15326942dn2601_3

Jones, S. S. (2009). The development of imitation in infancy. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1528). Retrieved from http://rstb.royalsocietypublishing.org/content/364/1528/2325.short

Kail, R. (1993). Processing Time Decreases Globally at an Exponential Rate during Childhood and Adolescence. *Journal of Experimental Child Psychology*, *56*(2), 254–265. https://doi.org/10.1006/jecp.1993.1034

Kenward, B. (2012). Over-imitating preschoolers believe unnecessary actions are normative and enforce their performance by a third party. *Journal of Experimental Child Psychology*, *112*(2), 195–207. https://doi.org/10.1016/j.jecp.2012.02.006

Kenward, B., Karlsson, M., & Persson, J. (2011). Over-imitation is better explained by norm learning than by distorted causal learning. *Proceedings of the Royal Society of London B: Biological Sciences*, *278*(1709). Retrieved from http://rspb.royalsocietypublishing.org/content/278/1709/1239.short

Kertzer, D. I. (1988). *Ritual, politics, and power*. Yale University Press. Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=osntCYbeXRYC&oi=fnd&pg=PP17&dq=ritual+defined&ots=dLcFs6lbTp&sig=-Nq9BXR1m8VyGCHfDL1vFvQGq-Q#v=onepage&q=ritual defined&f=false

Keupp, S., Bancken, C., Schillmöller, J., Rakoczy, H., & Behne, T. (2016a). Rational over-imitation: Preschoolers consider material costs and copy causally irrelevant actions selectively. *Cognition*. https://doi.org/10.1016/j.cognition.2015.11.007

Keupp, S., Bancken, C., Schillmöller, J., Rakoczy, H., & Behne, T. (2016b). Rational over-imitation: Preschoolers consider material costs and copy causally irrelevant actions selectively. *Cognition*, *147*, 85–92. https://doi.org/10.1016/j.cognition.2015.11.007

Keupp, S., Behne, T., & Rakoczy, H. (2013). Why do children overimitate? Normativity is crucial. *Journal of Experimental Child Psychology*, *116*(2), 392–406. https://doi.org/10.1016/j.jecp.2013.07.002

Keupp, S., Behne, T., Zachow, J., Kasbohm, A., & Rakoczy, H. (2015a). Journal of Experimental Child Over-imitation is not automatic : Context sensitivity in children ' s overimitation and action interpretation of causally irrelevant actions. *Journal of Experimental Child Psychology*, *130*, 163–175. https://doi.org/10.1016/j.jecp.2014.10.005

Keupp, S., Behne, T., Zachow, J., Kasbohm, A., & Rakoczy, H. (2015b). Over-imitation is not automatic: Context sensitivity in children’s overimitation and action interpretation of causally irrelevant actions. *Journal of Experimental Child Psychology*, *130*, 163–175. https://doi.org/10.1016/j.jecp.2014.10.005

Krohn, M. D., Skinner, W. F., Massey, J. L., & Akers, R. L. (1985). Social Learning Theory and Adolescent Cigarette Smoking: A Longitudinal Study. *Social Problems*, *32*(5), 455–473. https://doi.org/10.2307/800775

Kupán, K., Király, I., Kupán, K., Krekó, K., Miklósi, Á., & Topál, J. (2017). *Interacting effect of two social factors on 18-month-old infants' imitative behavior: Communicative cues and demonstrator presence*. *Journal of Experimental Child Psychology* (Vol. 161). https://doi.org/10.1016/j.jecp.2017.03.019

Lane, D. M., & Pearson, D. A. (1982). The development of selective attention. *Merrill-Palmer Quarterly*. Retrieved from http://psycnet.apa.org/psycinfo/1983-00792-001

Laor, Y. (2013). Cultural uniformity and religion. *Religion, Brain & Behavior*, *3*(3), 233–253. https://doi.org/10.1080/2153599X.2012.739409

Lashbrook, J. T. (2000). Fitting in: exploring the emotional dimension of adolescent peer pressure. *Adolescence*, *35*(140), 747–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11214212

Legare, C. H., & Nielsen, M. (2015a). Imitation and Innovation: The Dual Engines of Cultural Learning. *Trends in Cognitive Sciences*, *19*(11), 688–699. https://doi.org/10.1016/j.tics.2015.08.005

Legare, C. H., & Nielsen, M. (2015b). Imitation and Innovation: The Dual Engines of Cultural Learning. *Trends in Cognitive Sciences*, *19*(11), 688–699. https://doi.org/10.1016/j.tics.2015.08.005

Legare, C. H., Wen, N. J., Herrmann, P. A., & Whitehouse, H. (2015). Imitative flexibility and the development of cultural learning. *Cognition*, *142*, 351–361. https://doi.org/10.1016/j.cognition.2015.05.020

Lewis, H. M., & Laland, K. N. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2171–2180. https://doi.org/10.1098/rstb.2012.0119

Logan, C. J., Breen, A. J., Taylor, A. H., Gray, R. D., & Hoppitt, W. J. E. (2016). How

New Caledonian crows solve novel foraging problems and what it means for cumulative culture. *Learning & Behavior*, *44*(1), 18–28. https://doi.org/10.3758/s13420-015-0194-x

Longo, M. R., & Bertenthal, B. I. (2006). Common Coding of Observation and Execution of Action in 9-Month-Old Infants. *Infancy*, *10*(1), 43–59. https://doi.org/10.1207/s15327078in1001_3

LOUCKS, J., & MELTZOFF, A. N. (2013). Goals influence memory and imitation for dynamic human action in 36-month-old children. *Scandinavian Journal of Psychology*, *54*(1), 41–50. https://doi.org/10.1111/sjop.12004

Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The Development of Nonverbal Working Memory and Executive Control Processes in Adolescents. *Child Development*, *76*(3), 697–712. https://doi.org/10.1111/j.1467-8624.2005.00872.x

Lyons, D. E., Damrosch, D. H., Lin, J. K., Macris, D. M., & Keil, F. C. (2011). The scope and limits of overimitation in the transmission of artefact culture. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *366*(1567). Retrieved from http://rstb.royalsocietypublishing.org/content/366/1567/1158.short

Lyons, D. E., Damrosch, D. H., Lin, J. K., Macris, D. M., & Keil, F. C. (2011). The scope and limits of overimitation in the transmission of artefact culture. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*, 1158–1167. https://doi.org/10.1098/rstb.2010.0335

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

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

Marsh, L. E., Ropar, D., Hamilton, A. F. de C., Erdohegyi, A., & Csibra, G. (2014). The Social Modulation of Imitation Fidelity in School-Age Children. *PLoS ONE*, *9*(1), e86127. https://doi.org/10.1371/journal.pone.0086127

Marsh, L., Pearson, A., Ropar, D., & Hamilton, A. (n.d.). Children with autism do not overimitate. Retrieved from http://www.antoniahamilton.com/MarshPearsonRoparHamilton_CB_postprint.pdf

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

Mataric, M. J., & Pomplun, M. (1998). Fixation behavior in observation and imitation of human movement, 191–202.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., & Stallings, M. C. (2013). Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Childhood Research Quarterly*, *28*(2), 314–324. https://doi.org/10.1016/j.ecresq.2012.07.008

McGuigan, N. (2012). The role of transmission biases in the cultural diffusion of irrelevant actions. *Journal of Comparative Psychology*, *126*(2), 150–160. https://doi.org/10.1037/a0025525

McGuigan, N. (2013). The influence of model status on the tendency of young children to over-imitate. *Journal of Experimental Child Psychology*, *116*(4), 962–969. https://doi.org/10.1016/j.jecp.2013.05.004

McGuigan, N., & Burgess, V. (2017). Is the tendency to conform influenced by the age of the majority? *Journal of Experimental Child Psychology*, *157*, 49–65. https://doi.org/10.1016/j.jecp.2016.12.007

McGuigan, N., Gladstone, D., Cook, L., Macris, D., & Keil, F. (2012a). Is the Cultural Transmission of Irrelevant Tool Actions in Adult Humans (Homo Sapiens) Best Explained as the Result of an Evolved Conformist Bias? *PLoS ONE*, *7*(12), e50863. https://doi.org/10.1371/journal.pone.0050863

McGuigan, N., Gladstone, D., Cook, L., Macris, D., & Keil, F. (2012b). Is the Cultural

Transmission of Irrelevant Tool Actions in Adult Humans (Homo Sapiens) Best Explained as the Result of an Evolved Conformist Bias? *PLoS ONE*, *7*(12), e50863. https://doi.org/10.1371/journal.pone.0050863

McGuigan, N., & Graham, M. (2010). Cultural transmission of irrelevant tool actions in diffusion chains of 3- and 5-year-old children. *European Journal of Developmental Psychology*, *7*(5), 561–577. https://doi.org/10.1080/17405620902858125

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

McGuigan, N., Makinson, J., & Whiten, A. (2011b). From over-imitation to super-copying: adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal of Psychology (London, England : 1953)*, *102*(1), 1–18. https://doi.org/10.1348/000712610X493115

McGuigan, N., & Robertson, S. (2015). The influence of peers on the tendency of 3- and 4-year-old children to over-imitate. *Journal of Experimental Child Psychology*, *136*, 42–54. https://doi.org/10.1016/j.jecp.2015.03.004

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

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

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

McIntosh, D. N., Reichmann-Decker, A., Winkielman, P., & Wilbarger, J. L. (2006).

When the social mirror breaks: deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Developmental Science*, *9*(3), 295–302. https://doi.org/10.1111/j.1467-7687.2006.00492.x

Meltzoff, A. N. (n.d.). Imitation as a Mechanism of Social Cognition: Origins of Empathy, Theory of Mind, and the Representation of Action. In *Blackwell Handbook of Childhood Cognitive Development* (pp. 6–25). Malden, MA, USA: Blackwell Publishers Ltd. https://doi.org/10.1002/9780470996652.ch1

Meltzoff, A. N. (1985). Immediate and Deferred Imitation in Fourteen- and Twenty-Four-Month-Old Infants. *Child Development*, *56*(1), 62. https://doi.org/10.2307/1130174

Meltzoff, A. N., & Moore, M. K. (1989). Imitation in Newborn Infants: Exploring the Range of Gestures Imitated and the Underlying Mechanisms. *Abravanel & Sigafoos*, *25*(6). Retrieved from http://ilabs.washington.edu/meltzoff/pdf/89Meltzoff_Moore_DevPsy.pdf

Morgan, T. J. H., Laland, K. N., & Harris, P. L. (2015). The development of adaptive conformity in young children: effects of uncertainty and consensus. *Developmental Science*, *18*(4), 511–524. https://doi.org/10.1111/desc.12231

Moriguchi, Y. (2014). The early development of executive function and its relation to social interaction: a brief review. *Frontiers in Psychology*, *5*, 388. https://doi.org/10.3389/fpsyg.2014.00388

Mullally, S. L., & Maguire, E. a. (2014). Learning to remember: The early ontogeny of episodic memory. *Developmental Cognitive Neuroscience*, *9*, 12–29. https://doi.org/10.1016/j.dcn.2013.12.006

Nakano, T., & Kitazawa, S. (2017). Development of long-term event memory in preverbal infants: an eye-tracking study. *Scientific Reports*, *7*, 44086. https://doi.org/10.1038/srep44086

Nielsen, M. (2006). Copying actions and copying outcomes: social learning through the second year. *Developmental Psychology*, *42*(3), 555–65. https://doi.org/10.1037/0012-1649.42.3.555

Nielsen, M. (2013). Young Children’s Imitative and Innovative Behaviour on the

Floating Object Task. *Infant and Child Development*, *22*(1), 44–52. https://doi.org/10.1002/icd.1765

Nielsen, M., & Blank, C. (2011). Imitation in young children: when who gets copied is more important than what gets copied. *Developmental Psychology*, *47*(4), 1050–3. https://doi.org/10.1037/a0023866

Nielsen, M., Cucchiaro, J., & Mohamedally, J. (2012). When the Transmission of Culture Is Child's Play. *PLoS ONE*, *7*(3), e34066. https://doi.org/10.1371/journal.pone.0034066

Nielsen, M., & Hudry, K. (2010). Over-imitation in children with autism and Down syndrome. *Australian Journal of Psychology*, *62*(2), 67–74. https://doi.org/10.1080/00049530902758613

Nielsen, M., Kapitány, R., & Elkins, R. (2014). Evolution and Human Behavior The perpetuation of ritualistic actions as revealed by young children â€™ s transmission of normative behavior. *Evolution and Human Behavior*. https://doi.org/10.1016/j.evolhumbehav.2014.11.002

Nielsen, M., Moore, C., & Mohamedally, J. (2012). Young children overimitate in third-party contexts. *Journal of Experimental Child Psychology*, *112*(1), 73–83. https://doi.org/10.1016/j.jecp.2012.01.001

Nielsen, M., Simcock, G., & Jenkins, L. (2008). The effect of social engagement on 24-month-olds' imitation from live and televised models. *Developmental Science*, *11*(5), 722–731. https://doi.org/10.1111/j.1467-7687.2008.00722.x

Nielsen, M., & Tomaselli, K. (2010). Overimitation in Kalahari Bushman children and the origins of human cultural cognition. *Psychological Science : A Journal of the American Psychological Society / APS*, *21*(5), 729–736. https://doi.org/10.1177/0956797610368808

Nielsen, M., Tomaselli, K., Mushin, I., & Whiten, A. (2014). Exploring tool innovation: A comparison of Western and Bushman children. *Journal of Experimental Child Psychology*, *126*, 384–394. https://doi.org/10.1016/j.jecp.2014.05.008

Odean, R., Nazareth, A., & Pruden, S. M. (2015). Novel methodology to examine cognitive and experiential factors in language development: combining eye-tracking and LENA technology. *Frontiers in Psychology*, *6*, 1266. https://doi.org/10.3389/fpsyg.2015.01266

Óturai, G., Kolling, T., & Knopf, M. (2013). Relations between 18-month-olds' gaze pattern and target action performance: a deferred imitation study with eye tracking. *Infant Behavior & Development*, *36*(4), 736–48. https://doi.org/10.1016/j.infbeh.2013.08.005

Over, H., & Carpenter, M. (2012). Putting the social into social learning: Explaining both selectivity and fidelity in children's copying behavior. *Journal of Comparative Psychology*, *126*(2), 182–192. https://doi.org/10.1037/a0024555

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011). Imitation in infancy: Rational or motor resonance? *Child Development*, *82*(4), 1047–1057. https://doi.org/10.1111/j.1467-8624.2011.01610.x

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, *3*(9), 337–344. https://doi.org/10.1016/S1364-6613(99)01362-5

Perner, J., & Ruffman, T. (1995). Episodic Memory and Autonoetic Conciousness: Developmental Evidence and a Theory of Childhood Amnesia. *Journal of Experimental Child Psychology*, *59*(3), 516–548. https://doi.org/10.1006/jecp.1995.1024

Pfeifer, C., & Elsner, B. (2013). Preschoolers' encoding of rational actions: The role of task features and verbal information. *Journal of Experimental Child Psychology*, *116*(2), 532–544. https://doi.org/10.1016/j.jecp.2012.12.005

Plude, D. J., & Doussard-Roosevelt, J. A. (1989). Aging, selective attention, and feature integration. *Psychology and Aging*, *4*(1), 98–105. https://doi.org/10.1037/0882-7974.4.1.98

Postman, L., & Phillips, L. W. (1965). Short-term Temporal Changes in Free Recall. *Quarterly Journal of Experimental Psychology*, *17*(2), 132–138. https://doi.org/10.1080/17470216508416422

Rakoczy, H., & Schmidt, M. F. H. (2013). The Early Ontogeny of Social Norms. *Child Development Perspectives*, *7*(1), 17–21. https://doi.org/10.1111/cdep.12010

Reindl, E., Apperly, I. A., Beck, S. R., & Tennie, C. (2017). Young children copy cumulative technological design in the absence of action information. *Scientific Reports*, *7*(1), 1788. https://doi.org/10.1038/s41598-017-01715-2

Rigamonti, M. M., Custance, D. M., Previde, E. P., & Spiezio, C. (n.d.). Testing for Localized Stimulus Enhancement and Object Movement Reenactment in Pig-Tailed Macaques (Macaca nemestrina) and Young Children (Homo sapiens). https://doi.org/10.1037/0735-7036.119.3.257

Roderer, T., & Roebers, C. M. (2010). Explicit and implicit confidence judgments and developmental differences in metamemory: An eye-tracking approach. *Metacognition and Learning*, *5*(3), 229–250. https://doi.org/10.1007/s11409-010-9059-z

Roebers, C. M., Schmid, C., & Roderer, T. (2010). Encoding Strategies in Primary School Children: Insights From an Eye-Tracking Approach and the Role of Individual Differences in Attentional Control. *The Journal of Genetic Psychology*, *171*(1), 1–21. https://doi.org/10.1080/00221320903300361

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to Theory of Mind. *Cognition*, *122*(2), 228–235. https://doi.org/10.1016/j.cognition.2011.11.004

Sato, W., & Yoshikawa, S. (2007). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition*, *104*(1), 1–18. https://doi.org/10.1016/j.cognition.2006.05.001

Scarf, D., Gross, J., Colombo, M., & Hayne, H. (2013). To have and to hold: Episodic memory in 3- and 4-year-old children. *Developmental Psychobiology*, *55*(2), 125–132. https://doi.org/10.1002/dev.21004

Schachner, A., & Carey, S. (2013a). Reasoning about “irrational” actions: When intentional movements cannot be explained, the movements themselves are seen as the goal. *Cognition*, *129*(2), 309–327. https://doi.org/10.1016/j.cognition.2013.07.006

Schachner, A., & Carey, S. (2013b). Reasoning about “irrational” actions: When intentional movements cannot be explained, the movements themselves are seen as the goal. *Cognition*, *129*(2), 309–327.

https://doi.org/10.1016/j.cognition.2013.07.006

Schwanenflugel, P. J., Henderson, R. L., & Fabricius, W. V. (1998). Developing organization of mental verbs and theory of mind in middle childhood: Evidence from extensions. *Developmental Psychology*, *34*(3), 512–524. https://doi.org/10.1037/0012-1649.34.3.512

Sevlever, M., & Gillis, J. M. (2010). An examination of the state of imitation research in children with autism: Issues of definition and methodology. *Research in Developmental Disabilities*, *31*(5), 976–984. https://doi.org/10.1016/j.ridd.2010.04.014

Seyfarth, R. M., & Cheney, D. L. (2013). Affiliation, empathy, and the origins of theory of mind. *Proceedings of the National Academy of Sciences of the United States of America*, (Supplement 2), 10349–56. https://doi.org/10.1073/pnas.1301223110

Sharon, T., & DeLoache, J. S. (2003). The role of perseveration in children's symbolic understanding and skill. *Developmental Science*, *6*(3), 289–296. https://doi.org/10.1111/1467-7687.00285

Sheridan, K. M., Konopasky, A. W., Kirkwood, S., & Defeyter, M. A. (2016). The effects of environment and ownership on children's innovation of tools and tool material selection. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1690), 20150191. https://doi.org/10.1098/rstb.2015.0191

Shipton, C., & Nielsen, M. (2015). Before Cumulative Culture. *Human Nature*, *26*(3), 331–345. https://doi.org/10.1007/s12110-015-9233-8

Siegel, L. S., & Ryan, E. B. (1989). The Development of Working Memory in Normally Achieving and Subtypes of Learning Disabled Children. *Child Development*, *60*(4), 973. https://doi.org/10.2307/1131037

Simpson, A., & Riggs, K. J. (2011a). Three- and 4-year-olds encode modeled actions in two ways leading to immediate imitation and delayed emulation. *Developmental Psychology*, *47*(3), 834–840. https://doi.org/10.1037/a0023270

Simpson, A., & Riggs, K. J. (2011b). Three- and 4-year-olds encode modeled actions in two ways leading to immediate imitation and delayed emulation. *Developmental Psychology*, *47*(3), 834–840. https://doi.org/10.1037/a0023270

Simpson, A., & Riggs, K. J. (2011c). Three- and 4-year-olds encode modeled actions in two ways leading to immediate imitation and delayed emulation. *Developmental Psychology*, *47*(3), 834–840. https://doi.org/10.1037/a0023270

Southgate, V., Chevallier, C., & Csibra, G. (2009). Sensitivity to communicative relevance tells young children what to imitate. *Developmental Science*, *12*(6), 1013–1019. https://doi.org/10.1111/j.1467-7687.2009.00861.x

Spengler, S., Bird, G., & Brass, M. (2010). Hyperimitation of Actions Is Related to Reduced Understanding of Others' Minds in Autism Spectrum Conditions. *Biological Psychiatry*, *68*(12), 1148–1155. https://doi.org/10.1016/j.biopsych.2010.09.017

Strouse, G. a., & Troseth, G. L. (2008). "Don't try this at home": Toddlers' imitation of new skills from people on video. *Journal of Experimental Child Psychology*, *101*(4), 262–280. https://doi.org/10.1016/j.jecp.2008.05.010

Subiaul, F., & Schilder, B. (2014). *Working memory constraints on imitation and emulation*. *Journal of Experimental Child Psychology* (Vol. 128). https://doi.org/10.1016/j.jecp.2014.07.005

Subiaul, F., Winters, K., Krumpak, K., & Core, C. (2016). Vocal overimitation in preschool-age children. *Journal of Experimental Child Psychology*, *141*, 145–160. https://doi.org/10.1016/j.jecp.2015.08.010

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

Taylor, G., & Herbert, J. S. (2013). Infant and adult visual attention during an imitation demonstration. *Developmental Psychobiology*, *56*(4), 770–82. https://doi.org/10.1002/dev.21147

Taylor, G., & Herbert, J. S. (2014). Infant and adult visual attention during an imitation demonstration. *Developmental Psychobiology*, *56*(4), 770–782.

https://doi.org/10.1002/dev.21147

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

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2405–2415. https://doi.org/10.1098/rstb.2009.0052

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. Series B, Biological Sciences*, *364*, 2405–2415. https://doi.org/10.1098/rstb.2009.0052

Terrace, H. S., & Metcalfe, J. (2005). *The missing link in cognition : origins of self-reflective consciousness*. Oxford University Press. Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=hRdnDAAAQBAJ&oi=fnd&pg=PA3&dq=tulving+2005+episodic+memory&ots=BcT2VA7eTO&sig=XKEjNCMLyTlYKEH3GA709evj2KI#v=onepage&q&f=false

Thelen, Mark, H., Dollinger, Stephen, J., Kirkland, & Karen. (1979). Imitation and response uncertainty. *The Journal of Genetic Psychology*, *135*(1). Retrieved from http://search.proquest.com/docview/1297169971/fulltextPDF/88ACA121791340FEPQ/1?accountid=12753

Thompson, D. E., & Russell, J. (2004). The ghost condition: imitation versus emulation in young children's observational learning. *Developmental Psychology*, *40*(5), 882–889. https://doi.org/10.1037/0012-1649.40.5.882

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*(3), 495. https://doi.org/10.1017/S0140525X0003123X

Tomasello, M., Savage-Rumbaugh, S., & Kruger, A. C. (1993). Imitative Learning of Actions on Objects by Children, Chimpanzees, and Enculturated Chimpanzees. *Child Development*, *64*(6), 1688–1705. https://doi.org/10.1111/j.1467-8624.1993.tb04207.x

Troseth, G. L., Saylor, M. M., & Archer, A. H. (2006). Young Children ' s Use of Video as a Source of Socially Relevant Information, *77*(3), 786–799.

Tulving, E. (2002). E PISODIC M EMORY : From Mind to Brain. *Annual Review of Psychology*, *53*(1), 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Uniwersytet Warszawski, N., Setti, A., & Maurizzi, D. D. (2014). *Psychology of language and communication.* De Gruyter Open. Retrieved from https://cora.ucc.ie/handle/10468/2707

van Baaren, R., Janssen, L., Chartrand, T. L., & Dijksterhuis, A. (2009). Where is the love? The social aspects of mimicry. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1528). Retrieved from http://rstb.royalsocietypublishing.org/content/364/1528/2381.short

Vandierendonck, A., Kemps, E., Fastame, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology*, *95*(1), 57–79. https://doi.org/10.1348/000712604322779460

Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E. V., & Kliegel, M. (2013). Development of Affective Theory of Mind Across Adolescence: Disentangling the Role of Executive Functions. *Developmental Neuropsychology*, *38*(2), 114–125. https://doi.org/10.1080/87565641.2012.733786

Vetter, N. C., Leipold, K., Kliegel, M., Phillips, L. H., & Altgassen, M. (2013). Ongoing development of social cognition in adolescence. *Child Neuropsychology*, *19*(6), 615–629. https://doi.org/10.1080/09297049.2012.718324

Vivanti, G., Hocking, D. R., Fanning, P., & Dissanayake, C. (2017). The social nature of overimitation: Insights from Autism and Williams syndrome. *Cognition*, *161*, 10–18. https://doi.org/10.1016/j.cognition.2017.01.008

Vivanti, G., Nadig, A., Ozonoff, S., & Rogers, S. J. (2008). Journal of Experimental Child What do children with autism attend to during imitation tasks ? *Journal of Experimental Child Psychology*, *101*(3), 186–205. https://doi.org/10.1016/j.jecp.2008.04.008

Wang, Y., Newport, R., Hamilton, A. F. D. C., De, A. F., & Hamilton, C. (2011). Eye contact enhances mimicry of intransitive hand movements Eye contact enhances mimicry of intransitive hand movements, (April 2010), 10–14. https://doi.org/10.1098/rsbl.2010.0279

Wang, Y., Newport, R., & Hamilton, A. F. de C. (2010). Eye contact enhances mimicry of intransitive hand movements. *Biology Letters*. Retrieved from http://rsbl.royalsocietypublishing.org/content/early/2010/04/26/rsbl.2010.0279.short

Want, S. C., & Harris, P. L. (2001). Learning from Other People's Mistakes: Causal Understanding in Learning to Use a Tool. *Child Development*, *72*(2), 431–443. https://doi.org/10.1111/1467-8624.00288

Watson-Jones, R. E., Legare, C. H., Whitehouse, H., & Clegg, J. M. (2014). Task-specific effects of ostracism on imitative fidelity in early childhood ☆. https://doi.org/10.1016/j.evolhumbehav.2014.01.004

Wellman, H. M. (n.d.). Understanding the Psychological World: Developing a Theory of Mind. In *Blackwell Handbook of Childhood Cognitive Development* (pp. 167–187). Malden, MA, USA: Blackwell Publishers Ltd. https://doi.org/10.1002/9780470996652.ch8

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child Development*, *72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development*, *75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Wen, N. J., Herrmann, P. A., & Legare, C. H. (2016). Ritual increases children's affiliation with in-group members. *Evolution and Human Behavior*, *37*(1), 54–60. https://doi.org/10.1016/j.evolhumbehav.2015.08.002

Whitehouse, H., & Legare, C. H. (2013). The ritualistic child: imitation, affiliation, and the ritual stance in human development. Retrieved from https://ora.ox.ac.uk/objects/uuid:2b00ce6b-f281-4644-83fb-ef484701b5f6

Whiten, A., Allan, G., Devlin, S., Kseib, N., Raw, N., & McGuigan, N. (2016a). Social Learning in the Real-World: "Over-Imitation" Occurs in Both Children and Adults

Unaware of Participation in an Experiment and Independently of Social Interaction. *PLOS ONE*, *11*(7), e0159920. https://doi.org/10.1371/journal.pone.0159920

Whiten, A., Allan, G., Devlin, S., Kseib, N., Raw, N., & McGuigan, N. (2016b). Social Learning in the Real-World: “Over-Imitation” Occurs in Both Children and Adults Unaware of Participation in an Experiment and Independently of Social Interaction. *PLOS ONE*, *11*(7), e0159920. https://doi.org/10.1371/journal.pone.0159920

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

Wild, K. S., Poliakoff, E., Jerrison, A., & Gowen, E. (2012). Goal-Directed and Goal-Less Imitation in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *42*(8), 1739–1749. https://doi.org/10.1007/s10803-011-1417-4

Wilks, M., Collier-Baker, E., & Nielsen, M. (2015). Preschool children favor copying a successful individual over an unsuccessful group. *Developmental Science*, *18*(6), 1014–1024. https://doi.org/10.1111/desc.12274

Williams, J. H. G., Casey, J. M., Braadbaart, L., Culmer, P. R., & Mon-Williams, M. (2014). Kinematic Measures of Imitation Fidelity in Primary School Children. *Journal of Cognition and Development*, *15*(2), 345–362. https://doi.org/10.1080/15248372.2013.771265

Williamson, R. A., & Markman, E. M. (n.d.). Precision of Imitation as a Function of Preschoolers’ Understanding of the Goal of the Demonstration. https://doi.org/10.1037/0012-1649.42.4.723

Wimmer, H., & Perner, J. (1983a). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children’s understanding of deception. *Cognition*, *13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Wimmer, H., & Perner, J. (1983b). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2012). Context-dependent model-based biases in cultural transmission: children's imitation is affected by model age over model knowledge state. *Evolution and Human Behavior*, *33*(4), 387–394. https://doi.org/10.1016/j.evolhumbehav.2011.11.010

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Copy me or copy you ? The effect of prior experience on social learning q. *Cognition*, *127*(2), 203–213. https://doi.org/10.1016/j.cognition.2013.01.002

Yu, Y., & Kushnir, T. (2011). It's all about the game: Infants' action strategies during imitation are influenced by their prior expectations. *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, 1–5. Retrieved from papers3://publication/uuid/fcd4fc7f-b310-4f25-8bac-4fb869bff675

Yu, Y., & Kushnir, T. (2014). Social context effects in 2- and 4-year-olds' selective versus faithful imitation. *Developmental Psychology*, *50*, 922–33. https://doi.org/10.1037/a0034242

Zmyj, N., Daum, M. M., & Prinz, W. (2012). Fourteen-Month-Olds ' Imitation of Differently Aged Models, *266*(September 2011), 250–266. https://doi.org/10.1002/icd

Zmyj, N., & Seehagen, S. (2013). The Role of a Model ' s Age for Young Children ' s Imitation : A Research Review, *641*(August), 622–641. https://doi.org/10.1002/icd