

Navigating the Boundary Between the Actual and the Possible Through Reasoning,

Imitation and Language at Age Four

Katherine Sophia Scott

PhD Thesis

School of Psychology, Cardiff University

2015

## DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ..... (candidate)      Date  
.....

## STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed ..... (candidate)      Date  
.....

## STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ..... (candidate)      Date  
.....

## STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate)      Date  
.....

## STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans **after expiry of a bar**

**on access previously approved by the Academic Standards & Quality Committee.**

Signed ..... (candidate)

Date

.....

## Summary of Thesis

For children to become competent intentional agents they must constantly navigate between understanding how the world is, and how the world could be: they navigate between the actual and the possible. From their earliest experiences of the world and their interactions with those who inhabit it infants can extract information from and make inferences about the causal structure of both the physical and social worlds in which they reside. Seeing how an event in the world can change that world provides children with an opportunity to consider what else could have, and could be done to produce a similar change in the future. Imitating others allows children to replicate goal successes through the replication of some or all of another agent's behaviors. Imitation provides a short-cut to a possible world through standing on the shoulders of giants. In this thesis we investigate imitation and language through the lens of how well children aged four-years navigate these actual vs. possible worlds. In Chapters 2 through 4 we characterize our sample's Counterfactual Reasoning, Executive Functioning and Mental State Understanding. In Chapter 5 we utilize an experimental paradigm to assess the imitative behaviors of our sample and relate these to the earlier indices of cognitive development. We test the hypothesis that children who reason more successfully about counterfactual situations and are better able to infer behaviors based on mental states, are concurrently better equipped to imitate in order to achieve a given instrumental goal. We wanted to relate these cognitive reasoning skills not only to their proclivity for imitating different aspects of others' intentional acts but also their use of language as a tool. In Chapter 6 we investigated our sample's imitation of different aspects of language. Furthermore from our experimental paradigms we derived a measure of linguistic competence for use in Chapter 7 in an attempt to characterize our sample's use of language as a tool.

This thesis provides the first investigation of the relationship between children's reasoning about the actual versus the possible and their concurrent ability to move between the actual and the possible through their own physical or linguistic manipulations.

## Acknowledgements

This PhD research project was undertaken in, and funded by, the School of Psychology at Cardiff University. I would like to thank Professor Merideth Gattis for her supervision of this project and for all the hard work and dedication she puts in daily for the benefit of this, and other, projects undertaken within Development@Cardiff and beyond. Thank you to Professor Ulrike Hahn at Birkbeck, University of London, for her supervision, her endless encouragement and for always finding a way to make sure I knew where I was headed. To the families and little people who donated their time to this project, this thesis is for you.

To the School of Psychology, there are too many names to mention here, but to all the faces from Cognitive group, Friday talks and the halls of the Tower I feel very grateful to have shared this journey alongside you all. To Rob and Ed, thank you for listening. To my office friends, Georgie, Cass, Rachel, Mirko, Dan & Aimee you made the bad days better, often with bacon and/or beer. Scott and Katy, you did, do, and always will make me laugh, I love you both. To my friends and colleagues and Cardiff University Sport, thank you for letting me be a part of your world. Cardiff University Netball Club and Llandaff Athletic, thank you for every second. I will always be proud to wear black and red.

Nadine, Lynn, Gemma, Naomi, Fleur, Wendy and Danielle, thank you for being my friends and always being there when I needed you. To the Perkins' and Widleys, thank you for always giving me a home when I was stressing out. Mum, Dad and Mark, I couldn't have a better family and your unconditional love and belief makes anything possible. And finally to Tom, I love you for all the big and small things you did along the way, thank you for always walking beside me and holding my hand.

## Table of Contents

|  |    |
|--|----|
| Chapter 1: Navigating the Actual versus the Possible: An Introduction.....                                 | 13 |
| 1.1. Executive Functioning.....  | 14 |
| 1.2. Counterfactual Reasoning.....   | 15 |
| 1.3. Mental State Understanding.....   | 16 |
| 1.4. Imitation.....  | 17 |
| 1.5. Language.....   | 20 |
| 1.6. Research Aims .....   | 22 |
| Chapter 2: The Road to Reason: Conditional and Counterfactual Reasoning at Age<br>Four.....                | 26 |
| 2.1. Abstract.....   | 26 |
| 2.2. Introduction.....   | 27 |
| 2.2.1. Assessing the Development of Counterfactual Reasoning.....  | 36 |
| 2.2.2. The Relationship Between Counterfactual Conditional Reasoning and<br>Other Cognitive Abilities..... | 44 |
| 2.2.3. Aims.....   | 50 |
| 2.3. Method.....   | 54 |
| 2.3.1. Participants.....   | 54 |
| 2.3.2. Design.....   | 54 |
| 2.3.3. Materials.....  | 55 |
| 2.3.4. Procedure .....   | 57 |
| 2.4. Results.....  | 59 |
| 2.4.1. Slide Task.....   | 59 |
| 2.4.2. Story Task .....  | 61 |
| 2.4.3. Relationship Between Tasks.....   | 63 |
| 2.5. Discussion.....   | 66 |
| 2.5.1. Limitations.....  | 69 |
| 2.5.2. Conclusions.....  | 70 |
| Chapter 3: Executive Functions and the Selection of One Response Over All<br>Others.....                   | 72 |
| 3.1. Abstract.....   | 72 |
| 3.2. Introduction.....   | 73 |
| 3.2.1. What are Executive Functions? .....   | 74 |
| 3.2.2 Measuring EF and its Development in the Pre-School Years.....  | 79 |

|   |     |
|---|-----|
| 3.2.3. Executive Functions and Other Cognitive Abilities.....   | 84  |
| 3.2.4. Aims.....  | 86  |
| 3.3. Method.....  | 87  |
| 3.3.1. Participants.....  | 87  |
| 3.3.2. Measures.....  | 87  |
| 3.3.3. Procedure.....   | 87  |
| 3.3.4. Coding and Analysis.....   | 91  |
| 3.4. Results.....   | 92  |
| 3.5. Discussion.....  | 94  |
| 3.5.1. Limitations.....   | 98  |
| 3.5.2. Conclusions.....   | 98  |
| Chapter 4: Moving Between Your Actual and their Possible: Conditional Reasoning<br>About the Mind of Another.....                                 | 99  |
| 4.1. Abstract.....  | 99  |
| 4.2. Introduction.....  | 100 |
| 4.2.1. What is Mental State Understanding? .....  | 101 |
| 4.2.2. Developmental Investigations into MSU.....   | 103 |
| 4.2.3. The Relationship Between Young Children’s MSU and Other<br>Developing Abilities.....   | 108 |
| 4.2.4. Structural and Conceptual Similarities Between Executive Functioning<br>Tasks, Counterfactual Reasoning an Mental State Understanding..... | 110 |
| 4.2.5. Potential Benefits of Mental State Understanding for Children.....   | 114 |
| 4.2.6. Aims.....  | 115 |
| 4.3. Method.....  | 119 |
| 4.3.1. Participants.....  | 119 |
| 4.3.2. Materials.....   | 119 |
| 4.3.3. Procedure.....   | 120 |
| 4.4. Results.....   | 123 |
| 4.4.1. Mental State Understanding and Executive Functions.....  | 125 |
| 4.4.2. Mental State Understanding and Counterfactual Reasoning.....   | 128 |
| 4.4.3. Counterfactual Reasoning and Executive Functions .....   | 128 |
| 4.5. Discussion.....  | 131 |
| 4.5.1. The Measurement of MSU.....  | 132 |
| 4.5.2. The Measurement of Counterfactual Reasoning.....   | 134 |



|  |     |
|--|-----|
| 4.5.3. Limitations .....   | 135 |
| 4.5.4. Conclusions.....  | 135 |
| Chapter 5: Standing on the Shoulders of Giants: The Development of Imitation and its<br>Roots in Social Cognition.....     | 137 |
| 5.1. Abstract.....   | 137 |
| 5.2. Introduction.....   | 138 |
| 5.2.1. An Introduction to Imitation .....  | 138 |
| 5.2.2. The Beneficial Effects of Imitation.....  | 139 |
| 5.2.3. The Paradox of Faithful Imitation.....  | 142 |
| 5.2.4. The Centrality of Goals and Intentions to Imitation.....  | 150 |
| 5.2.5. Imitation’s Relationship to other Cognitive Skills.....   | 154 |
| 5.2.6. Aims.....   | 160 |
| 5.3. Method.....   | 167 |
| 5.3.1. Participants.....   | 167 |
| 5.3.2. Materials.....  | 167 |
| 5.3.3. Procedure.....  | 169 |
| 5.3.4. Data Analysis.....  | 170 |
| 5.4. Results.....  | 172 |
| 5.4.1. Faithful Imitation and Sensitivity to Causality.....  | 172 |
| 5.4.2. Comparing the Tower and Rake.....   | 176 |
| 5.4.5. Cross-task Relationships.....   | 181 |
| 5.5. Discussion.....   | 191 |
| 5.5.1. Object Imitation, Inhibitory Control and Cognitive Flexibility.....   | 196 |
| 5.5.2. Object Imitation and Counterfactual Conditional Reasoning.....  | 198 |
| 5.5.3. Object imitation and Mental State Understanding.....  | 201 |
| 5.5.4. Limitations.....  | 203 |
| 5.5.5. Conclusions.....  | 204 |
| Chapter 6: Imitation of Language and Linguistic Competence: Grammaticality, Novel<br>Lexical Items and Complex Syntax..... | 206 |
| 6.1. Abstract.....   | 206 |
| 6.2. General Introduction.....   | 207 |
| 6.2.1. Research Questions and Aims.....  | 211 |
| 6.3.1. Experiment 1: Introduction.....   | 214 |

|   |     |
|---|-----|
| 6.3.1.1. Aims.....  | 218 |
| 6.3.1.2. Methods.....   | 219 |
| 6.3.1.2.1. Participants.....  | 219 |
| 6.3.1.2.2. Materials.....   | 220 |
| 6.3.1.2.3. Procedure.....   | 220 |
| 6.3.1.2.4. Data Analysis.....   | 221 |
| 6.3.1.3. Results.....   | 221 |
| 6.3.1.4. Experiment 1: Discussion.....  | 224 |
| 6.3.2.1 Experiment 2: Introduction.....   | 227 |
| 6.3.2.1.1. Introduction to the Passive Verb Construction .....  | 228 |
| 6.3.2.1.2. Developmental Patterns with the English Passive Verb<br>Construction.....                            | 230 |
| 6.3.2.1.3. Abstract Schemas vs. Verb Islands: Assessing Productivity with<br>the Passive Experimentally.....    | 232 |
| 6.3.2.1.4. Imitation versus Innovation: How Language Sits Astride the<br>Actual and the Possible.....           | 238 |
| 6.3.2.1.5. What Facilitates the Imitation of Syntactic Structure? .....   | 243 |
| 6.3.2.1.6. Why Might the Imitation of Language Structures Itself be Useful<br>for Young Language Learners?..... | 245 |
| 6.3.2.1.7. Why Explore the use of Passive Verb Competence at Age 4?....   | 248 |
| 6.3.2.3. Method.....  | 251 |
| 6.3.2.3.1. Participants.....  | 251 |
| 6.3.2.3.2. Materials.....   | 252 |
| 6.3.2.3.3. Procedure.....   | 253 |
| 6.3.2.3.4. Data Analysis.....   | 255 |
| 6.3.2.4 Results.....  | 256 |
| 6.3.2.4.1. Novel verb utterances .....  | 256 |
| 6.3.2.4.2. English Utterances.....  | 263 |
| 6.3.2.4.3. Passive Verb Comprehension.....  | 268 |
| 6.3.2.5. Experiment 2: Discussion.....  | 269 |
| 6.3.3.1. Analysis 3: Introduction.....  | 276 |
| 6.3.3.1.1. Imitation of Actions and Speech.....   | 277 |
| 6.3.3.1.2. Aims .....   | 277 |
| 6.3.3.2. Results.....   | 278 |

|  |     |
|--|-----|
| 6.3.3.2.1. Language Imitation: Between Task Relationships.....   | 278 |
| 6.3.3.2.2. Imitation of Language and Actions: Cross Domain Relationships.....  | 282 |
| 6.3.3.3. Discussion.....   | 287 |
| 6.4. General Discussion.....   | 292 |
| 6.4.1. Limitations.....  | 296 |
| 6.4.2. Conclusions.....  | 298 |
| Chapter 7: The World as it Is, Was, Could Have Been and Could Be: The Relationship Between Reasoning About, and Manipulating the World Through Language..... |     |
| 7.1. Abstract.....   | 300 |
| 7.2. Introduction.....   | 301 |
| 7.3. Method.....   | 306 |
| 7.4. Results.....  | 313 |
| 7.4.1. Counterfactual Conditional Reasoning.....   | 313 |
| 7.4.2. Executive Functioning.....  | 315 |
| 7.4.3. Mental State Understanding.....   | 316 |
| 7.4.4. Object Imitation.....   | 317 |
| 7.5. Discussion.....   | 318 |
| 7.5.1. Counterfactual Conditional Reasoning and Language Competence.....   | 319 |
| 7.5.2. Executive Functioning and Language Competence.....  | 320 |
| 7.5.3. Mental State Understanding and Language Competence.....   | 321 |
| 7.4.5. Imitation and Language Competence.....  | 322 |
| 7.5.5. Limitations.....  | 322 |
| 7.5.6. Conclusions.....  | 325 |
| Chapter 8: Navigating the Boundary Between the Actual and the Possible at Age Four: Conclusions.....   |     |
| 8.1. Summary of Results.....   | 327 |
| 8.2. General Discussion.....   | 333 |
| 8.2.1. Reasoning at Age Four.....  | 333 |
| 8.2.2. Reasoning Skill versus Action on the World.....   | 335 |
| 8.2.3. Imitation Across Actions and Utterances.....  | 336 |
| 8.2.4. Limitations.....  | 338 |
| 8.2.5. Conclusions.....  | 341 |

|  |     |
|--|-----|
| 8.2.6. Future Directions.....  | 343 |
| References.....  | 345 |
| Appendix A: Counterfactual Reasoning Story Task.....                           | 382 |
| Appendix B: Location Change Counterfactual (Slide) Task.....                   | 387 |
| Appendix C: Excerpt from the NIH Toolbox Scoring and Interpretation Guide..... | 388 |
| Appendix D: Mental State Understanding Resources.....                          | 393 |
| Appendix E: Object Imitation Resources.....                                    | 397 |
| Appendix F: Verbal Imitation Sentences.....                                    | 398 |
| Appendix G: Verbal Imitation Apparatus.....                                    | 399 |

## **Chapter 1. Navigating the Actual and the Possible: An Introduction**

For children, becoming a competent intentional agent requires that they learn to navigate between understanding how the world was, is, and how it could be. Children are constantly required to navigate between the actual and the possible in their lives: it is this navigation between the actual and possible that this thesis is interested in. Children must incorporate information from the past, present and future when making their decisions. As humans, the ability to detach from a present reality and reason about past or future realities is arguably one of our central cognitive skills (Markovits & Barrouillet, 2002). In addition we learn to perceive the world in a less egotistical way and are able to think about situations from the perspective of others: we are able to reflect on their experiences and understandings of the world in addition to our own. With experience of the causal worlds in which we reside comes the increasingly complex ability to generate and reason about different possibilities and alternatives that are temporally or perceptually distinct from the current state of affairs (Diamond, 2013). That is we can model worlds that didn't or haven't happened, or worlds as experienced or observed by others, and subsequently reason about and make decisions based on these independently generated models. It is this ability to navigate the actual and the possible that provides us as humans with the machinery to support unrestrictive choices and behaviors.

Experiential cues such as conditional probability and temporal regularity are thought to contribute to young children's developing knowledge about the causal world (e.g., Frosch, McCormack, Lagnado & Burns, 2012; Tenenbaum, Kemp & Griffiths, 2006). With experience they begin to mentally move backward and forward in time, applying nascent causal knowledge to models of the world in order to make inferences and predictions about how they may have been or may be altered to

achieve a different outcome. These predictions may go on to provide their own additional causal information to further inform our future instrumental and social decision-making and behaviors. Flexible thinking allows for the assimilation of new causal information and symbiotically new causal information supports flexible thinking.

### **1.1. Executive Functioning**

The number and diversity of decisions we make throughout a lifetime are unfathomable: from the inconsequential to the life changing, the simple to the complex. Many behaviors that we engage in are based upon pre-defined associative, stimulus-response pairings, for example when we learn to stop at a red light and go on a green: if the light is red then we stop, if the light is green then we go. This constitutes a very basic form of conditional reasoning. Many tasks designed to assess an individual's executive functioning rely on the imposition of novel conditional rules in an attempt to capture the ease and speed with which we select one option over another. Using a novel task affords a glimpse at decision-making processes free from belief or prior knowledge constraints. Executive functions (EFs) themselves shall be discussed in more detail in Chapter 3, but briefly they refer to the domain general processing abilities required for goal directed thought and action (e.g. Carlson, 2005; Gioia, Isquith & Guy, 2001). In the context of this thesis investigating the ease with which children apply newly presented if/then causal statements is an important starting point from which to challenge their nascent actual versus possible navigational abilities. Once we can have a sense of children's domain general conditional reasoning ability we can then look to more specific forms of conditional reasoning that place additional demands on the reasoner. Learning to incorporate

temporal information, or social information into a decision for example increases its' complexity.

## **1.2. Counterfactual Reasoning**

Conditional reasoning provides a basis from which inferences can be drawn: *If* this happened/happens *then* that would have happened/will happen. Conditional reasoning can be applied to a diverse range of physical and social scenarios and relate to both physical and abstract causal relationships. However, basic conditional reasoning in isolation, such as stopping on red and going on green, is not the most demanding strategy we can adopt. It can be carried out free from context and with little or no regard to the actual set of circumstances being experienced (Johnson-Laird & Byrne, 2002). Counterfactual reasoning, however, is a more complex form of conditional reasoning (Perner & Rafetseder, 2011). Counterfactual conditional reasoning requires a model of the world being generated that is as close as possible to the actual and varies only as a function of those elements that are necessarily altered (e.g., Johnson-Laird & Byrne, 2002; Perner & Rafetseder, 2011). These models allow inferences to be made about novel situations that have not happened and as such can lead to innovative and flexible predictions and responses being produced. This ability is part of what gives humans their unique cognitive flexibility to navigate between and contrast the actual and the possible in such a way as to learn from the resulting counterfactual conclusions. This ability is truly fascinating as surprising outcomes can be reached that are neither true nor false in the actual world due to their counterfactual nature. Counterfactual outcomes, as we reflect on them in this thesis are those that were once possibilities, but never came to fruition. That is not to say, however, that the same causal counterfactual principles could not be used in the future in pursuit of the same or similar effects. Counterfactual reasoning, although temporally removed

from the actual, can inform the possible. This ability to move between how the world was and how the world could or could have been allows us to generate and select from almost infinite behavioral possibilities: flexible thinking allows for the uniquely human capability for flexible decision-making.

From their earliest social interactions infants can infer and extract information about the causal structure of both the physical and social worlds in which they reside. As motor development progresses and linguistic skills emerge children become increasingly able to make their own behavioral manipulations and interventions in the world, through actions on objects and vocalizations to others. Seeing how an event in the world causally impacts upon that world provides children with an opportunity to consider what else could have been, and could be done to produce a similar change in the future. Imitation is the replication of some or all of this modeled event sequence. Imitating the actions and speech of others allows children to achieve instrumental and social goals through the replication of some or all of another agent's behavior and negates the need for innovation in many situations: imitation provides a short-cut to a possible world through standing on the shoulders of giants.

### **1.3. Mental State Understanding**

Reasoning conditionally about the behavior of others, however, requires an enhanced set of skills when compared to reasoning about more concrete causal relations. Where cause and effect relationships such as motion or loss of integrity can be directly observed, the causal structure of human intentional action can largely pass undetected. A person may act in an intentional way based upon desire, belief, knowledge access or emotion, properties that may only be inferred not observed by another. Mental state understanding (MSU) skills in typically developing children emerge during the pre-school years and are discussed in more detail in Chapter 4.



When a child sees another person push a ball they can identify a cause/effect relationship directly. When a child sees another agent choose one of two equally appealing confectionary options they can infer a preference, but not directly observe it. Classic mental state understanding tasks such as the Sally-Anne task probe how well children reason about another's intentional behavior when their knowledge access is superior in the given experimental situation (Baron-Cohen, Leslie & Frith, 1985). Around age four typically developing children begin to reason in accordance with another's false belief rather than their own knowledge. They show an increasing amount of sophistication at identifying the boundaries between their own knowledge and how that affects their own actions, and how the knowledge available to others similarly affects theirs.

#### **1.4. Imitation**

During imitation this boundary, between ones own and others' causal knowledge and resultant intentions can become entwined. An individual's causal net or framework can be updated based on their own, and others interventions on the world. On the one hand children bring a wealth of their own causal knowledge and their own intentions to any situation. On the other hand, however, others may have more experience or expertise in a given situation and hence you may draw information from their choice of behaviors and perceived goals. During imitation tasks children are expected to infer the intentions of another agent and ultimately replicate their actions to the same instrumental end. Imitation, as opposed to mimicry, (discussed at length in Chapter 5) relies on this intentional understanding and as such often generates some initially paradoxical behavioral manifestations. Faithful or over-imitation refers to a situation in which one individual replicates both the goal of another's behavior, and crucially also the manner in which it was achieved. This is in

contrast to selective imitation where only those behaviors causally instrumental to goal success are replicated (again discussed in detail in Chapter 5). In many ecologically valid scenarios faithful-imitation is an eminently sensible strategy to adopt as individuals would be expected to, and will often chose the most efficient means through which to achieve their goals, and as such their intentional actions need not be altered, simply replicated. However, in some situations unnecessary or even counter-productive elements make their way into the behavioral sequence. Further, some situations can render certain elements of a behavioral sequence irrelevant. A child watching a peer climb on a box to switch on a wall light may copy their actions despite being several inches taller and not needing the extra height provided by the box to carry out the same action. Children from the age of 18-months increasingly incorporate these superfluous, redundant actions into their replications of another's intentional behaviors in an experimental setting (Nielsen, 2006). They replicate not only the intended goal, but also the intentional actions of the model and they do so more often than not in our target age group. Such over-imitation findings are robust and over-imitation persists even when children have the necessary causal knowledge to identify the additional actions as unnecessary (Lyons, Young & Keil, 2007; Nielsen & Tomaselli, 2010). Findings such as these show just how much emphasis children put on the intentions of others, even at the expense of physical economy. As a result their flexible representations of causal structure go beyond just physical need, but also incorporate social, intentional information.

Imitation is a smart strategy, in a way similar to counterfactual reasoning, as it reduces the requirement for personal, explicit trial and error experience of the world. Understanding that there is a high level of self-other equivalence between you, and other agents you may observe, makes possible this additional opportunity for causal

learning independent from one's own practical experience. Observing the behavioral steps another intentional agent undertakes to achieve some physical or social goal provides a similar short-cut to counterfactual reasoning (CFR) when it comes to planning and executing one's own behaviors. Seeing someone else operate a novel object, or use language in some communicatively novel way affords one the opportunity to represent said actions or utterances, reason about their causal implications and, if the context arises, replicate some or all of the observed behaviors to their own ends. Just like the mental modeling and rehearsal involved in counterfactual reasoning, where using one's own knowledge of, and inferences about causal structure can guide future behaviors, imitation too provides an additional source of causal information above and beyond one's own experimentation that can be used to guide future behaviors. Both practices, counterfactual thinking and imitation, reduce the need for hands on, trial and error experience of an object or situation having been able to acquire knowledge through some other cognitive strategy. Where counterfactual reasoning asks children to reason about something that once was, but no longer is possible, imitation tasks observe how children reason about and respond to something that has happened and can happen again. For counterfactual reasoning children must generate a novel world, different only by virtue of a counterfactual alteration. Imitation requires children to reverse their role in the world from observer to agent, they move from what someone else did to what they could do. When children imitate they have at their disposal the recent memory of another's behavior and often the required steps to reach some desirable goal state. For counterfactual reasoning there is much less available information or contextual support as the to be reasoned about scenario did not happen and could not now happen. However, there are many striking similarities between the demands of both

tasks. Both CFR and imitation require a judgment to be made based upon some passed event. In CFR a model of this passed event must be altered to reflect some counterfactual change. In imitation the passed event must be translated from the perspective of another into the actions of the self. The act of imitation requires that the imitator is aware of the causal structure of the behavior in question to achieve a given goal. It further requires mental state understanding as imitation specifically refers to the replication of intentional goals. Furthermore over-imitation necessarily requires intentional understanding as children who over-imitate copy not only the goal but also those actions that although carried out intentionally, are instrumentally redundant to a perceived goal.

### **1.5. Language**

Language is a powerful causal tool that can be used to achieve goals both physical and social in nature. Language connects us to other people: we can impact upon and shape their thoughts as they can do with ours through a shared language. As such, language affords us formidable collective powers (Pinker, 1995, pp. 16). When we talk about the imitation of behaviors we would often think first of physical behaviors such as playing with a toy or game. However, one of the most freely available modeled behavioral resources that typically developing children experience is spoken language. By the age of four, language is one of the most striking socio-cognitive skills that continues to rapidly develop. Spoken language, and communication more generally, is a central part of the human experience. Language learners have numerous aspects of language to attend to when constructing their own utterances such as their choice of words, grammar, syntax and cadence. Language, however, is a unique ability in that it is both necessarily imitative and productive. To be understood, children must be replicating items from the shared lexicon of their

native language whilst simultaneously using these tokens, and the structures in which they reside flexibly to achieve their novel communicative goals. Language, however, goes beyond servicing communicative goals. It gives children a framework in which to conceptualize complex and/or abstract ideas, to think about possibilities, to reflect on things temporally distant, and to give and receive information about things not present or observable. Language is tool with many functions, none less than allowing us to reflect on alternatives to the actual, and realize the possible.

Not only does language allow us to communicate about the actual versus the possible: language itself embodies the actual versus possible distinction. Language can be highly conventional using only utterances heard before, highly innovative combining linguistic units in a novel way, or somewhere in between. It is the combination of imitative and novel aspects that led us to include language as a measure in our study. We chose two language tasks to collect a breadth of information about linguistic imitation.

Children copy behaviors to induce physical outcomes: children also copy behaviors to induce psychological outcomes for both themselves and others. Children begin copying actions and vocalizations in selective, instrumental ways but with time they begin to copy many intentional behavioral features, both causally important and redundant in nature. They move from thinking and behaving in terms of some physical, instrumental goal to becoming more socially oriented in their endeavors. As mentioned earlier this move from instrumental to social imitation is a prevalent and highly robust strategy used by four-year-olds in the action domain. We were interested in concurrently extending this line of research into the linguistic domain. Is the strategy of imitation one uniformly deployed across a diverse number of tasks or is imitation a highly selective strategy that works independently across different

contexts? We were also interested in whether one's imitativity in both the action and linguistic domains was positively associated with language competence more generally, i.e., its effective use as a tool.

## **1.6. Research Aims**

In this thesis I investigated various forms of conditional reasoning, and the imitation of actions and speech through the lens of how well children navigate actual versus possible worlds. We wanted to generate an idea of how competent children were in their counterfactual and conditional reasoning. Chapter 2 introduces counterfactual conditional reasoning and provides an outline of its suggested developmental trajectory. We investigated the performance of our sample on two measures designed to assess counterfactual conditional reasoning in young children. We contrasted performance on concrete versus more abstract tasks, and also contrasted conditions where prior knowledge may or may not bolster performance to help elucidate the circumstances under which children at this age were more or less successful at reasoning. The chosen tasks probed our sample's competence in reasoning about a world different by virtue of temporal and perspectival affordances, and as such contingent on a diverse set of causal structures. Some of the items could be answered based on prior knowledge alone where others required more complex, and crucially context dependent reasoning procedures.

Chapter 3 is dedicated to the exploration of our sample's domain general cognitive abilities, such as those characterized as executive functions. Two standardized measures of executive functioning taken from the NIH Toolbox (Version 1) were administered to our sample. These included the Flanker task of inhibitory control and the Dimensional Card Change Sort task of cognitive flexibility. As mentioned earlier, during counterfactual and mental state reasoning children need to

be able to disengage from current reality, the actual, and switch their attention to a new representation of the world, the possible. The ability to select one answer from another, reject one option in favor of another, switch attention between different features and flexibly integrate information from multiple contrasting sources are some of the hall marks of maturing executive functions.

Executive functions make it possible for us to represent information abstractly, move flexibly between these contrasting representations and, plan and execute responses based upon them. Executive functions constitute those abilities required to think abstractly and manipulate ideas and are crucial for the discovery and retention of new information about the world (Anderson, 1998). The tasks used to assess executive functions, counterfactual reasoning tasks and indeed mental state understanding measures share a need to infer, based on current reality, some feature of a future, counterfactual or mental state reality. Each are subject to tension between two or more conflicting options, representing different representations of the world, the actual and the possible.

Chapter 4 focuses on mental state understanding and how children's reasoning about others' minds may impact upon their reasoning and decision making in other realms. We first wanted to assess how well our sample could predict the behaviors or mental states of others. We then explored the possible relationships between mental state reasoning, executive functioning and counterfactual reasoning. These proposed relationships were then tested to identify their existence and strength statistically.

In Chapter 5 we present experimental data on the pattern of children's object imitation and how this is influenced by the causal necessity of modeled actions. Children saw two novel objects, in different conditions, on which two actions could be completed. For one condition both actions were causally necessary, for the other

only one of the two completed was causally necessary. Children's patterns of response to these objects and conditions were then correlated with their reasoning task performances collected in Chapters 2-4. We hoped to speak to the question of whether children's imitation of causally necessary and unnecessary actions was related to both their causal counterfactual and mental state reasoning competencies.

In Chapter 6 experimental data on children's imitation of various features of language is presented. This chapter focused on the levels of imitation and innovation in the productive language of our sample. Experiment 1 features children's imitation of simple sentences, some of which contain grammatical or ungrammatical repetitions. These two target conditions were chosen to chime with the causally necessary and unnecessary conditions of the object imitation paradigm. Experiment 2 focused on the imitation of novel lexical items and syntactic structure during a novel verb-learning paradigm. We wished to investigate the patterns of imitation children produced in response to these varying features of language and whether any conceptual contiguity was seen with the imitation of actions. Finally Experiment 3 specifically examined the relationships between the imitation of these different features of language and the object imitation response patterns presented in Chapter 5.

In Chapter 7 language remained the focus for study, however, we moved from being interested in the paradox of imitation and innovation both being hallmarks of linguistic productivity to language's functionality as a tool within the world. Specifically we investigated the relationship between the reasoning competence of our sample and their imitation of our three key features of language. Furthermore we outlined the relationships between their skills at reasoning counterfactually and about mental states to their linguistic competence as assessed by a novel composite score extracted from our experimental measures.



Finally Chapter 8 the General Discussion integrates the experimental data that has been reported in the earlier chapters. The relationship between: children's reasoning about events in the world that are temporally or cognitively set apart from their own egocentric view and their actions and utterances in the world is explored.

We assess the hypothesis that children who reason more successfully about counterfactual situations, and are more able to infer behaviors based on mental states, are concurrently better equipped to imitate in order to achieve a given instrumental goal. We relate their counterfactual and mental state reasoning skills to their proclivity for imitating different aspects of others' intentional acts, including actions on objects, the repetition of ungrammatical language and the use of novel verbs in a complex syntactic structure. In addition we probe the possibility that imitation as a more general strategy across both action and verbal domains can account for linguistic competence more generally.

All the experimental data we collected was an endeavor to delve deeper into two interconnected theoretical questions. Firstly, we wanted to understand how well our sample reasoned with conditional rules in a range of situations, and how well they could successfully select a possible incongruent, counterfactual or psychological world based upon features of an actual target world. Secondly, we wanted to investigate whether this same sample's reasoning performances related to their selection and execution of possible actions or utterances modeled in the actual world.

## **Chapter 2. The Road to Reason: Conditional and Counterfactual Reasoning at Age Four**

### **2.1. Abstract**

From moment to moment, as we explore the physical and social worlds that we inhabit, we are exposed to information from which we may infer causality. Regardless of whether this information reaches us passively, or we have sought it out, we are potentially able to use it to formulate or improve our own behaviors (Hume, 1739/1978). Infants who are new navigators of the world are less able to use causal information in this way. In some cases the causal information required for a task has never been available to them: in other cases the ability to successfully process causal information that they have encountered is not fully functional. Children are not restricted to putting causal information to use in a physical way, they may also use newly obtained causal information to mentally rehearse current or future scenarios. In addition children's reasoning based on causal knowledge need not be restricted temporally, that is reasoning only about present or future events: they may also reason causally about events in the past.

In this chapter I will discuss the development of children's causal conditional reasoning in the context of counterfactual thinking, making inferences about a situation that is counter to fact. I will examine the tasks traditionally used to assess childhood counterfactual reasoning ability and compare them with newer, exigent measures exposing the limits of young children's counterfactual reasoning performance. I will present experimental data collected from a group of four year-old children using both traditional and newer, more demanding reasoning measures. The experimental data contrasts performance between the tasks. In addition the effect of

typicality, or congruence between subject matter heuristics and counterfactual outcomes is explored.

## **2.2. Introduction**

Everyday we find new information and learn new skills. We may, after several attempts, succeed in operating a novel object, or we may observe someone else operating this object and in future utilize the same, or similar skills with the same or similar object. Perhaps we offer, or we may see someone else offer, our friend some chocolate when they appear sad, and observe them magically become a little cheerier as a consequence. Not only do we observe such relations but we also have the ability to reflect on, or even extend them. Perhaps, for example, we see the power of the chocolate in current reality and this triggers our memory of a past reality where another person was sad. We may then mentally simulate the effect chocolate might, but no longer could have had on this person. This type of mental rehearsal is called counterfactual reasoning and constitutes a powerful tool for adults and children. Children become able to cognitively test the causal information that they possess, for example in response to a poor outcome in their physical or social world, be it an enigmatic object or an upset friend. Children can mentally contrast the actual and counterfactual outcomes of a situation given some key change in the causal chain. Such complex abilities are based on causal and conditional reasoning as expanded upon below.

Causal knowledge is the cornerstone of planning and predicting future outcomes in the world (Hagmayer, Sloman, Lagnado & Waldman, 2007) and such knowledge increases with age (Sobel, Tenenbaum & Gopnik, 2004). Through observation, events that frequently occur together are inferred to have some

cause/effect relationship (Hume, 1973/1978). Causal reasoning is a form of inductive reasoning as conclusions are based on knowledge about what happens all or most of the time given a set of circumstances. Additionally through intervention, or action in the world potential causal structures can be identified or tested (Hagmayer et al., 2006). Existing causal knowledge obtained through causal reasoning can go on to support more complex forms of conditional reasoning. Where causal knowledge specifically refers to the identification of cause and effect relationships between events/objects, conditional reasoning more broadly refers to reasoning based on if-then relationships that need not require additional identification of causal structure (Ali, Schlottmann, Shaw, Chater & Oaksford, 2010). Conditional reasoning refers to the extraction of inferences from statements with an *if-then* structure and is central to human inference (Ali et al., 2010). The *if* element forms the argument's antecedent, with the *then* component forming the argument's consequent. Formal logicians often use abstract rules such as " $a \rightarrow q$ " to investigate conditional reasoning performance (eg; Markovits, 1985). However, semantically meaningful statements can also be utilized: "*If it snows, then the class will be cancelled*". Without basic causal reasoning abilities, that essentially work to provide some structure to our representation of the causal world, we would not be able to make conditional inferences without being expressly informed of the relevant causal structures involved. An example of conditional reasoning being supported by existing causal knowledge follows. When asked a question such as "If the egg falls from the counter, what will happen then?" adults can plausibly infer that the egg will crack based on their prior experience of gravity and egg fragility. The latter question is an example of Basic Conditional Reasoning (BCR). More specifically, the given example requires reasoning about imagined events based on previous causal reasoning. Basic conditional reasoning

tasks can be described as timeless due to their propensity to be expressed in the present or neutral tense and relate to established regularities in the world (Perner & Rafetseder, 2011, p.91-98). They can be introduced and reasoned about in isolation. Basic conditionals have been described as mostly independent from context and as having antecedent events that are possibilities, and consequents that can transpire as a result of these possibilities being fulfilled (Johnson-Laird & Byrne, 2002). Basic conditional reasoning relies upon default assumptions that are plausible or typical in any given context and do not rely on some specific content knowledge pertinent to a specific sequence of causal events (Perner & Rafetseder, 2011, p.91). Basic conditionals can refer to “if-then” or “if-then possibly” type statements and often include a causal relationship, be it absolute or possible.

Counterfactual Reasoning (CFR) requires a strategy that is taken counter to actual events. In contrast to basic conditional reasoning, the assumptions used during CFR must be imported from the event’s sequence itself and not rely solely on plausible assumptions (Perner & Rafetseder, 2011, p.91). Unlike basic conditionals where the antecedent is a possibility, counterfactual antecedents are necessarily impossible in the context of current reality, but treated as possible in the context of a counter-reality (Johnson-Laird & Byrne, 2002, p.4). Say the event sequence was described as follows, “A toy egg was rolled to the edge of a counter but stopped before it reached the end and fell to the floor.” Then the following counterfactual antecedent was introduced, “What if the egg wasn’t a toy, would it have cracked?” The correct counterfactual conclusion (although not possible in the context of the actual world) would incorporate all the events from the actual event sequence including the assertion that the egg stopped before it reached the end of the counter. As a result the counterfactual and actual consequents would be the same: the egg

remains intact by virtue of it never reaching the end of the counter. If, however, not all the events were imported, such as the egg stopping before the edge, the counterfactual consequence would be different and the egg would have broken. Although this consequence is still plausible, with people focusing on the altered consequent of the egg's properties, it does not remain true to the actual sequence of events.

CFR has the potential to go far beyond this simple scenario, however, when the causal knowledge being used is inferred from much sparser probabilistic information, what Tenenbaum, Griffiths and Kemp (2006) refer to as an “inductive leap”. From a young age children are able to make causal judgments and predictions based on inferences drawn from diverse sources, often not specifically related to the problem in hand. Attempts have been made to characterize how children combine their probabilistic knowledge and new “data” and experience when decision-making. Studies applying the formal mathematical approach of Bayesian belief networks, for example, combine both probabilistic information and causal relationships within directed causal graphs representing causal relationships between variables (e.g., Glymour, 2001; Schulz, Gopnik & Glymour, 2007). Bayes nets then allow for flexible representations of causal structure to be based on both observations and interventions, with information obtained from both these sources resulting in an update or change to the representation of causal structure. These causal nets provide flexible representations of causal structure that can be continually altered or embellished to reflect new information. In addition such causal nets support judgments about the impact of interventions and counterfactual situations. In combining Bayesian and development approaches, some studies have put forward strong claims that children's nascent representations of causal structure can indeed support predictions about

intervention and counterfactual reasoning (e.g., Gopnik, Glymour, Sobel, Schulz, Kushnir & Danks, 2004; Schulz, Kushnir & Gopnik, 2007; Schulz et al., 2007). In particular experimental evidence suggests that between the ages of three and four years children's learning moves more towards a Bayesian structure, that is less reliant on association alone, and indeed supports the generation of novel actions based on their causal graphs (Sobel, Tenenbaum & Gopnik, 2004).

Returning then to CFR, reasoning about the possibilities in a counterfactual world, is highlighted as a critical component of adult cognition (Markovits & Barrouillet, 2002). Being able to consciously appraise the outcomes of our own, and other's actions in the landscape of what could have been, or could be, has beneficial effects on our decision making, analytical reasoning, feelings of regret and relief, and emotions and motivation (Smallman & Roese, 2009; Kray, Galinsky, & Wong, 2010; Guttentag & Ferrell, 2004; McCrea, 2008). Being able to incorporate causal information not only about what was, but crucially what could have been, gives us additional information upon which to base our future decisions. Counterfactual thinking abilities increase the amount of causal knowledge that we might have access too, making for a more informed decision, in so far as the results of our counterfactual musings were causally sound.

Recall the previous example once again, "*A chef places an egg on the counter: the egg falls and cracks*". A counterfactual question often negates an antecedent of the statement's consequent, for example "*What if the egg **hadn't** fallen?*" One possible response is what Riggs, Peterson, Robinson and Mitchell (1998) describe as a 'realist response' (a term borrowed from the earlier mental state understanding literature of Wimmer & Perner, 1983). A realist response is given when the reasoner answers in line with the actual world and not the counterfactual. In this case a realist

response would be saying the egg would still crack. Adults, however, can be expected to easily conclude the egg would not have cracked in the majority of cases. When true counterfactual reasoning is taking place the reasoner must be generating more than one target model. The first target model represents the actual sequence of events. Additional models are also based upon the same representational target. However, the content of these models must differ as a function of the altered element or elements involved, and include any causal repercussions these alterations exert in the target world (see Figure 2.1.). Counterfactual worlds are inextricably bound to the reality from which they depart and must be coordinated for successful CFR to have taken place (Byrne, 1997, p.108; Rafetseder, Schwitalla & Perner, 2013). As discussed earlier counterfactual reasoning is only successful when all elements of the counterfactual event sequence remain the same, with the exception of those necessarily altered by the counterfactual premise. As such it is necessary for the actual and counterfactual worlds to be co-ordinated in the mind of the reasoner as so far as to maintain this nearest possible world (Lewis, 1973). In Figure 2.1 element 1b is based upon element 1a and differs only in the counterfactual antecedent. Causal knowledge can then be used to generate element C, based on element 1b. If additional changes are made then the counterfactual conclusion cannot be valid. Although much of this section, and indeed the literature, focuses on concrete physical counterfactuals the same principles can be applied to social situations. Take for example reasoning about the effect of one individual's behavior (offering chocolate) on another individual's mood.

In an event sequence like the egg falling off the table the effect of prior knowledge often produces an interesting effect on reasoning. Counterfactual reasoning questions involving situations that are typical or regular can instead be



successfully reasoned about using a basic conditional reasoning strategy. Often our knowledge or memory of such situations can generate default assumptions about causal relationships, independent to the actual context being presented. In Figure 2.1 the basic conditional comes straight from prior knowledge rather than the generation of some new model based on the actual world. Recall that the CFR question was “*What if the egg hadn’t fallen?*” Using a simple basic conditional reasoning strategy involving only one fictive representational model extracted from past experience, (*if an egg doesn’t fall, then it doesn’t break*), is just as effective as reasoning counterfactually, that is from two models, one of actual events and one of fictive events. This introduces a problem for CFR research: the ability to reason counterfactually can only be attributed when the use of any other reasoning strategy would result in an incorrect inference. This distinction between basic and counterfactual conditional reasoning is important in the context of this thesis as we wished to explore the limits of children’s reasoning abilities about actual and possible worlds and, how these abilities might contribute to other social and cognitive skills. Being able to effectively navigate the causal repercussions of changes to the world provides the basis for many future decisions about ones own interaction with objects and people in the world. As such the distinction between basic conditional, and the more challenging counterfactual conditional reasoning affords very different opportunities for judgment and decision making. The ability to mentally rehearse a counterfactual situation provides a short cut for children, as they do not need to have experienced an alternative course of action to be able to predict its causal outcome.

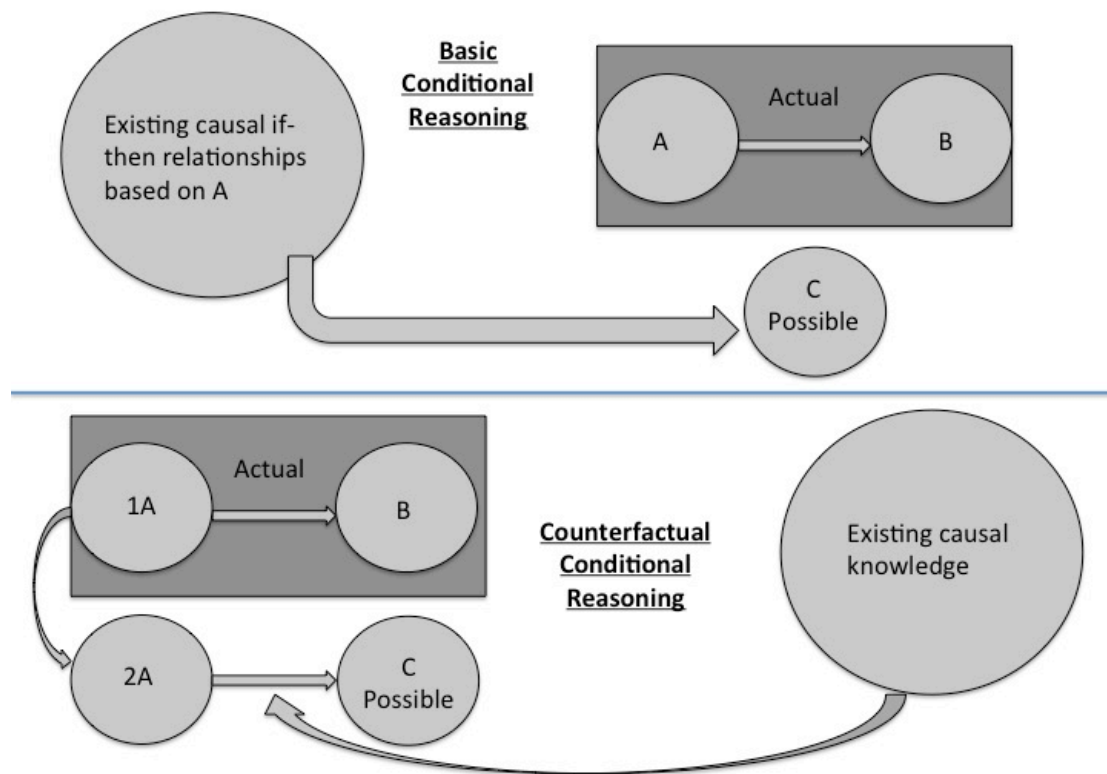


Figure 2.1. Visualization of the Difference Between Basic and Counterfactual Conditional Reasoning.

Given the distinction between these two forms of conditional reasoning we wanted to examine both in our sample of four year-old children. We tested basic and counterfactual reasoning ability to understand firstly, if they are related to one another and secondly, if they are related to other cognitive indices. In the following sections we will explore some of the measures used to gauge children’s reasoning ability and the conclusions drawn from them.

It is problematic then that some measures specifically developed to investigate counterfactual reasoning can in fact be successfully answered using basic conditional reasoning principles. This paradox is especially pertinent when examining the ability of children and young adults to reason counterfactually. As a result only situations in

which basic conditional reasoning and CFR generate conflicting answers can be used to identify which individuals can successfully utilize CFR strategies (Rafetseder & Perner, 2010; Rafetseder, Christi-Vargas & Perner, 2010; Rafetseder et al., 2013).

Such an example can be seen in the sweet story (Figure 2.2. Rafetseder, et al., 2010). The situation is set up like so, “Mum buys sweets and leaves them in the kitchen. Simon can only reach the top shelf because he has a broken leg and cannot bend down. Julia can only reach the bottom shelf, as she is not as tall as her brother. When they find sweets the children always bring them to their rooms. Mum buys sweets and leaves them on a high shelf. Simon comes into the room” Consider being asked the following question after listening to the story, “What if the little girl had come into the kitchen, where would the sweets be now?” If the reasoner applied a basic conditional reasoning strategy relying on default assumptions about the desire for sweets they would conclude that should a young girl come into the kitchen and there be sweets, she would take the sweets to her own room. This basic conditional reasoning strategy ignores much of the relevant causal information such as the fact that the sweets are still on a high shelf and cannot be reached by said little girl. Similarly had the reasoner been asked, “What if the mother had put the sweets on the bottom shelf, where would the sweets be now?” Again the basic conditional reasoning strategy would focus on the assumption that someone would get the sweets and therefore conclude the sweets would be in the little girl’s room, ignoring the fact that the boy remained the character who entered the kitchen subsequently, not the little girl. Only a true CFR strategy, one that appropriately imported all the relevant information about; the sweets’ location, the character’s abilities, and the searcher’s identity could successfully answer this problem. If the girl and not the boy had come in, or had the mother put the sweets on the bottom shelf - the sweets would remain on

the shelf in the kitchen. In the majority of developmental investigations, however, tasks that can satisfy this requirement of disentangling CFR and basic conditional reasoning have not been used (Rafetseder & Perner, 2011).

### **2.2.1. Assessing the Development of Counterfactual Reasoning**

In the developmental literature the majority of paradigms used to assess CFR are counterfactual conditionals in which children must reason about consequences given false antecedents (eg: Riggs, Peterson, Robinson & Mitchell, 1998; Rafetseder et al., 2010; Beck & Crilly, 2009; Rafetseder & Perner, 2010; Rafetseder et al., 2013). Typically children will be asked about one of the antecedents causally related to the outcome event in the story, “What if X hadn’t happened, where would Y be now?” or similar. Harris, German and Mills (1996) showed children aged 3-4 years a story in which a protagonist with muddy shoes walks over a clean floor. Children were then asked what would happen if the agent’s shoes had been removed. The majority of children (75%) successfully inferred that the floor would be clean, leading to the claim that they were reasoning counterfactually. Indeed the children were reasoning counter to fact: however, a simple basic conditional reasoning strategy based on children’s existing causal knowledge would also result in the correct answer: *if a person with clean socks walks across a clean floor, then the floor will be clean*. Both, Beck, Robinson, Carroll and Apperly (2006) and Rafetseder and Perner (2010) suggested that CFR questions that may be successfully answered in isolation are easier than those in which reference to the actual sequence of events is necessary. Indeed 3.5 year-olds perform better in CFR questions in which the correct location is one semantically linked to the story’s main character, thereby creating a default assumption as to where the actor would be (Rafetseder & Perner, 2010). The typical condition in the Doctor story (see Figure 2.3.) benefited from the default assumption

about where a Doctor would normally be – a hospital. This is in contrast to conditions where the correct answer required the Doctor to be in a semantically unrelated location like a park that carries no default association. Despite the boost in performance for the typical condition, across condition performance was still poor in 3 and 4-year-old samples (Robinson & Beck, 2000; Rafetseder & Perner, 2010).

Evidence such as that presented by Rafetseder and Perner (2010) suggests that young children may not be consistently reasoning in a truly counterfactual way. Tasks that may only be consistently answered successfully through the use of counterfactual conditional reasoning highlight the fragility of children's developing reasoning abilities. In the absence of other supporting factors or congruent expectations children struggle to reliably identify a counterfactual outcome. In the case of the kind of story tasks mentioned, young children do not have the same strength of prior expectations and knowledge to aid their decision-making as in the dirty floor task. Story tasks that set up novel situations about the location and accessibility of a pack of sweets provide little opportunity for past experience of a similar scenario being imported into the counterfactual model when compared to a task relating only to the effects of dirty shoes upon a clean floor. Similarly an expectation about the location of a person or object can only be present if that person or object, and representations of their semantically related locations, are familiar and easily accessible in memory. A lack or weakness of any default assumptions at the group level about Doctors or other familiar professionals that could have been used to drive basic conditional reasoning, then may impact upon the poor performance shown in conditions where these defaults would have been congruent with the counterfactual conclusion (Rafetseder & Perner, 2010; Robinson & Beck, 2000).

Given the integral positioning of counterfactual reasoning within adult human cognition its development provides an exciting area for examination in the wider context of cognitive and communicative development. Counterfactual reasoning allows children to generate inferences about how any behaviors they intend to carry out would work to achieve their goals: it allows them to mentally test the impact of these possible actions on the actual world. As children begin to appreciate their own, and others intentionality the task of mentally forecasting the causal repercussions of one's actions or vocalizations is considerable. Reasoning about how reality could be, or crucially in the case of counterfactual reasoning, could have been changed, is a priceless resource for those starting out on their own journey as an intentional agent.

Often the content of a reasoning question can have dramatic influences upon an individual's ability to reason successfully. Location and state change counterfactuals involve physical, observable movements or changes for the objects being reasoned about. This contrasts with CFR tasks involving necessarily abstract or unseen changes such as those involving mental states. Given that CFR requires the generation of, and comparison between different mental representations of the same situation, counterfactuals that refer to concrete, and therefore more easily represented situations, will be more successfully dealt with than those requiring more abstract representational inputs. Mental state reasoning is one such example of a skill where pre-school children continue to perform poorly, as representing the minds of others is more challenging than representing the movement of a visible object (Birch & Bloom, 2004). Using paradigms that depict concrete observable actions helps remove the additional difficulties young children have with reasoning about any abstract or unseen properties within a vignette. Beck et al. (2006) reported that children as young as 3-4-years were competent, achieving an 87% success rate, during concrete

counterfactual questioning about a ball being dropped down an inverted Y shaped slide. In this “standard” condition they were asked “What if the ball had gone the other way, where would it be?” Furthermore Harris et al. (1996) also attributed impressive counterfactual reasoning ability to 3-4-year-olds. 75% could conclude that a model floor would be clean had a protagonist taken off their muddy shoes (Harris et al., 1996).

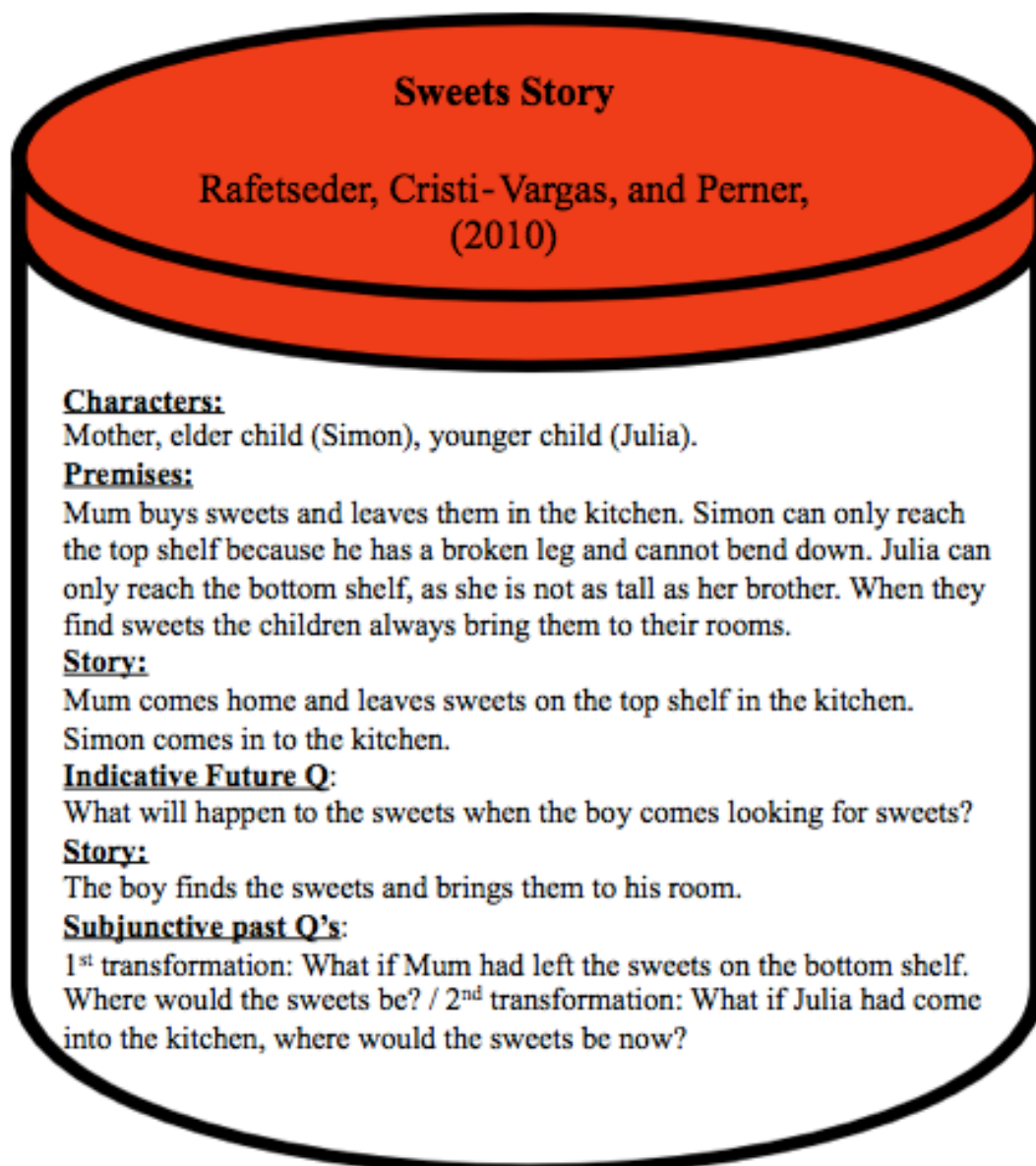
Despite this evidence, 3-year-old children completing more complex belief-free (although not intention free) counterfactuals involving increased amounts of causal information often answer as realists (e.g., Riggs et al., 1998). Recall realist answers are given when the reasoned answers in line with the actual world rather than the counterfactual world. Returning to the egg example then a realist answer to the counterfactual question “What if the egg hadn’t fallen?” would be “The egg would have cracked.” Realists do not alter the outcome despite the causal information available. In the Riggs et al. (1998) study children were asked to reason about a fireman named Peter who was ill in bed, but on hearing there was an emergency raced to help: where would he be had the fire not started? Many responded that Peter would still be at the fire. They fail to inhibit what they know to be true and persevere with answering that the outcome would be the same, regardless of changes to events earlier in the causal chain. Children seem to begin relinquishing this realist strategy only in their 4<sup>th</sup> year (Riggs et al., 1998; Guajardo & Turley-Ames, 2004), and, only then, answer in accordance with the counterfactual outcome at the group level.

Rafetseder et al. (2010) challenged the conclusion that even results such as these are evidence for true counterfactual reasoning ability. Instead they suggest that many tasks purporting to index nascent CFR skills are instead driven by the application of the basic conditional reasoning processes discussed earlier (Rafetseder

et al., 2010). Children are argued to reason in isolation from the actual events and construct only one model, compatible with the premise, rather than the two + representations necessary to attribute CFR.

To demonstrate children's adherence to basic conditional reasoning processes Rafetseder et al. (2010) utilized an exigent task in which CFR & basic conditional reasoning produced conflicting conclusions (see Sweets Story in Figure 2.2). We too opted to use different conditions that could dissociate basic and counterfactual conditional reasoning skill. In Rafetseder et al. (2010) the sample of 3-6-year-olds' performance is poor when compared to the 87% of correct responses reported by Beck et al. (2006) by 3-4-year-olds on a standard CFR question "What if it had gone the other way?" Indeed only 6% of the sample in Rafetseder et al. answered subjunctive past (counterfactual) questions correctly. That is only 6% were able to maintain the pertinent information within the story to correctly identify that the sweets would remain in the kitchen.





*Figure 2.2.* Example of an exigent counterfactual reasoning paradigm in which only true counterfactual reasoning can reliably produce the correct response. Adapted from Rafetseder, Cristi-Vargas, and Perner, (2010).

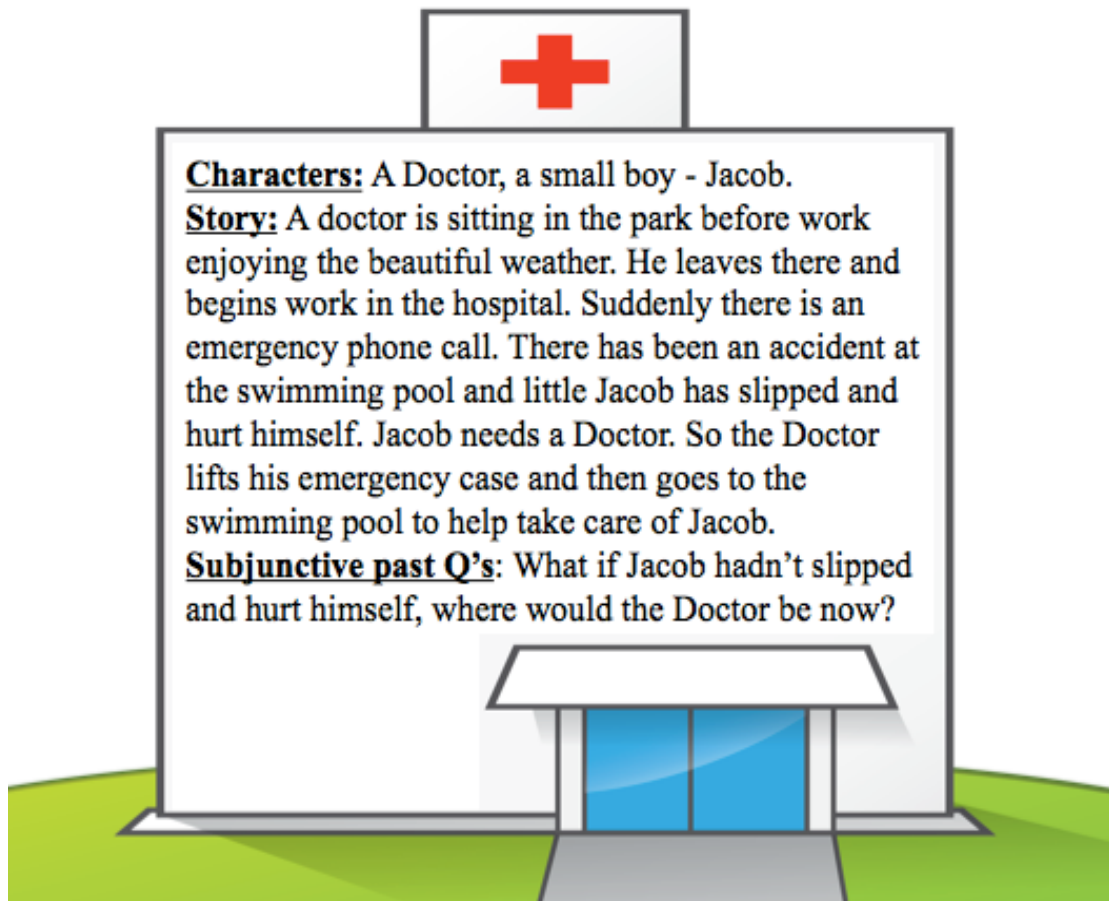


Figure 2.3. Example of a typical (congruent) story used to assess the effect of typicality on counterfactual reasoning judgments. Adapted from Rafetseder & Perner, (2010).

6-year-olds alone did perform better at around 24%, a figure still far below the success rates of adult subjects (100%). A subsequent paper identified 12-14-year olds' as reaching statistically similar attainment levels as adults (88:100), with questions in which BCR and CFR lead to incompatible conclusions (Rafetseder et al., 2013). In this chapter we aimed to unite these findings that young children perform well on traditional location change counterfactuals when compared to the exigent story tasks developed by Rafetseder and colleagues where performance is strikingly poor throughout childhood. The traditional tasks could be successfully answered using

basic or counterfactual strategies, whilst the exigent tasks could crucially only be reliably completed using a counterfactual strategy.

Young children who have relinquished a realist strategy and engage in CFR tasks do seem to appreciate that if the antecedent is false, then the consequents will change, therefore rejecting realist answers. In tasks that can be solved using default assumptions then the basic conditional reasoning abilities in place can help formulate the correct response. However, applying basic conditional reasoning in situations where there are 2+ other locations or states available, and where neither of these locations have a significantly more salient relationship to the person or object involved in the transformation, means children may randomly select one of the other options, (Rafetseder, Eckmaier & Perner, unpublished manuscript). The use of basic conditional reasoning before CFR could then lead to distinctive patterns of responding in children, given that the material being reasoned about brings to bear some pertinent prior knowledge. If basic conditional reasoning relies on default assumptions about the object or person whose location or state stands to change, then it follows that for situations in which defaults would aid CFR children's performance should increase.

In the example Sweets Story (Figure 2.2.) children reasoned that the sweets would be in the room of whichever child came into the kitchen, regardless of causal validity. Rafetseder and Perner (2010), wished to elucidate whether a typicality bias about familiar professions would similarly influence responding. The addition of a typicality bias gave the authors a way to identify whether highly salient default reasoning assumptions did indeed interfere with CFR in young children, by way of altering performance. Indeed as seen in the previous section performance was greatly improved in typical conditions (see Figure 2.3.) where the CFR and typical locations were congruent, that is the correct CFR response was the location most semantically

salient to the protagonist. This provides further evidence that in such CFR tasks, young children rely more so on basic conditional reasoning than CFR. Something is stopping children from performing well on these types of tasks.

It seems clear then that those tasks traditionally used to assess CFR reasoning ability are potentially only giving conclusive evidence of one distinct form of conditional reasoning – basic conditional reasoning. CFR attribution can only be made using tasks that require those processes supporting basic conditional reasoning to be inappropriate for reasoning success in this case. Some additional ability or set of abilities is required to ensure reasoning can go forward successfully. It is clear that the structure and content of such tasks are generally cognitively demanding. But it also seems likely that one cannot develop CFR without first developing basic conditional reasoning processes. Basic conditional reasoning, like counterfactual conditional reasoning relies on the understanding of possibilities (Beck et al., 2006). In the case of the basic conditional the possibility is not temporally bound. In the counterfactual, however, the possibility can only exist in one reality, one that is temporally abstract due to its opportunity of fruition having passed.

In addition to being governed by different temporal constraints, basic conditional reasoning involves reasoning from 1 model, where CFR demands reasoning from 2+. Basic conditionals can exist in isolation whilst counterfactuals are inextricably bound to the realities from which they have departed. Is performance on these types of task correlated? And if so do any additional cognitive indices feature prominently in this relationship?

### **2.2.2. The Relationship Between Counterfactual Conditional Reasoning and Other Cognitive Abilities**

One potential bottleneck for young children when reasoning are their executive functions (EFs). Executive functions are introduced further in Chapter 3. From a basic stance EFs allow children to select a response from a set of options. Children's comprehension and subsequent selection may not be congruent – therefore incorrect answers may be due solely to weak executive functions. Inhibitory control broadly refers to the ability to inhibit a pre-potent tendency associated with some stimulus (Gerstadt, Hong & Diamond, 1994). Inhibitory control is one of the domain-general processes included under the umbrella of executive functions. Inhibitory control is hypothesized to be crucial to CFR due to the need for children to inhibit a realist response, a response type seen often in 3 year-old children (Riggs et al., 1998). Recall, a realist response is one in which the participant ignores a counterfactual transformation and fails to acknowledge that the situation would change. Children who give realist responses could be failing to reject the lure of the most salient, pre-potent, actual world response.

Cognitive flexibility and working memory then are thought to be utilized due to a need both to generate and compare representations of the actual, and counterfactual worlds in order to reason successfully (Guajardo et al., 2009). If children are not able to generate and switch between world-views, of the actual and possible/counterfactual, then they will be unable to consistently perform well in CFR situations. It seems the basic task demands of CFR necessarily involve a certain level of EF proficiency to satisfy the procedural task demands.

However, investigations directly scrutinizing the role of EF in CFR have yielded mixed results, and have generally been restricted to traditional CFR tasks: that is those theoretically passable through BCR alone. Beck et al. (2010), utilized a CFR battery to better assess its' relationship with EF. Although they failed to find an effect

of working memory on children's reasoning they did report a predictive effect of inhibitory control and receptive language. Beck et al. (2010) argued this effect was realized through children's increased ability to inhibit what they know to be true (preventing realist answers). Additional support for a reasoning and inhibitory control link in older children was reported in Simoneau and Markovits, (2003). However, a study of regret, assessed using CFR stories, found that switching (CF), and not working memory or inhibitory control, was a significant predictor (Burns, Riggs & Beck, 2012). Earlier work by Riggs et al. (1998) dismissed inhibitory control as a predictor of CFR as future hypothetical problems (that require the same inhibitory control demands to complete as reasoning problems) were solved significantly more readily. In addition, manipulating the inhibitory control demands of a task did not alter performance in Robinson and Beck (2000). Guajardo et al. (2009) found working memory and cognitive flexibility positively predicted 5-year-old's ability to generate antecedent counterfactuals.

From the numerous mixed results it seems clear that with different conditional reasoning tasks, varied demands will be made of the reasoner. Certainly the process of inhibiting realist answers, and comparing reality to a possible scenario require EFs. However, these domain general abilities alone will not suffice across differing tasks. No single EF has emerged as a clear predictor of conditional reasoning: rather a more general developmental progression has been suggested. It is reasonable to hypothesize that in the pre-school years, a time in which there remain large developmental and individual differences in EF, measures that assess domain general cognitive functions will relate with performance in some way. The size and strength of the relationship, however, will crucially depend on the task in question.

Beyond EF demands, however, each type of counterfactual task can differ along numerous dimensions. For children to reason successfully, that is make valid inferences, they must understand the causal relationships embedded in the content. Specific content may be more or less familiar, causal relationships may be more or less concrete, simple or complex and, the amount and diversity of information can differ vastly. It is not only children's procedural, cognitive ability to reason that is being tested, but also their ability to reason about the causal affordances of certain scenarios. To return to the example of the egg: children who have the necessary cognitive capacities may not have the necessary causal knowledge, through lack of experience, about the fragility of eggs. If there is no causal structure from which to reason about counterfactual changes, then performance will be poor. The robustness of any causal structure representations can indeed impact upon counterfactual judgments being made (Frosch, McCormack, Lagnado & Burns, 2012). Alongside EFs then, solid causal knowledge about the content is needed in order to solve counterfactual reasoning tasks.

Rafetseder et al. (2013) further pointed out that it is unlikely to be weak EFs alone that account for the specific problems faced by children during CFR, as opposed to basic conditional reasoning. Instead they put forward the claim that problems are due to children not understanding the "nearest possible world constraint" (Lewis, 1973. in Rafetseder & Perner, 2010; Stalnaker, 1968). Children are not adhering to the process of retaining the structure of the world faithfully in all but those features necessarily altered by the counterfactual statement. This nearest possible world constraint constitutes a form of task specific knowledge related uniquely to CFR, as basic conditional reasoning need not be modeled in this form. Only elements that are causally dependent on the counterfactual assumption need to

be changed. All other causally independent elements must be retained. Young reasoners do not appreciate the complexities of what does and does not need to be integrated in the counterfactual model (Rafetseder & Perner, 2010). This problem does not occur in basic conditional reasoning due to the lack of additional models being generated and concurrent reliance on default assumptions about causal structure. Evidence for this comes from CFR problems in which either the location or the searcher is altered in the counterfactual world. In this case children fail to integrate the location or searcher information that was not altered, whether that information was causally before or after the transformation that was made, as in the Sweets Story (Rafetseder et al., 2010). At this point the influence of typicality, and how it diverges from the actual sequence of events is marked. Instead of retaining the particulars of the world being reasoned about in all cases except those altered by the counterfactual statement, children often revert to reasoning based on a typical world. Rather than utilizing the nearest possible world as dictated by the context, they use a more prototypical version of the world. This familiar, typical world benefits from having more stable representations of causal structure, based on real world experience and aligned with the child's self generated representation of causal structure. Often this overriding propensity to reason in a way that is congruent to typicality results in success, as shown by traditional CFR measures. However, when typicality and reality are incongruent, then basic reasoning processes alone are insufficient.

Utilizing the Sweets story again presented earlier (see Figure. 2.2.) for both subjunctive-past questions children overwhelmingly replied that the sweets would be in the little girl's room despite the fact that in both counterfactual worlds the sweets would remain on the shelf. In the 1<sup>st</sup> transformation question only the location of the sweets, and not the identity of the searcher changed (the sweets would remain on the



shelf). In the 2<sup>nd</sup> transformation although the sweets' location remained the same, the searcher differed. Julia could not reach the sweets in this location but children again decided the sweets would be in her room regardless. This provides evidence that reasoning is resistant to where the transformation occurs, performance is still poor and this seems best explained as a result of a failure to integrate the relevant causal information from reality into the counterfactual model. It seems then that working memory cannot be wholly to blame (as the location of the transformation does not systematically alter success). Additionally, in even the youngest group, there were relatively few realist errors, suggesting that inhibitory control, at the group level at least, cannot be accounting for performance. It is not that children are failing to understand the world has changed, it is that they are not using a correct representation, that in which only causally necessary changes are made, on which to base their inferences.

Counterfactual reasoning ability, as measured by more traditional tasks has also been related to other developing cognitive abilities and skills. CFR mediates emotions such as regret and relief: in these cases it modulates perception of the current reality (Beck & Crilly, 2009; O'Connor, McCormack & Feeney, 2012; Rafetseder & Perner, 2011), improves creative association generation in this case modulating future reality (Kray, Galinsky & Wong, 2006) and, correlates with pretense abilities (Buchsbaum, Bridgers, Weisberg & Gopnik, 2012).

Guajardo and Turley-Ames (2004) found high inter-correlations between theory of mind skills and counterfactual reasoning. This echoed work in 3-4 year olds that suggested around 25% of variance in theory of mind was explained by CFR performance (Riggs et al., 1998). Support for a theory linking CFR and theory of mind has also been found in autistic children (Peterson & Bowler, 2000). However,

only 3-16% of the variance in theory of mind found by Guajardo and Turley-Ames could be accounted for by CFR alone. Language and age explained a larger proportion of variance. A later study (Guajardo, Parker & Turley-Ames, 2009) included CFR, theory of mind and EF measures and again reported only limited unique variance in theory of mind being attributable to CFR. Instead they found a mediation pattern through working memory and cognitive flexibility performance. Furthermore Perner, Sprung and Steinkogler (2004) found dissociation between theory of mind and CFR in 3 to 5-year old children, prompting them to describe a link between the two as pure speculation. The relationship between counterfactual reasoning and other cognitive abilities will be discussed in greater depth in Chapter 4 of this thesis.

### **2.2.3. Aims**

We had multiple aims for this first experimental chapter. We chose to use two different paradigms, the Slide task adapted from Beck and Guthrie (2011) and the Story task presented in Rafetseder and Perner (2010). We chose these two tasks to examine and compare both basic conditional and counterfactual conditional reasoning in the same sample. Our first two aims were to replicate the findings of these original investigations.

The easier of the two tasks then, that requiring only basic conditional reasoning, was a location change counterfactual adapted from Beck et al. (2006) and, Beck and Guthrie (2011). This task, henceforth called the Slide task, involved a ball being rolled down an inverted Y shaped pipe-slide following which counterfactual questions about the ball's location were asked. Reasoning performance was reportedly above chance in the Standard condition of this task for both the younger and older groups described in Beck and Guthrie (2011). 49% of the younger group (mean age=

3.9 years) and 74% of the older group (mean age= 5.0 years) answered both standard questions correctly. The standard question asked children to indicate where the ball would be had it gone the other way. In the second condition of the Slide task, the Open condition, children saw the same scenario but were asked “Could the ball have gone any other way?” This question is much less constrained and theoretically children could answer in an almost infinite number of ways and not be correct or incorrect. There are many possibilities associated with this form of counterfactual question. Our first aim then was to replicate the above chance performance rates in the standard condition. As our sample had a mean age of 4.4 years we expected between 49% and 74% of our sample to answer both standard questions correctly.

The second, more challenging task was replicated from Rafetseder and Perner, 2010 and shall be referred to as the Story task throughout. Here a short story was told about a protagonist moving through three locations given some emergency. After the story children were asked a counterfactual question about this protagonist. This task was chosen so that at least one condition would concretely assess true counterfactual reasoning ability while the others could be answered by basic conditional reasoning alone with a reliance on a typicality heuristic. In comparison to the location change Slide task the vast majority of the sample of three to six year olds reported in Rafetseder and Perner (2010) failed this task. For our third aim then we further expected that children would be more successful in the typical condition of this task when compared with the atypical condition. In the typical condition the correct answer was also the semantically congruent location and this congruency was reported to significantly improve performance in the original version.

Moving to our fourth aim for Chapter 2 then we chose multiple tasks so that we would have a range of challenges for children to compare. For the first task (the

slide task) basic conditional reasoning only was required for success. In the second task the two conditions afforded us the opportunity to look at scenarios where either both basic and counterfactual conditional reasoning could be used (typical condition), or only counterfactual conditional reasoning would suffice (atypical condition). Difficulty was expected to go in that order. These tasks were administered in a within participants design so we could compare performance. Performance was expected to be high in four year olds where basic conditional reasoning alone was required as in the Slide task. Where counterfactual reasoning only could suffice, performance was expected to be low, the atypical condition of the Story task.

There are several additional key differences between the Slide and Story tasks that can also be expected to generate differences in performance. First of these is the availability of the relevant causal structure. For the Slide task the causal structure of the event sequence is concretely defined and visually observable in the apparatus: the ball can either fall to the left or to the right, and one of these will be the realist location. In the story task, however, there are three locations introduced that are physically independent of one another and with no inherent causal structure visually discernable. Unlike the causal path of the ball, the causal path of the story protagonist relies on the verbal information presented in the story itself. Therefore an abstract representation of the causal structure must be generated from which to reason successfully.

Secondly there is no intentional agent involvement in the Slide task. An inanimate object, the ball, drops down one of two chutes at chance. In the story task, however, the unseen intentions of the protagonist must be taken into account, a decision must be made about their trajectory that is no longer based on chance.

Finally, the counterfactual question being asked in both tasks is quite different. For the Slide task the counterfactual is “What if it had gone the other way, where would it be?” For the story task the much more complex “What if event x hadn’t happened, where would the protagonist be now?” For the Slide task there is not only linguistic markers as to the correct location i.e., “the other,” but there is no causal chain to alter and reconstruct given the very simple causal pathways involved. In the story task, however, no such linguistic support is given in the question and there must be a second contrasting representation generated that defines a new causal structure given the removal of a key causal event. No such removal or change is required to the representation of the Slide task’s structure.

We were not only looking for differences in the tasks, however, but also positive correlations between them. Our fifth and final aim then was to identify whether performance on the two types of task was related. Often in the literature counterfactual reasoning has been assessed using one measure only. We wished to have a more comprehensive and challenging assessment of CFR. Secondly we wanted to generate individual performance scores for each child’s CFR that we could then use to relate to their performance on our other cognitive measures. These analyses are reported in Chapter 4, sections 4.4.2 and 4.4.3 where the relationship between conditional reasoning, executive functions and mental state understanding are explored.

Given this thesis’ primary goal of assessing how children’s navigation between actual and possible worlds impacts upon their behavior, counterfactual reasoning is the natural place to start. Counterfactual reasoning demands children reason concurrently about what is, what was and what could be, and it is this skill that

goes on to drive many of our successful manipulations of the world to achieve our intentional goals.

## **2.3. Method**

### **2.3.1. Participants**

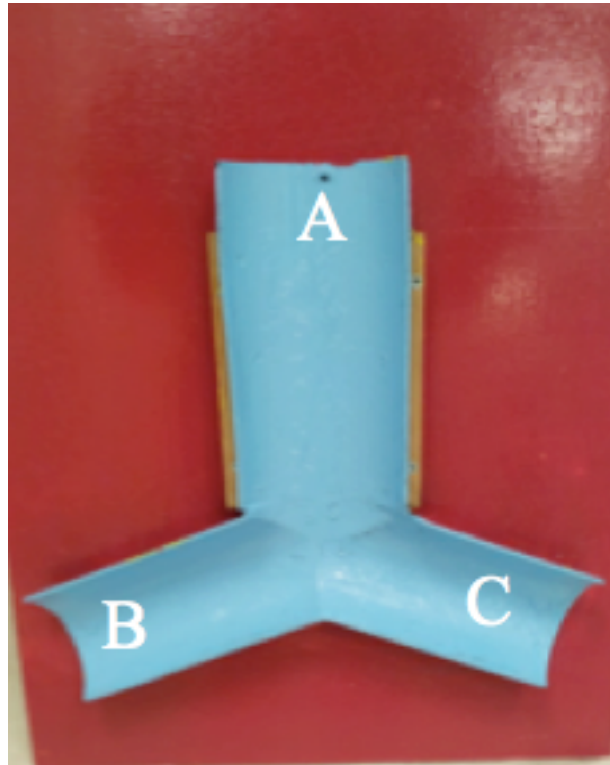
Fifty-five 4.4-year olds took part. N=28 of these had previously taken part in a longitudinal study from birth to 24-months named First Steps, (see Ellis-Davies, Sakkalou, Fowler, Hilbrink & Gattis, 2012). First Steps was set up to longitudinally assess various aspects of socio-cognitive development through the collection of experimental, observational and parent report data. Thirty-seven infants remained in the study at eighteen months and twenty-eight remained at fifty-two months. An additional twenty-seven participants were recruited for the fifty-two month phase to increase the sample size for the current study. These additional children were tested concurrently with the original sample participants. Participants were recruited through community groups. Children came from a range of socioeconomic and maternal education backgrounds. The majority were first language English speakers (N=45), with 10 bi-lingual or second language speakers. Testing took place over two visits on campus at Cardiff University. Participants were given £15 of shopping vouchers and a small gift bag on completion of the two visits. A total of five task groups were given to children in the course of these two testing sessions: counterfactual reasoning; mental state understanding; executive functioning; imitation and; language. The counterfactual reasoning tasks of interest in this chapter were both completed in the second testing session.

### **2.3.2. Design**

A within subjects design was utilized. All participants received all tasks and conditions. Tasks types were administered in the same order across participants. Children first completed all conditions of the stories tasks. Once completed the testing room was cleared and then the apparatus and testing procedure for the Slide task were introduced. This order was chosen for operational reasons. In the Story task twelve separate locations were used presented, three for each vignette. These scenes comprised of landscapes, small toy figures and props that took several minutes to assemble and as such were prepared prior to the participant's arrival at the University to avoid the disruption of data collection. The Slide task had no such set-up constraints as the required materials were smaller in number and could be stored easily behind testing curtains and produced when required with minimal set up being required.

### **2.3.3. Materials**

The location change Slide task was adapted from Beck, Robinson, Carroll and Apperly (2006) and Beck and Guthrie, (2011). Two slides were made from semi-circular tubes each 12cm in diameter. These slides were affixed to a wooden board with the inner surface of the tubes facing forward so the slide's contents could be seen at all times. The yellow slide was a straight tube with one opening at either end. The blue slide resembled an inverted Y shape with one opening at the top and two at the bottom (see Figure 2.4.). One yellow and one blue ball were used. These balls were only seen descending down their correspondingly colored slide, the blue ball always was shown going down the blue slide. The yellow ball was only shown going down the yellow slide.



*Figure 2.4.* Photo of the inverted Y shaped pipe-slide used in the location change slide task. A lightweight blue plastic ball was dropped down at point A and could travel down the slide to exit at either point B or point C. Children in the standard condition were asked “What if it had gone the other way, where would it be?” The correct answer would be to select the option B or C that contrasted with where the ball actually went during the model. An incorrect answer would be to select the same exit option as the model.

The Story task was adapted from Rafetseder and Perner (2010). Four vignettes about a Policeman, Doctor, Fireman and Teacher were used. For each of these stories small scenes were created, clearly depicting the three locations included in the dialogue (See Appendix A). The scenes were set up using cardboard illustrations and Playmobil toys. All actors in the stories were distinguishably different and were depicted by Playmobil figures. For the Doctor story the three locations were a



hospital, a park and a swimming pool. For the Fireman story the three locations were a fire station, a sitting room and a forest. For the Policeman story the three locations were a police station, a convenience store and a car park. Finally, for the Teacher story the locations were a classroom, a playground and a TV room. These scenes were set up on the four corners of a large black rug (150cm x 100cm). Each was separated from one another so as to distinguish the groups of three locations that belonged together. The child sat in the middle of the mat while the tester sat outside the mat and moved position for each story.

The vignettes all followed the same structure with three separate locations. “Before work a Doctor was sitting in the park enjoying the beautiful weather. She left the park to go to work in the hospital. When she got there he received an emergency phone call. Look, there has been an accident at the swimming pool and little Jacob has slipped and hurt himself. Jacob needs a Doctor. The Doctor lifts her emergency first aid case and walks from the hospital to the swimming pool to help Jacob.” A comprehensive script for all the stories can be found in Appendix A.

#### **2.3.4. Procedure**

Parents and children were invited to the on-campus lab for the two visits over the course of the fifty-two month testing phase. Each of these visits began with a short warm up time before the tasks were introduced in a separate room. The CFR tasks were administered during the second visit in a purpose built testing room with inbuilt digital audio and visual recording capabilities.

The Story task came first with the play scenes set up in four corners of the room on a large dark colored play mat covering the floor. Examples of the scenes can be found in Appendix A. For each vignette children were first asked to point out the various locations and were corrected if they identified them incorrectly. The four

story contents were counterbalanced with all four having both typical and atypical structure versions. Children either heard a typical-atypical-typical-atypical order, or an atypical-typical-atypical-typical order. Children were told one story at a time with each being followed immediately by two questions relating to that story, one control and one counterfactual question. In total then each participant was asked four control and four counterfactual questions. The question structure was the same for each of the four stories; control – “Where is the (protagonist) now?” and cognitive flexibility- “What if the (incident) hadn’t happened, where would the (protagonist) be now?” Once this procedure was finished children were asked to help the experimenter tidy up the story scenes and figures so as to make room for some new toys.

The two brightly coloured slides and balls were placed in the centre of the play mat. It was explained that the yellow ball could only go down the yellow slide, and the blue ball could only go down the blue slide. The yellow slide was straight with one opening at either end. The blue slide was an inverted-Y with one opening at the top and two at the bottom (See Appendix B). The yellow slide served as a filler item between some of the target trials. Trials were carried out in the same order for each participant. The four counterfactual questions were asked in a fixed alternating order (filler-standard-open-filler-standard-open). There were two target conditions both administered using the blue inverted Y shape slide, the standard condition and the open condition. After the experimenter let the ball roll down the blue slide the subject was asked “What if it had gone the other way, where would it be” (standard condition) or “This time it went that way, could it have gone anywhere else” (open). For the standard condition there was a very clear alternative answer, the chute that the ball had not gone down on that occasion. For the open condition the semantics of the question are quite different and could instead prompt a myriad of possible answers not

necessarily true or untrue in the context. The ball could have gone towards the ceiling having been abducted by aliens. Conversely the question could be interpreted in a deterministic way hence any other possibility would be ruled out.

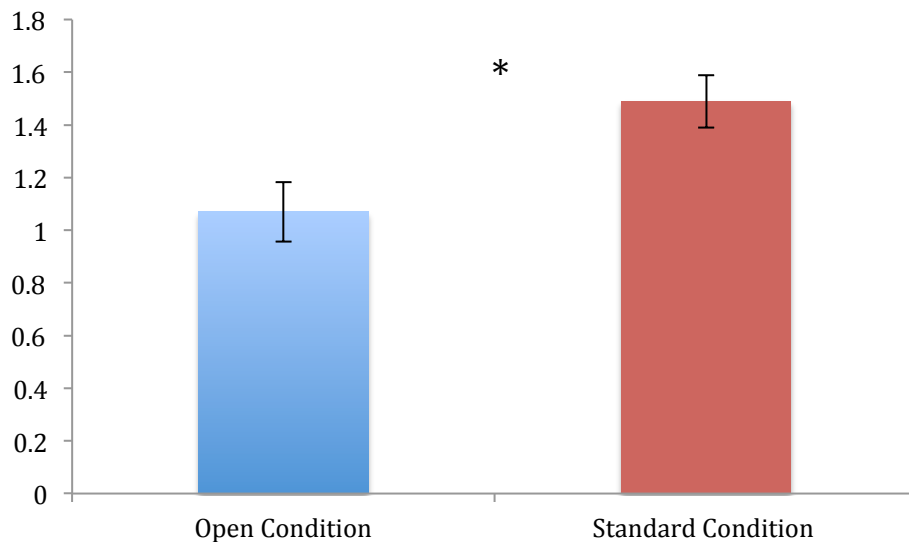
Answers were coded by a primary and secondary coder from the recordings after all sample data had been collected. The primary coder was the experimenter. The secondary coder was another PhD student who was blind to the hypotheses of the experiment. For both tasks coding began once the experimenter had finished asking the target question. For the Story task children's responses were coded when they either pointed to or verbalized a location in the story. If they did not respond the question was asked a second time. After two repetitions a forced choice option was given, (10% of sample). Responses were coded as incorrect if participants indicated either of the two remaining locations, one of those always being the realist location. Examples of the correct and incorrect answers can be found in Appendix A. Responses were also coded as incorrect if they indicated or referred to any other location not presented in course of the story.

For the Slide task children scored one point for each target trial if they indicated the exit of the slide not taken during the model. For each condition this gave a maximum possible score of two. Responses were coded as incorrect if the participant indicated the exit the ball had taken (realist answer), the entrance to the slide, to any other location not presented during the model or any other incomprehensible answer. Inter-rater reliability for both reasoning tasks was 100%.

## **2.4. Results**

### **2.4.1. Slide Task**

Each participant was asked four target questions, two open and two standard. Their final coded answers for each target question were used in this analysis giving a score of 0-2 for both the standard and open conditions. The mean score for the open condition=1.1. The mean scores for the standard condition=1.5. At the group level Standard Slide performance exceeded chance set at 1 with the degrees of freedom equally N-1,  $t(54)=4.91$ ,  $p<.001$ . 64% of our participants answered both Standard slide questions correctly. In the Open slide condition performance was not different from chance  $t(54)=.65$ ,  $p=.5$ , with only 38% of children answering both of the Open questions correctly. Interestingly, however, only 8 children (15%) got both Standard questions wrong while 17 children (31%) got both Open questions wrong. Across participants performance on the Open question is at chance level (See Table 2.1.) and is significantly worse than Standard performance,  $t(53)=3.2$ ,  $p=.002$  (see Figure 2.5.). Performance on the two conditions does correlate positively and just reaches significance  $r=.27$ ,  $p=.046$ .



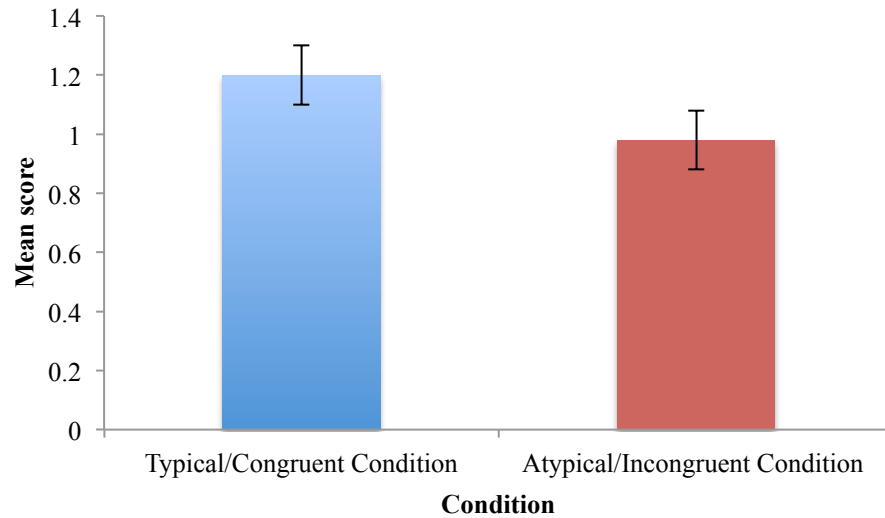
*Figure 2.5.* Mean scores for the Open and Standard conditions of the slide task. Error bars represent standard errors.

Table 2.1. *Total number of participants who answered 0, 1 or 2 reasoning questions correctly for both of the questions for the slide task.*

| Condition | Number Correct |    |    |
|-----------|----------------|----|----|
|           | 0              | 1  | 2  |
| Standard  | 8              | 12 | 35 |
| Open      | 17             | 17 | 21 |

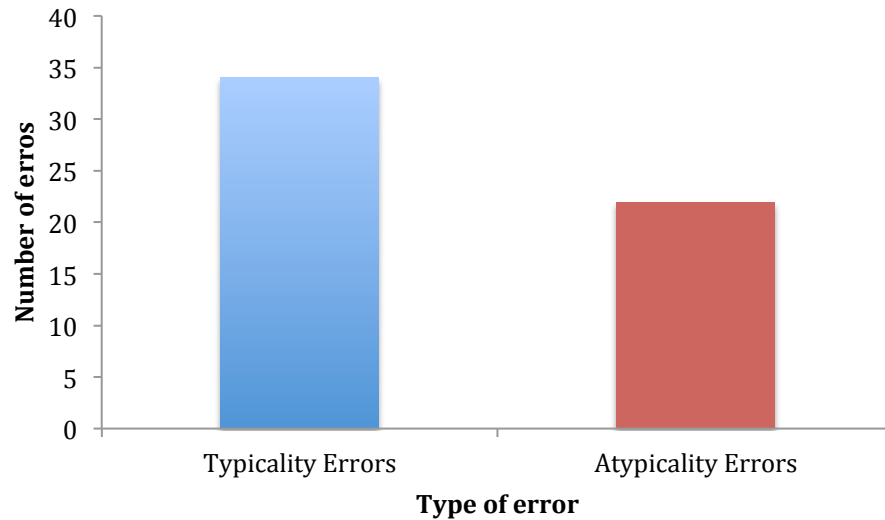
#### **2.4.2. Story Task**

Total story performance was the combined score for the two atypical and two typical conditions (0-4). A one sample t-test was carried out for total story performance against chance levels set at 2. Scores did not differ significantly from this chance level,  $t(54)=.88$ ,  $p=.38$ . Performance in the typical condition ( $M=1.2$ ) exceeded that of the atypical condition ( $M=.98$ ). At the group level, performance in both conditions was not significantly different from chance. However, in the case of typical condition performance was approaching significance against the chance level,  $t(54)=1.9$ ,  $p=.062$ . Even when children who had committed realist errors were omitted from the analysis, in order to examine only those children who were able to inhibit realist responses, the difference between the two conditions did not reach significance.



*Figure 2.6.* Mean performances on the Typical and Atypical conditions of the Story task. Error bars represent standard error.

Although the means for our Story tasks varied to those reported in Rafetseder and Perner (2010), the pattern of performance was similar with performance in the typical condition exceeding performance in the atypical condition (Figure 2.6.). However, in our sample, the difference between performance in these two conditions, typical and atypical was not significantly different from chance as measured by a paired samples t-test,  $t(54)=1.9$ ,  $p=.08$ . Mean performance across the conditions correlated positively with one another,  $r=.3$ ,  $p=.027$ . There were more typicality error types ( $M=.8$ ), than atypicality error types ( $M=.5$ ) (see Figure 2.7.). However, this difference was only approaching significance,  $t(54)=1.8$ ,  $p=.07$ .



*Figure 2.7.* Column chart showing number of typical and atypical errors recorded during the stories task.

### **2.4.3. Relationship Between Tasks**

The open condition proved difficult for children and as Beck et al. (2006) pointed out there was no single correct answer available for children. In the open question children could theoretically have identified an infinite number of other possibilities as to where the ball could have gone, this was reflected in some children pointing to the ceiling or elsewhere in the room when asked the open question. Additionally the form of these questions was designed to elicit other possibilities from the children and not specifically identify one counterfactual outcome, this is despite asking the children to mentally represent another course that at one point could have, but could no longer have been taken. Given that we wished to compare performance on two specific types of causal reasoning in which a defined correct and incorrect answer could be identified it was deemed prudent to remove the open condition results from this phase of the analysis as a direct comparison could no longer be

drawn. Due to the performance distribution on the open condition of the Slide task it was decided to remove it from the comparison analyses with the story task.

Mean performance in the standard condition of the Slide task ( $M=1.5$ ), exceeded that of the Story Task as analyzed using a paired sample test (adjusted  $M=1.1$ ), (Figure 2.8.). This difference is significant,  $t(54)=-.42$ ,  $p<.001$ .

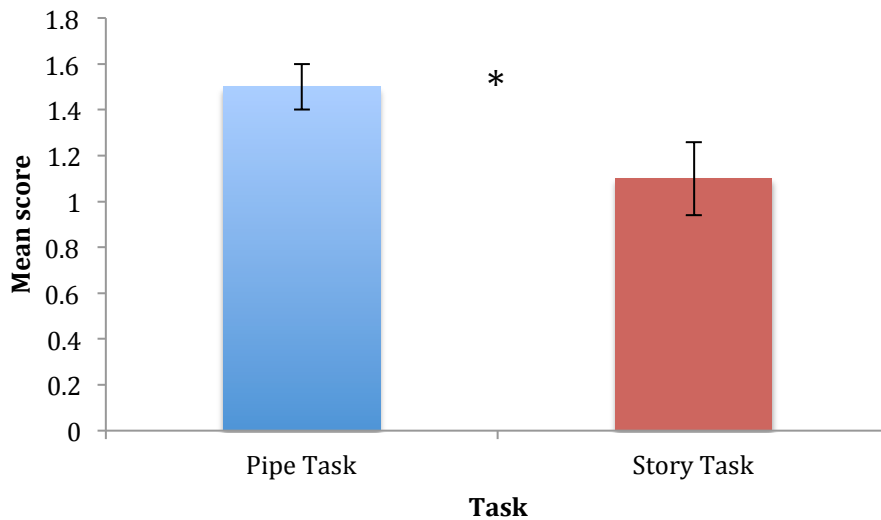


Figure 2.8. Means performance on the Slide and Story tasks. Error bars represent standard error. \* Represents a significant difference between bars  $p<.05$ .

Table 2.2. shows the correlations between the conditions of our two tasks. Across task significant moderate positive correlations were demonstrated between the Standard condition and the Typical condition, in addition to the Atypical condition (see Table 2.2.). As expected the Standard condition was significantly easier than both the Typical ( $t(54)=-2.4$ ,  $p<.05$ ), and Atypical conditions of the story task ( $t(54)=4.5$ ,  $p=.001$ ).

Table 2.2. *Pearson's correlations between the stories task and standard condition of the Slide task.*



|          | Typical | Atypical | Standard |
|----------|---------|----------|----------|
| Typical  | 1       | .298*    | .275*    |
| Atypical |         | 1        | .36**    |
| Standard |         |          | 1        |

\* p<.05                      \*\*p<.01

In addition, following a regression analysis of the error types made in the Stories task, performance on the Standard condition of the Slide task emerged as a significant negative predictor of the number of Realist errors across the Story Task  $R^2=.25$ ,  $F(1,53)=17.7$ ,  $p<.001$ . Realist errors are those in which the subject continues to place the protagonist at the same end location as in the initial dialogue, despite the false antecedent being introduced. Participants making fewer realist errors in the standard Slide task also generated fewer realist errors in the Story task. When realist errors are partialled out the correlation between tasks disappear.

There were more Typical errors =34 (incorrect selection of the typical location), than Atypical errors =22 (incorrect selection of the atypical location). 62% of children made one or more typicality error, whilst only 40% made one or more atypicality error. However, this difference just failed to reach significance  $t(54)=1.8$ ,  $p=.07$ .

Finally we wanted to establish whether the order of performance seen, with the Slide task, standard condition being passed first and the story task second was statistically significant, i.e., that passing the Slide task was a precursor to passing Story task. The Slide task was significantly easier than the story task and when performance was cross-tabulated it was evident that children very rarely passed the Story tasks having failed the Slide task. Success on the standard Slide task was a pre-requisite for success in the story task, Pearson  $\chi^2=11.73$ ,  $df=1$ ,  $p<.001$ .

## 2.5. Discussion

Our first aim was investigate how our results compared with those of previous authors using the same tasks (Beck et al., 2006; Beck & Guthrie, 2011; Rafetseder & Perner, 2010). The location change counterfactual Slide task's standard condition saw 64% of participants answer both test items correctly with a mean of 1.5. It is difficult to make a direct comparison with the data reported in Beck and Guthrie (2011) as their sample groups had a mean age of 3.9 years for the younger group and 5.0 for the older. However, their mean scores for the standard condition were 1.3 and 1.6 respectively so as expected our sample of mean age 4.4 scoring 1.5 fell between these two. Also as expected, the 64% of children who answered both standard questions correctly also fell between the two groups reported in Beck and Guthrie (2011), 49% for the younger group and 74% for the older group. In the open condition our samples mean score of 1.1 again fell between the two means reported in Beck and Guthrie of 1.2 for the younger group and 0.8 for the older (Beck & Guthrie, 2011). As with our sample, both of their groups were better than chance in the standard condition but not the open condition. We too identified the previously reported pattern of unpredictable responding in the Open condition of the Slide task (Beck & Guthrie, 2011). Due to the erratic performance in this condition for our sample we chose to withdraw the open condition results from future planned analyses. The structure of the question was not being consistently interpreted as a counterfactual, or even a basic conditional reasoning problem given the chance responding reported. The patterns of performance, and how they relate to other measures should be investigated more comprehensively in future works but are not suitable to be included here specifically as an index of counterfactual reasoning.

Story task performance for the atypical condition in our study reached almost 50% with a mean of .98: children were performing at chance levels. The typical

condition was slightly more successful with a mean of 1.2. These scores are markedly higher than those for the 4.4 year olds reported in Rafetseder and Perner (2010). However, only 13 children of this age were tested in the original study compared to 55 in the current study.

Rafetseder and Perner (2010) further reported that the majority of the four year-olds in their study picked out the realist location. This response was fairly rare in our sample with only 13 of the 55 children generating one or more realist errors. Only 30 realist errors were recorded out of 220 responses (14%). We must then disagree with their conclusion that four-year-old reasoners preferentially pick the real location in this task.

Rafetseder and Perner also reported that younger three-year-old participants answered in a basic conditional reasoning congruent way, that is they preferentially made typicality errors in the atypical condition but answered correctly in the typical condition. This younger group of children made fewer realist errors than the older group. The authors suggested that the younger children reason in a very basic conditional reasoning structure where they focus on the problem in isolation and answer in a semantically congruent way. The older children, however, are suggested to make more realist errors as they understand that the question relates to where the protagonist is now, but fail to integrate all the relevant information about said protagonist in relation to the counterfactual antecedent so answer with reality. Only when children become even older (Rafetseder & Perner suggest) do they begin to reject the realist answer but still not have integrated information successfully enough to reliably pick the correct location, when it is semantically incongruent. In our sample only 7% of children answer as realists across the board, 7% answer systematically in a basic conditional way and 13% answer in a consistent

counterfactual manner. That leaves the overwhelming majority of 71% answering using a mix of strategies. We must conclude then that by fifty-two months of age children do not answer as realists when given counterfactual reasoning problems, however, they do not appear to have adopted a consistent strategy when it comes to reasoning about counterfactual situations.

Also in contrast to the findings reported by Rafetseder and Perner we did not find a significant difference in performance between the typical and atypical condition. Although showing the same trend with performance slightly better in the typical condition in which the correct answer was also the semantically congruent option, there was not enough of a typicality bias present to statistically alter the group response patterns in our sample. This result was initially surprising due to our use of the same paradigm as their original study. However, Rafetseder and Perner (2010) report this specific between condition difference only in reference to their sample as a whole. Their complete sample numbered 133, more than double that of our own. In addition the age range of this sample was greater with children aged 3;2 to 6;10 years taking part. Our sample consisted wholly of children aged 4;4 years. Rafetseder and Perner's 4;4 year sample contained only 13 individuals and no typical versus atypical condition difference was reported for this portion of the sample alone. Visually inspecting Figure 1., of their results suggests that it is at this age, 4;4 years, at which typicality begins playing a more dominant role with performance in the atypical condition being poorer to a greater extent in the latter age groups than the former. Perhaps then had our sample been several months older, or larger, we may have replicated their initial finding.

Our final aim for this chapter then was to determine whether performance on the two types of CFR task was related. The Slide task represented a more established

form of counterfactual in which the location of an object would have been different given a counterfactual antecedent. The story task provided a more exigent measure of CFR that challenged children to integrate elements of reality up to the point where they diverged from the counterfactual statement. The story task could only be successfully navigated using CFR and not a more basic conditional reasoning process that would have been equally successfully regardless of the particulars of the counterfactual situation. Despite the difference in difficulty between the two types of task there was a strong correlation between performance on the standard condition of the Slide task and performance on the stories task. This appeared to be due to the pattern of realist errors observed across tasks. It seems then although the demands are very different for the tasks, and performance on one is significantly superior to the other, there is still a strong positive relationship between them.

### **2.5.1. Limitations**

The inclusion of the Open condition in the Slide task proved problematic. Participants answered at chance levels and the structure of the question was, as the name suggests, very open. Children were asked “This time the ball went that way. Could it have gone anywhere else?” No counterfactual change has been made and there was no change to the causal structure suggested. These elements combined led to the question being ambiguous and not a sensitive test for the kind of reasoning ability we were interested in. We wanted to see how well children could reason about causal structures in a counterfactual way. The other questions all suggested a counterfactual antecedent of sorts and as such provided clues for children to reason accordingly. Beck and Guthrie (2011) also reported difficulties with this question finding responses unpredictable.

With regard to the Story task in the context of this thesis it may have been

beneficial to have included a different exigent measure capable of identifying true counterfactual reasoning. The example of the Sweets story given in Figure 2.2 for example may have served as a useful template for us to generate similarly structured questions in order to give participants a CFR score. The Sweets story example, unlike our chosen story task, did not include contrasting conditions in which prior knowledge could have impacted upon performance. Although the comparison between the typical and atypical condition was an interesting one it may have benefited us to have included additional examples of true counterfactual reasoning. Having multiple questions, all reliant on true CFR would have afforded us the ability to identify children who were consistently able to reason counterfactually

### **2.5.2. Conclusions**

At age four most children are able to reason about an alternative outcome to a past event when; the options are limited, the causal structure is visible and the question is unambiguous. Being given all the causal information necessary for successful reasoning in a concrete and structured way allows children to reason successfully about how a change in a past world could result in a change to a counterfactual world.

When there are several possible counterfactual options, however, and when the causal structure of the counterfactual world is not so concretely visible children's reasoning performance is reduced. Having a less structured and more complex problem highlights the immaturity of young children's counterfactual reasoning ability. Children are very inconsistent with how they reason about counterfactual worlds of this sort, worlds that involve other intentional agents and require the integration of multiple pieces of information from reality alongside new pieces of information held within the counterfactual statement.

Despite there being distinct differences in performance between these two forms of counter to fact reasoning task, there is a strong positive relationship between them. Being able to reason in a more structured problem space predicts reasoning ability in a less constrained and more challenging counterfactual situation. Such results suggest that conditional reasoning ability is fairly consistent across tasks. Children who are able to reason more effectively about the consequence of a counterfactual antecedent in a highly constrained, concrete situation are also more likely to reasoning effectively about the significantly more abstract altered movements of a person, in a more complex causal situation, given a counterfactual antecedent. Children show contiguity in understanding how an event happening in a past reality could impact upon a future reality differing only by virtue of the altered element. In the following chapters we shall investigate other expressions of causal and conditional reasoning and how effectively they can operate at fifty-two months.

## **Chapter 3. Executive Functions and the Selection of One Response Over All**

### **Others**

#### **3.1. Abstract**

In order to make valid inferences about a certain problem children must have specific content knowledge about the problem itself in addition to the domain general processes necessary to reason about and execute any appropriate responses. Where conditional reasoning tasks require the utilization of knowledge obtained, executive functioning tasks require the utilization of knowledge provided. Executive functioning tasks require no inductive reasoning as the conditional rules and resultant actions are distinctly stated. Conditional reasoning tasks require inferences to be drawn based on the information provided or obtained through causal reasoning. Executive functioning tasks therefore have two inherently informative aspects to their construction in the context of this thesis. Executive functions are not only required for tasks measuring children's reasoning about the actual versus the possible but they are themselves measures of actual versus possible responses. In this short chapter I will introduce several of the accounts of executive functioning most dominant in contemporary literature. Next I shall discuss briefly executive functioning's development and measurement techniques in young children. Finally I will present data collected using two standardized executive functioning tasks (the Flanker and DCCS from the NIH Toolbox, Version 1) with the same sample of children reported in Chapter 2. Our aim was to generate individual executive functioning scores that could be compared with our other collected measures of cognitive and communicative development. In addition we will go on to assess the relationship between executive functions, counterfactual reasoning and mental state understanding in Chapter 4.



### 3.2. Introduction

In the previous chapter I discussed the ability to reason about specific changes to the world. Executive functioning (EF) refers to the higher order processes required for all effortful, purposeful and goal directed thought and action, (e.g. Carlson, 2005; Gioia, Isquith & Guy, 2001). The main abilities that fall under this umbrella term, inhibitory control, cognitive flexibility and working memory have been greatly studied in both developmental and adult data sets (see Diamond, 2013 for a review). In essence executive functions allow us to represent information, move between these different representations and, plan and execute responses based upon them. They constitute the abilities required to think abstractly and manipulate ideas. Executive functions are crucial for the discovery and retention of new information about the world (Anderson, 1998). As such they have been shown to support children's development in a diverse range of areas including social interactions, cognition, mental health, school readiness, school success and beyond into adulthood (e.g. Anderson, 2002; Blair & Razza, 2007; Diamond, 2005). Verbal ability is also often strongly related to EF performance (Carlson & Moses, 2001; Hughes, 1998). Executive functioning abilities develop across childhood and into adolescence. Age is a significant factor in performance for many of the tasks used to investigate EF (Carlson, 2005). Additionally performance on different EF tasks can mature at different rates (Walsh, Pennington & Groisser, 1991).

For the purpose of this thesis we choose to include measures of inhibitory control and cognitive flexibility in our test battery of executive functioning. These two measures of executive functioning were chosen due to their structure being inherently complimentary to our interest in the actual versus the possible. Both of these executive functions are measured using paradigms that challenge participants to

select one of two possible responses based on conditional rules referring to an actual presentation. Both the Flanker and DCCS challenge children to reason about a possible world given an actual world.

### **3.2.1. What are Executive Functions?**

There are many more or less altered descriptions of what executive functions do. Anderson (1998) describes the ability to shift between response sets, learn from mistakes, devise alternative strategies, divide attention and process multiple sources of information. Luria (1973) determines EF to be responsible for the synthesis of external stimuli, preparation for action, formation of programs, allowing actions to take place and verifying the success of said actions. In a recent review Diamond (2013) refers to executive functions as a family of top-down cognitive processes deployed when reliance on more automatic processing would be inappropriate. She deems these processes as necessary when playing with or manipulating ideas or, resisting temptation. What these, and other (e.g. Miller & Cohen, 2001; Zelazo, Müller, Frye, Marcovitch, Argitis, Boseovski, ... & Carlson, 2003) descriptions converge on is the idea that executive functions are at the interface between the intention and planning to act and, the subsequent actions and their appraisal: executive functions facilitate the movement between reality, the actual world to a future, possible world. As discussed in Chapter 2 the ability of children to incorporate their' past experiences and knowledge to their future decision making has huge benefits. To achieve an optimal possible state of affairs it is necessary to react to reality in the most successful way possible. This requires information about the past and prior knowledge being combined in such a way as to initiate the actions and behaviors most likely to bring about desirable results. Mature intentional behavior in humans goes beyond a simple stimulus response relationship. We may usually get up

to answer the telephone when it rings, but if it happens during dinner, at the same time as a sales call the day previously, we may instead ignore the ringing as if it is important a genuine caller will leave a voicemail. Information from numerous and/or diverse sources can be reflected upon in order to come to a reasoned decision, as opposed to uniformly deploying the most salient or familiar type of behavior, regardless of the specific attributes of a given physical or social situation. In this way responses can be innovative and appropriate in a way that goes beyond an automatic processing and appraisal of the world. Executive functions are what facilitate these possibilities through allowing reflection and inhibition of the most common response. Being able to flexibly respond to the actual world to move towards a better possible one is the hallmark of strong EF. Executive functions facilitate us recruiting the pertinent prior knowledge related to current reality; they allow us to abstractly manipulate, appraise and select possible courses of action or behaviors and they work to ensure these optimum outcomes are achieved. In this chapter we discuss the tasks used to assess EF, and expand upon how the successful deployment of EF is the juncture at which children move from the current state of affairs, to a new and better world.

EFs can be identified in terms of what they accomplish more so than how they accomplish it (Zelazo et al., 2003). Tasks designed solely to assess EF rely on novelty and complexity in order to judge how adept children are at executing “if-then” rules in an experimental context. Novelty of both stimulus and task limits any practice effects influencing results. EF tasks are interested in the general processes not knowledge content of reasoning about the actual versus the possible. CFR and MSU tasks in contrast are interested in the process of reasoning as applied to counterfactual knowledge and mental state knowledge. Zelazo et al. (2003) identified four main

strands of account that focus on different aspects and proposed mechanisms of executive function that are briefly described below: Complexity theories, working memory, inhibitory control and re-description accounts.

Complexity theories such as the Cognitive Complexity and Control theory (CCC), (e.g. Zelazo & Frye, 1998; Zelazo et al., 2003) suggest that with age and experience children are able to represent and reflect on cognitive rules in an increasingly hierarchical fashion. Initially children struggle to respond appropriately to if-then type statements if there is some form of conflict between rules, for example in a card sorting tasks where test cards differ on two dimensions, shape and color. When cards are initially sorted according to shape, then participants are asked to sort the same cards by color, if there is no integration between the two-rule dichotomies, participants will continue to sort along the initially implemented dimension. They must appreciate that one card can be either red or blue, whilst also being either a square or a triangle. In addition they must be able to selectively attend to the relevant dimension in response to the explicit if-then rule presented by an experimenter. According to CCC, only when different rules can be incorporated in a higher order structure are they able to flexibly move between behaviors based on these rules, allowing correct pre and post switch sorting to occur. CCC presumes that children are able to represent a nested model of the target stimuli, in this case a card, where two sets of conflicting rules can reside concurrently: one set distinguishing the cards on the dimension of color, and one set distinguishing on the basis of shape. Prior to such hierarchical integration then children will tend to perseverate with a certain pattern of responding. This is described as being result of their inability to efficiently build and use a suitable problem representation and switch attention to a new dimensional feature of the stimuli in question. With increasing levels of task complexity then

performance will become harder due to the rules needing to be represented in this hierarchical fashion, from which behavioral inferences must then be extracted.

In contrast, memory accounts (e.g. Gordon & Olson, 1998; Morton & Munakata, 2002) claim it is not the complexity of rules that dictate performance on EF tasks. Instead, it is the number and/or strength of rules to be held in memory and acted upon that determine performance. As memory capacity increases or memory processing mechanisms mature, an increasing number of rules can be comprehended and applied. Evidence in opposition to such theories has been presented by Zelazo et al. (2003). Zelazo and colleagues contrasted performance on a task utilizing four rules and two rules and no performance difference was evidenced. More rules do not necessarily result in poorer performance, it is the dimensionality of the rules that impact upon complexity and performance (Zelazo et al., 2003).

Accounts that focus on the development of inhibitory processes, specifically in EF performance, deem the ability of a mechanism to suppress behavior as key (e.g. Carlson, Moses & Hix, 1998; Luria, 1973). Other related accounts postulate inhibitory control, alongside memory developments as the mechanism responsible to EF performance (e.g. Gerstadt, Hong & Diamond, 1991; Roberts & Pennington, 1999). Such accounts blame poor EF performance on an inability to overcome some pre-potent response when challenged to behave in accordance with some other explicitly presented if-then rule. Pre-potent responses can include semantically congruent responses such as pressing a button on the right when a right arrow is presented. Pre-potent responses can also be generated when participants are asked to follow some if-then rule regarding a stimulus, then a new rule is presented relating to this same stimulus. For example being asked to sort blue/red pictures of squares/circles by color, and then by shape. In young children 2-5 years of age, sorting based on color

will become a transient pre-potent response due to having applied this rule repeatedly in the recent past. Inhibition failures, however, need not only occur at the point of response control. Kirkham, Cruess & Diamond (2003) suggest that an inability to inhibit a certain representation of the problem will hinder performance. Children can understand new to-be-applied rules but fail to disengage their attention from previously applied or otherwise more salient rules when executing behavior. According to Kirkham et al. (2003) children find it challenging to switch their attentional mindset despite being able to identify the relevant dimensions. Kirkham et al. (2003) have labeled this phenomenon as “attentional inertia” and characterize it as a source of cognitive rigidity in contrast to the flexibility that is characteristic of mature executive functioning. The authors also highlight the difficulty children face in redirecting their attention to a newly relevant dimension. Children’s representation of their world, and the tasks they are undertaking within it seem to become more flexible with age as children become better able to switch between different representations and subsequently alter their behaviors.

Redescription accounts (e.g. Perner, Stummer, Sprung & Doherty, 2002) suggest that children’s difficulty in switching between rules is based upon an inability to re-describe a single stimulus in terms of a different dimension. Red objects are red objects, blue are blue, when asked to think of these colored objects along some other dimension such as shape (squares or circles), there is some difficulty in the re-description process. Such accounts have been influenced by theories about; children’s striking egocentrism, the development of their ability to represent multiple perspectives of a given scene and, the idea of mutual exclusivity (Markman, 1991; Piaget & Inhelder, 1948; 1956). The authors put forward the hypothesis that general developments in understanding different perspectives or representations around this

age positively impact upon other specific reasoning abilities. When, at around four years, children begin to understand multiple perspectives of a single object or event, they begin to be able to base their behavior on this understanding (Perner & Lang, 2002).

These accounts differ in their focus and emphases regarding what elements ultimately determine EF performance; rule complexity, working memory capacity, inhibitory control maturity or re-description/perspective taking ability. As such they do not present mutually exclusive theories. Rather, elements of each are qualitatively informative with regards the kind of abilities and processes necessary for successful EF task performance.

### **3.2.2 Measuring EF and its Development in the Pre-School Years**

The aforementioned genres of EF account, although differing in perspective and emphasis, assume that executive functioning encompasses the abilities to inhibit or suppress some behaviors based on some selected dimensional representation of an “if-then rule”. One of the hallmarks of EF failure then is an inflexibility of responding when faced with conflicting possibilities (Zelazo et al., 2003). This inflexibility may result from an underlying failure to inhibit a response or, a failure to inhibit or generate an appropriate representation of the task. These types of inflexibility may be a result of restricted WM capacity, immature IC skill, more complex rules than can currently be supported or an inability/difficulty in changing perspective. We chose to include two widely used measures of executive function. The Flanker task was designed to index response control inhibition whilst the Dimensional Card Change Sort (DCCS) was designed to tap the flexibility with which participants switch between multiple representations of the same stimuli. Specifically we chose to use newly developed, computerized versions of these tasks that comprised part of the

larger NIH Toolbox Version 1 (Zelazo, Anderson, Richler, Wallner & Allen, Beaumont, & Weintraub, 2013).

Inhibitory response control tasks such as the Flanker (IC) challenge participants to select a response based on a central stimulus that is incongruent to the other stimuli flanking it. The Flanker task for inhibitory control (e.g. Eriksen & Eriksen, 1974; Zelazo et al., 2003) requires selective attention; that is attending to and responding based upon a central stimulus regardless of whether the surrounding stimuli are congruent or incongruent. A classic example is being instructed to press the left key when a left pointing arrow is presented centrally, flanked by right pointing arrows on either side. Flanker tasks typically rely on applying only one rule dimension: direction. It allows us to reject a strong internal or external pre-potent response (Diamond, 2013). Inhibitory control begins to be seen in the first year of life, however, perseveration errors can be seen in tasks measuring IC through the life span. Typically perseveration errors, those resulting from a failure of inhibitory processes, decline during adolescence (Chelune & Baer, 1986.) Individual differences in both speed and accuracy on simple inhibitory control tasks are still developing through late childhood and into adolescence (Diamond & Taylor, 1996; Luna, 2009). Deploying IC enables an individual to resist the conditioned responses elicited by our environments (Diamond, 2013). Without IC we would become enslaved to the repetition of the same responses in any given situation without the ability to modify our actions in pursuit of different goals and in line with our intentions. Such inhibitory control failures would result in continual realist errors during reasoning and the absence of any innovation in behavior in the pursuit of possible goals divergent from the actual world. Inhibitory control specifically has been related to positive behaviors ranging from better concentration during childhood to reduced likelihood for anti-



social behaviors and ill health in later life (Moffitt, Arseneault, Belsky, Dickson, Hancox, et al., 2011).

Cognitive flexibility (CF) refers to an individual's ability to rapidly switch between response sets based on a change in some explicitly presented abstract rule or condition. It supports an individual's thinking about contrasting semantic, spatial or interpersonal perspectives: it allows us to change the way in which we are thinking (Diamond, 2013). Crucially cognitive flexibility demands we move between different representations of the same stimulus. This can be as straightforward as representing a simple object in terms of its shape having previously been interested only in its color. A classic task assessing CF involves sorting post-switch cards based on these two simple and directly observable dimensions. Participants must sort by color for example, having previously been sorting pre-switch cards based on shape (Milner, 1964). In adult versions of this task the sorting rule will change unannounced and subjects must find and adhere to a new rule. In childhood versions the rule is enforced by a voice across trials (e.g. Frye & Zelazo, 1998). Young children often persevere with sorting according to one rule, despite being competent in identifying the target features and having been able to sort them independently in a previous block (e.g. Zelazo, Frye & Rapus 1996). To succeed children must inhibit perseveration behaviors and select a new representation that supports correct sorting behavior. This contrasts with inhibitory control in that two or potentially more representations are competing. With inhibitory control tasks generally the same dimension is being attended to, it is the resultant responses that are competing: congruent and incongruent.

Such abilities emerge between 3-4 years of age, but show great improvements between ages 7 and 9 (Anderson, Anderson, Northam & Taylor, 2000). CF has been

documented as appearing and maturing later than IC (Diamond, 2013). Often tasks measuring CF also rely heavily on the inhibitory processes discussed previously in addition to working memory. Not only do children have to represent multiple contrasting rules, they must also inhibit the recently primed rule in favor of a negatively primed alternative (Zelazo et al., 2003). Cognitive control in particular has been related to creativity (Zabelina & Robinson, 2010).

Despite their different procedural demands, both types of EF are typically assessed through tasks requiring one response to be selected and another rejected. For inhibitory control the one to be rejected is highly salient due to some semantic relationship. For cognitive flexibility one response adheres to the current or most recent abstract sorting rule based on some representation of the target while one does not, a contrasting representation of the same stimuli must be attended to. In the most basic sense then both IC and CF measure the ability of an individual to not only differentiate and comprehend two possible responses, but to correctly select one given an explicitly imposed rule. Reasoning tasks then, where children are asked to select a response based on their inferences about some content information and the pertinent rules implicit there-within, rely heavily on IC and CF. With EF tasks the process itself is interesting as the rules are externally generated. Free from the constraints of relying on ones own causal knowledge children should be better able to carry out selection tasks with a greater deal of proficiency. In addition children should be performing on a more level playing field, so to speak, given that no prior knowledge is required and the to be applied rules are as a result of the same magnitude for all participants. In content based reasoning tasks the type of information being reasoned about is equally as interesting and influential as the processes involved due to the requirement that rules be extracted from the representations independently.

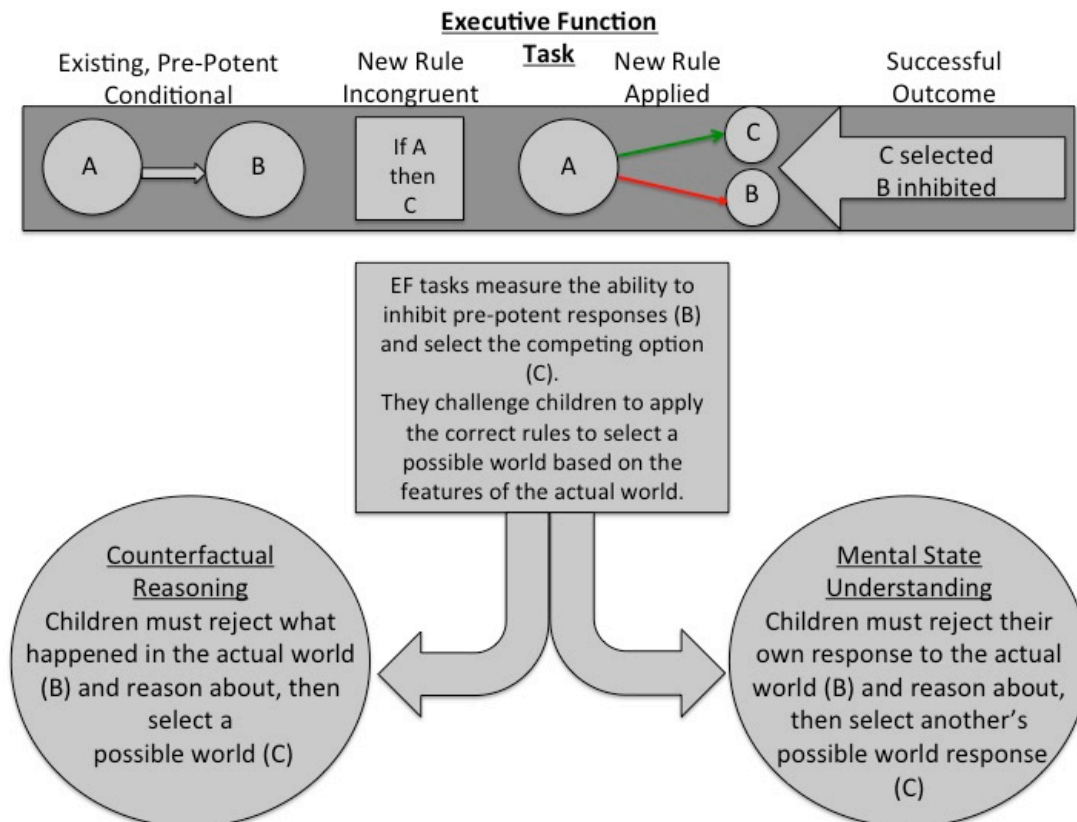
When children engage in a Flanker or DCCS tasks they sit at the interface between understanding a rule, intending to act based upon it and executing that action. In addition they are in a sense at the same interface between reality and counter-reality seen in CFR and MSU tasks. If situation A happens in reality they must then execute action A. They must infer their future behavior based on the characteristics of the current reality. In EF tasks the current reality is free from many of the constraints of prior knowledge. Children are being challenged in the target conditions to act incongruently with reality: they must change their perspective in order to make the correct inference about their future behavior. For the Flanker they must answer in terms of stimulus A = response not A. For the DCCS post switch phase they must inhibit the positively primed pre-switch sorting dimension in favor of the negatively primed post-switch dimension. For both Flanker and DCCS they must inhibit either semantically or procedurally pre-potent responses in favor of the response that satisfies the if-then rule currently in place. In addition the DCCS, like CFR and MSU requires the reasoner to switch their attentional processing to a different representation of the same problem. The DCCS requires participants to switch their attention from one representation of a card to another. CFR requires participants switch their attention from an actual to a once possible representation of a sequence of events. MSU requires participants to switch their attention from their own representation of the world to another's potential representation of that same world.

There are manipulations to EF tasks that have been shown to influence performance either negatively or positively. Labelling, demonstrating or otherwise scaffolding post switch rules in the DCCS increases performance in even the worst 3-4 year-old participants (Towse, Redbond, Houston-Price & Cook, 2000; Kirkham et al., 2003). When the salience of a newly presented rule is reinforced attentionally it is

more readily adopted in planning and executing behaviors. Leaving pre-switch cards upwards, so serving as a reminder of pre-switch rules, works in the contrasting manner and lowers post-switch performance in this same demographic (Kirkham et al., 2003). However, in children who select the wrong options themselves, there are still low levels of error detection when asked to evaluate the selections made by a puppet (Jacques, Zelazo, Kirkham & Semcesen, 1999). This contrasts with evidence that children have the required rule knowledge, yet it dissociates from actual rule use (e.g. Frye, Zelazo & Palfai, 1995). It seems not that young, 3-4 year-old children are unable to inhibit some pre-potent response, but more that they are for some reason unable to integrate their knowledge with their planned behaviors; they are not attending to the correct features when planning and executing their responses. By age 5 children are becoming much more competent in engineering their future responses to align with their knowledge (Frye et al., 1995; Jacques et al., 1999).

### **3.2.3. Executive Functions and Other Cognitive Abilities**

Executive function tasks challenge children to infer correct future behaviors based upon the application of explicit exogenous rules related to features of current reality: If reality A then behavior C (where B represents the pre-potent option). Counterfactual reasoning tasks challenge children to infer a counterfactual state of affairs based upon the application of causal rules grounded in current reality but applied to an alternative world where some alteration has taken place: If not B then state of affairs C. Finally mental state understanding tasks challenge children to infer mental states based upon the application of mental state knowledge and rules grounded in current reality but applied to an alternative world as experienced by some other agent: If not my mental state B then their mental state C (see Figure 3.1.).



*Figure 3.1.* Visualization of Common Actual vs. Possible World Demands Made on Children During Executive Function Tasks, Counterfactual Reasoning Tasks and Mental State Understanding Task.

These three types of tasks share the need to infer, based on current reality, some feature of a future, counterfactual or mental state reality (see Figure 3.1.). All three are subject to tension between two or more conflicting options, representing different representations of the world. What will you do if B? What would the world be like if not B? What is their mental state if B? For all three, children must disengage from one attentional dimension, be it a sorting rule from the recent past, an actual causal chain of events in reality or their own mental state relating to a situation. Once disengaged children must switch to a new representational perspective based-upon either; an explicit new rule, a counterfactual alteration or, the mental state of another

agent. Lack of flexibility in selection and inhibition between these contrasting representations will impede successful reasoning. Conflicts between these contrasting representations can be weaker or stronger based on factors such as complexity, abstractness, salience, recency, egocentricity or prior knowledge.

Indeed inflexibility, the hallmark of executive function failings, has been implicated in poor reasoning about physical causality (Frye, Zelazo, Brooks & Samuels, 1996) and mental state, namely false belief, understanding (Wellman, Cross & Watson, 2001). It would be expected that functional EF skills are necessary but not sufficient for other forms of counter to fact reasoning. Therefore tasks developed to measure individual differences in EF processes would be expected to relate positively with Mental State Understanding and Counterfactual Reasoning task performances. However, it is crucial to be mindful that in Mental State Understanding and Counterfactual Reasoning the to-be applied rules that ultimately guide responding are not explicit, therefore should be harder to identify and apply than in EF tasks. Transformations must be guided by individual differences in knowledge about mental states and physical causality. There is less explicit information provided with which to generate the necessary representations to be used in supporting subsequent inferences. In Chapter 4 we will explore these relations within our own data, looking at how executive functions correlate with counterfactual and mental state reasoning.

#### **3.2.4. Aims**

Our aim for this chapter was to characterize the executive functioning performance of our sample through the application of the Flanker and DCCS tasks. These tasks on a very basic level challenge the participants to successfully select one of two possible world options based on the actual stimuli presented to them. We wanted to generate individual executive function scores for our future planned

analysis in order to begin answering the question of whether they alone could account for the reasoning skill of young children.

### **3.3. Method**

#### **3.3.1. Participants**

Results presented are from the same sample as described in section 2.3.1.

#### **3.3.2. Measures**

Computerized NIH Toolbox Version 1.0 Flanker and Dimensional Card Change Sort (DCCS) measures were used (Zelazo., 2013). This program was run using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). As described above the Flanker task is a measure of inhibitory control (IC) and the DCCS is a measure of cognitive flexibility.

#### **3.3.3. Procedure**

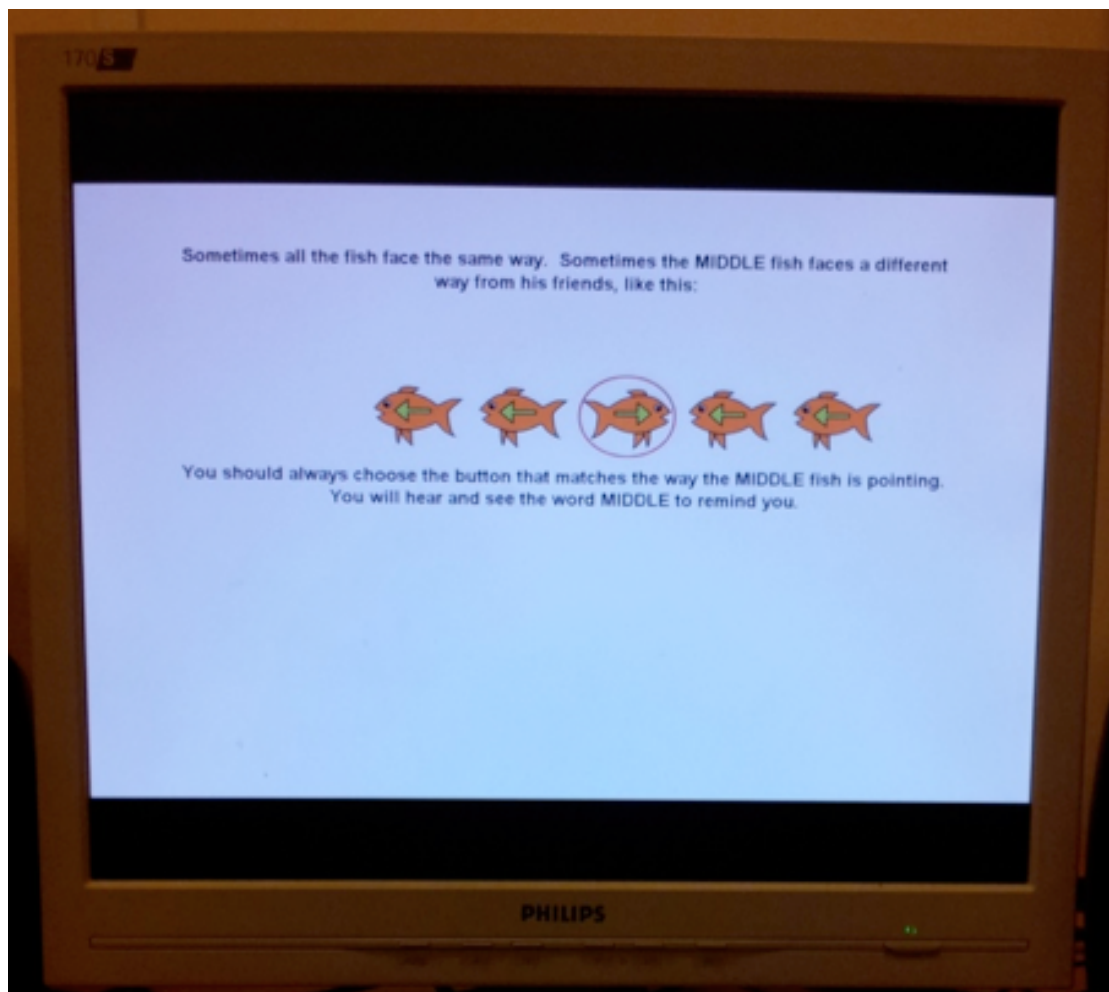
As described in section 2.3.3 children were invited to the University for two visits during which all testing took place. The DCCS task was completed in session 1, followed by the Flanker approximately one week later in session 2. Parents were with their children throughout the testing period. For both tasks participants were seated on a fixed chair facing a computer monitor approximately 60 cm away. This monitor was connected to a laptop running E-Prime 2.0 being operated by the tester and whose screen was not visible to the children. Speakers were placed either side of the monitor and the sound fixed at a suitable level for children to hear clearly. A modified keyboard containing only two buttons was placed in front of the monitor.

**Flanker task.** The official Toolbox instructions can be found in Appendix C. Upon opening the running file training procedures began. The experimenter read the

accompanying text displayed upon the screen during the experiment as the sample children were not yet old enough to read these competently unaided. Rather than using contrasting arrows as with adult versions of this task children saw static, computerized fish on screen that incorporated arrows and were oriented either toward the left or right of the screen (see figure 3.2.). Children were told that they were to help feed these fish by pressing the keyboard buttons (arrows) that corresponded to the direction of the middle fish. They were asked to point to the middle fish to ensure they could correctly identify it. Next they were shown instances where all the fish were pointing the same way, and instances where the target fish was swimming a different way from his friends. They were reminded again that they must always choose the button that corresponds to the way the middle fish was facing. Children were then able to undertake practice trials. On each trial the program stated the word middle to reinforce that there were to respond based on the middle fish. Following incorrect responses the computer highlighted the central fish and reminded the participant to press the corresponding button. Failure to respond within five seconds was also treated as an incorrect response by the program. When  $\frac{3}{4}$  of the practice trials were responded to correctly then Phase 1 of the experiment proper began. After two sets of practice trials if this cut off had not been met, the testing session was terminated with the final thank you screen appearing. In Phase 1 twenty trials were presented and for each the word 'middle' was repeated once by the program. If children exceeded a score of 75% in these 20 test trials in Phase 1 they moved to Phase 2. Participants who scored below this threshold finished the experiment at this point. In Phase 2 the task was the same except there were arrows only instead of fish incorporating arrows to respond to across another 20 trials. Children were reminded



of the rules by the experimenter reading the printed instructions, to press the arrow button corresponding to the target central arrow on the display.



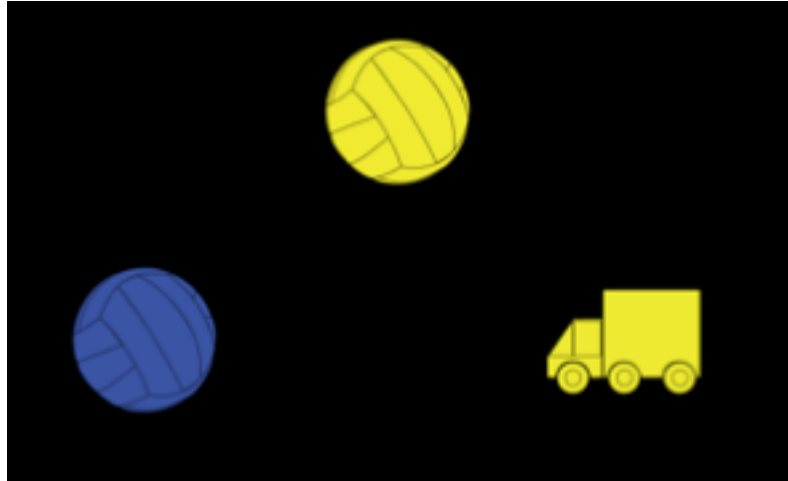
*Figure 3.2.* Photo of the screen presentation used during the Flanker task. Static fish incorporating arrows in their body were used in Phase 1.

**Dimensional card change sort.** Full instructions can be found in Appendix C. Upon opening the program children were introduced to the matching game where they had to sort cards by color and shape. The computer presented the written instructions and the experimenter read them to the participants and highlighted the buttons that were to be pressed. The experimenter completed a demonstration and

then the participant was instructed to try this. Children either played the SHAPE or the COLOR game (determined by the program). In both the practice and test trials two pictures were presented that varied along two dimensions, in this case their shape and color. At the start of each testing block the computer verbally informed the participant which sorting dimension they were to use, shape or color. The central stimulus presented was the target and sorting was dependent on its features. For example see Figure 3.3. If children were tasked with sorting by shape they would determine the central stimulus was a ball and select the other ball. If the sorting dimension was color they would select the other yellow object. Children were given four practice trials for each of the two dimensions. If the incorrect response was chosen, the experimenter would highlight the correct response. Once  $\frac{3}{4}$  of practice trials were answered correctly the test trials could commence. If the child failed to reach this cut off on the first practice block it was reminded of the rules of the game and given another demonstration. If after two sets of practice trials this 75% correct cut off had not been met, the testing session was terminated with the final thank you screen appearing.

Participants then completed 5 test trials sorting along one dimension. If 4/5 were answered correctly they sorted a further five items along the contrasting dimension. If children scored less than 4/5 in the first test block the testing was terminated and the final thank you screen appeared. Those children who sorted 4/5 correctly in the second block then moved to Phase 2. In Phase 2 children completed 30 test trials again sorting along two dimensions, color and shape. Again the experimenter read the instructions presented on screen. The computer selects either shape or color before each trial and the child must listen for which dimension has

been selected on each trial. In Phase 1 sorting dimensions remained stable across trials within a block, in Phase 2 the sorting dimension could change in every trial.



*Figure 3.3.* Example of sorting screen presentation for DCCS. Participants were tasked to sort based on the central stimulus using either color or shape.

### **3.3.4. Coding and Analysis**

In line with the NIH Toolbox Scoring and Interpretation Manual two Executive Function accuracy scores were generated that reflected the total number of correct trials children had completed in the Flanker and DCCS tasks. A correct trial was that in which the appropriate response had been executed in response to the stimulus. Each correct trial, out of a possible 40, was given a score of 0.125. Final accuracy scores for each task were between 0-5, the product of  $0.125 * \text{Number of correct responses}$ .

For participants with an accuracy response rate of greater than 80% a reaction time vector was further derived in line with the procedure described in the Scoring and Interpretation Manual. This score was based on a log<sub>10</sub> transformation of reaction times for correct incongruent trials in the Flanker task and correct non-dominant trials

in the DCCS task only. A score of 0-5 was generated. Accuracy and reaction time scores (if applicable) within each task were then combined for an over all score between 0-10. All of this follows exactly the Scoring and Interpretation Manual guidelines.

### 3.4. Results

Our mean score for the DCCS task was 3.5 (see Table 3.1). The NIH Toolbox Technical Manual presents norming data for all the Toolbox measures. For children aged four years they reported a mean score of 2.2 and for age five they report a mean of 3.5. Our mean score for the Flanker task was 4.9 (see Table 3.1). This exceeds the norms reported for age four (2.8) and age 5 (4.4) but not age 6 (5.8).

Executive Function computed scores correlated positively with one another,  $r=.45$ ,  $p<.001$ . However, when looking at the distribution of scores on these EF measures there seemed to be distinct High versus Low groups emerging around a mean based on performance, (See Figure 3.4.).

Table 3.1. *Table showing mean executive functioning scores for the Flanker and DCCS tasks.*

| Executive Function Task | Mean | Standard Dev. |
|-------------------------|------|---------------|
| Flanker                 | 4.9  | 2.9           |
| DCCS                    | 3.5  | 3.4           |

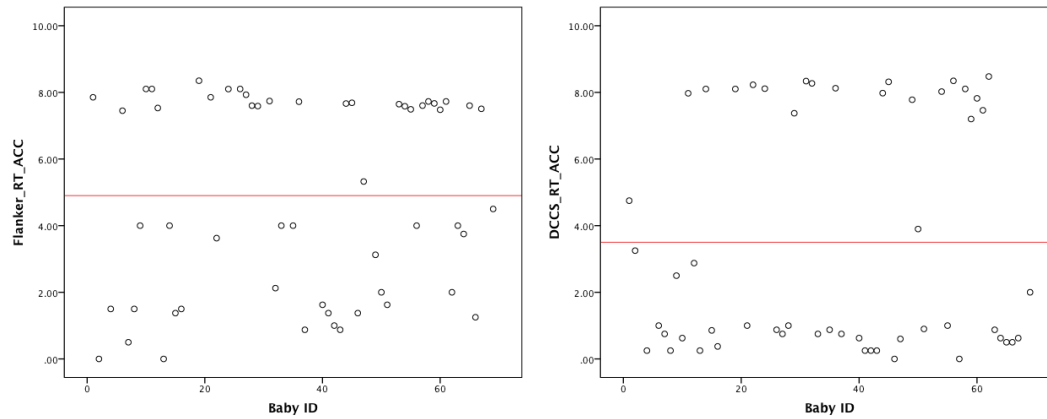
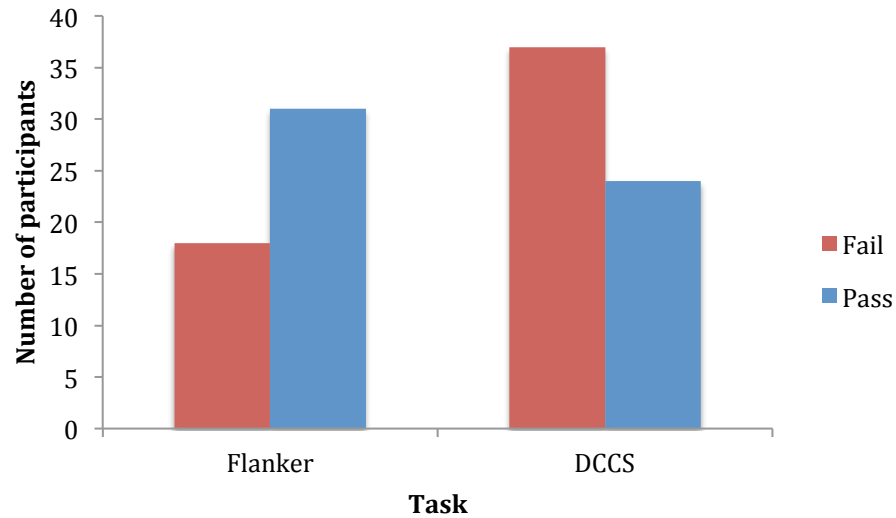


Figure 3.4. Scatterplots showing our samples distribution of combined accuracy and reaction time scores on the Flanker (left) and DCCS (right) tasks. Red lines indicate the mean score.

This pattern emerged as a consequence of the format of the NIH Toolbox. Children were required to achieve a success rate of >75% on target trials to progress from Phase 1 to Phase 2 of the computerized measure. Participants who moved to Phase 2 then had increased accuracy scores due to an increased number of congruent and incongruent trials. Only those children who scored in total had an accuracy score >80% were calculated a reaction time score. This score then boosted their overall performance scores resulting in the large standard deviations seen in Table 3.1. For the Flanker 18 participants failed Phase 1 while 37 passed. In the DCCS 31 failed Phase 1 and 24 passed (see Figure 3.5.). At this age participants are not mostly consistently passing or failing the first phase of these tasks. This suggests that both inhibitory control and cognitive flexibility are still very much developing abilities at this age.



*Figure 3.5.* Number of participants who passed and failed Phase 1 in the two computerized EF tasks.

### 3.5. Discussion

We chose to look at Executive Functioning using these two tasks as it allowed us to measure how well children respond when at the interface between the actual and the possible. Children needed no prior knowledge, they did not need to engage in causal reasoning or chose from a large set of response options. In the target trials for both tasks children must reject the semantically or procedurally pre-potent response in order to satisfy the imposed conditional rule. These demands align with those placed on children during counterfactual reasoning as they must reject the actual world and adhere to the nearest possible world. Further in Mental State Understanding children must reject the actual world as they experience it in order to reason in line with the world as experienced by another.

Within both the Flanker and DCCS tasks comparable numbers of children passed and failed Phase 1. Based on the distribution of the computed scores for these tasks children were divided into high and low groups determined by a median split. 79% of the sample passed phase one of the Flanker while 56% passed phase one of

the DCCS. Proportionally then more children performed better when asked to reject incongruent stimuli at this age than when required to move flexibly between different sorting rules. Although not directly comparable tasks there were many matched aspects of the procedure that are informative. Both tasks were presented in the same computerized format and in the same testing conditions. Children received the same number of trials and needed to meet the same accuracy cut-off thresholds. The general picture then fits with previous research outlining the emergence and development of these two different EF measures with inhibitory control maturing before cognitive flexibility (Anderson et al., 2000; Diamond, 2013).

The DCCS was more demanding in that it requires children to integrate new information in an online way in order to plan their subsequent behavior. In both counterfactual reasoning and mental state understanding children must also update and change their representations of the world and how these changes will effect their own behavior and the behavior of others. Children must efficiently use incoming information and apply rules to it in order to infer or create a new reality from a set of possibilities.

For both the DCCS and Flanker tasks our sample outperformed reported norms. DCCS performance was more consistent with children aged 5 and Flanker performance with children aged 6. The demographics of our sample may have contributed to these increased performance levels. The norming sample was chosen to be representative of the population of the United States of American and was drawn from a number of cities across the country. Families were recruited through the databases of market research companies. Further, families were offered a significant monetary reward of \$90 for completing the single required testing session. A total of 200 participants per age group (defined as 4 years and 0 days to 4 year and 364 days

old) were included in the study. 100 females and 100 males were recruited for each age group.

Our sample all resided in a similar geographical area and were tested within two weeks of being 4 years and 4 months old. Our participants were recruited through community groups and 28 had been involved in earlier data collections taking place monthly for the first 18 months of life. During the four-year testing phase our participants took part in two separate university visits as opposed to one. Furthermore families were given a much smaller monetary reward in the form of £15 gift vouchers and a gift bag. These differences may have altered the diversity of the samples recruited.

Adult analyses of these tasks take into account both accuracy and reaction time scores. Our computed scores were generated in the same way as those computed for adults. However, for many of the participants in our sample their accuracy scores were below the threshold for computing reaction time scores. As such their scores could never exceed 4 on the 0-10 scale. As such the distribution of our scores around the mean were polar and the standard deviation high (see Figure 3.3.). It is clear that for many of the children in our sample the processing of explicit rules related to planned behavior is far from effortless. As such when it comes to the everyday planning of the motor or vocal behaviors that make up children's planned actions their processing abilities will also be far from effortless. Integrating concrete, explicitly input information dictating the planning and execution of simple selection behaviors in EF tasks is challenging for four-year-olds. Even when given all the rules needed to make the correct decision children's moves from the actual to the possible are fraught with errors. In both the EF tasks reported in this chapter the need for knowledge is



minimized and the number of response options are limited and still many children struggled to accurately select the correct possible option based on the actual stimulus.

When further faced with novel content, and a need for implicitly generated rules and the integration of causal knowledge we would expect to see the same children experiencing difficulty. Counterfactual reasoning and Mental State understanding both require the integration of content knowledge in order to correctly infer how some situation will unfold, or to inform the reasoner what the best course of action would be. Further more response options are not constrained in the same way as the discrete options presented during and EF task. Inhibitory control and cognitive flexibility limitations then may hinder children's effective reasoning in these domains as they are unable to effectively and efficiently manipulate and select the required causal information in order to plan and execute the most appropriate goal directed action.

In the next chapter we will investigate the Mental State Reasoning of our sample and relate this to Executive Functioning and Counterfactual Reasoning performances. Mental State Understanding requires the integration of knowledge about ones own, and others' mental states in order to reason about some aspect of an example situation. Further it tasks children with responding based not on their own pre-potent knowledge and experience of a situation, but the mental states of others. Much like with EF tasks children must use a causal if-then rule to predict how another person will act or what they believe about the world.

In the final chapter of this thesis we shall be looking explicitly at how children's moves from the actual to the possible, as measured by all the tasks presented in this thesis, are related. We shall be examining the relationship between

EF and the competence with which children deal with varying facets of the actual versus possible.

### **3.5.1. Limitations**

As shown in Figure 3.1 the distribution of scores in our sample was not as we had hoped for. Although the tasks were user friendly, reliable and provided information rich output we had hoped to capture a more normally distributed data set for each EF. As children get older their accuracy scores improve and with that comes the ability to generate reaction time scores for all the participants. Perhaps had we used a task that did not move through phases based on accuracy performance we may have yielded such a score. A future study should consider carefully the likely data distribution before selecting EF tasks for this age group to be used in a correlational way.

### **3.5.2. Conclusions**

There was great variation in the Executive Functioning performance of the children in our sample. 67% of our sample reached the second phase of the Flanker task and 56% reached the second phase of the DCCS. It is clear then that at age 4;4 many children still struggle with immature inhibitory control and cognitive flexibility. These tasks allowed us to see how well children could move from the actual world to another possible world when knowledge was minimized and options constrained. Although our sample's performance was slightly above reported norms for the two tasks, particularly in the case of the Flanker task, performance was still far from ceiling in both tasks. Children at this age are still having problems moving from the actual to the possible despite clear conditional rules and simple stimuli.

## **Chapter 4. Moving Between Your Actual and Their Possible: Conditional Reasoning About the Mind of Another**

### **4.1. Abstract**

Young children can be relatively successful when reasoning using simple, physically observable, or pre-defined causal rules. When the conditional rules necessary for reasoning are internally generated or experienced by another intentional agent, however, things become more challenging. Classic theory of mind research often utilized false belief tasks to investigate children's understanding of how other's beliefs impacted upon their behavior (e.g. Baron-Cohen, Leslie & Frith, 1985; Wimmer & Perner, 1983). More recently a richer story has been presented in which typically developing children largely follow the same trajectory when it comes to reasoning about many types of mental state, not restricted only to the representation of false beliefs (e.g. Peterson & Wellman, 2009; Kristen, Thörmer, Hofer, Aschersleben & Sodian, 2006; Wellman & Lui, 2004). There are, however, broad individual differences within typically developing groups, as well as group level differences when comparing typically and atypically developing young children such as those with Autism Spectrum Disorder (Baron-Cohen et al., 1985). Although mental state understanding shares many procedural similarities with other types of conditional reasoning the content to be represented and reasoned about can have a huge effect on success. In this chapter I will review the development of mental state understanding while discussing some of the methods used to assess it. Next I will put forward some suggestions about the nature of mental state understandings relationship with other nascent abilities. Finally I shall present data again collected from the same sample of four year-old children as Chapters 2 and 3 that characterizes the nature of

their mental state understanding and provides us with individual differences information for subsequent analyses.

## **4.2. Introduction**

An interest in conditional reasoning and the relationship between the actual and the possible for the developing child should take an interest in Mental State Understanding for several distinct reasons. Firstly, experimental tests of Mental State Understanding frequently (or even typically) involve conflicts between representations of the state of the world in order to assess genuine understanding in much the same way that Counterfactual Reasoning tasks are important to understanding the child's ability for conditional reasoning. Many Mental State Understanding tasks involve a conflict between an independent character's (inner) beliefs about the world and the actual state of the world. This makes direct comparisons between conditional reasoning and Mental State Understanding tasks theoretically interesting. Second, appreciating the difference between beliefs about the world and the way the world actually is seems integral to a full ability to modify the world in order to suit one's needs. One cannot seek out an attractive possible world without understanding the actual one. And, finally, one of the main sources of information to human beings about how the world could be changed is the overt behavior of others. We will be examining this source of information in the subsequent chapters on imitation. For the moment, it suffices to say that for others' behavior to be maximally informative, understanding their goals and intentions, that is Mental State Understanding itself, is essential.

In the previous two chapters we focused almost exclusively on the kinds of concrete conditional decisions children can make, those based upon physical and

observable changes in the world. However, when one begins to speak about those causal conditional relationships that are not directly observable in a conventional sense a new literature comes to the fore: Mental State Understanding. When children see an egg fall to the floor, the resultant causal effect is readily observable, a rather messy and slimy area of broken egg. When a person makes a hurtful comment, the direct causal effect of this comment may or may not be seen directly and either way may not truly be reflecting the resultant psychological outcome. Conditional reasoning based on mental states rests upon the same tenets as other forms of conditional reasoning. However, the content and causal rules used to generate and reason from them rely on inferred representations of someone else's psychological world. It is this representation of the minds of others that marks mental state understanding as a "special case".

In previous research elements of MSU have been reported under various labels such as false belief understanding (FB), mentalizing and theory of mind (ToM). As such the research groups investigating the phenomenon of MSU have postulated a broad host of theories trying to account for its' development and cognitive underpinnings. This genre of study is broadly interested in the same set of abilities when children must reason using information about the mental states of other. No longer are children reasoning about their own view of reality, or a counterfactual alternative to it; instead children are challenged to reason about the world as represented by another intentional agent.

#### **4.2.1. What is Mental State Understanding?**

Mental states can include but are not limited to intentions, desires, thoughts, beliefs, dreams and pretense (Baron-Cohen, Ring, Moriarty, Schmitz, Costa & Ell, 1994). Researchers have sought to gather evidence of mental state understanding

through probing the limits of children's knowledge about how mental states are causally linked to perceptual inputs, behaviors and of course other mental states (Flavell, 2000). MSU is one of the socio-cognitive tools that enable us to navigate our social worlds efficiently through being able to understand past behaviors, and predict future ones (Moore & Frye, 1991). Some researchers have gone as far as to claim that mental state understanding is so important that adults implicitly track the knowledge and belief states of others, regardless of whether there is any explicit pressure to do so (van der Wel, Sebanz & Knoblich, 2014).

Mental state understanding (MSU) as we shall measure it refers to an individual's ability to understand and reason about the thoughts, feelings, intentions, desires and beliefs of others (Wellman & Liu, 2004). Flobbe, Verbrugge, Hendriks and Kramer (2008) do not see MSU as just one mental ability, but a collection of lesser processes being deployed as and when required. To classify as MSU these processes are necessarily related to reasoning about others' mental states. In a thread not dissimilar to the re-description account of EF explored in the previous chapter (Perner et al., 2002), mental state understanding requires children move away from a Piagetian egocentric world view and towards an appreciation of multiple perspectives (Flavell, 2000). Children must remove or pause their own mental states from the reasoning equation.

Many of the investigations into MSU have used the popular marker of false belief understanding as a proxy for MSU more generally (Shahaeian, Peterson, Slaughter & Wellman, 2011; Wellman & Liu, 2004). However, MSU can be indexed using any task that requires the participant to infer a person's behavior based on their internal mental states (Shahaeian et al., 2011). Although commonly tested using paradigms where participants must explicitly indicate their answers, mental state

understanding has also been inferred from participant's implicit looking behaviors (e.g. Call & Tomasello, 1999; Clements & Perner, 1994; He, Bolz & Baillargeon, 2012; Low, 2010; Scott, He, Baillargeon & Cummins, 2012.). Both types of task are interested in identifying whether participants will correctly predict the mental state, or mental state driven behaviors, of another individual.

#### **4.2.2. Developmental Investigations into MSU**

The large volume of research carried out into mental state understanding in the past two decades has resulted in a corresponding variety of paradigms designed to assess this ability. Classic theory-of-mind research most often used false-belief paradigms such as the Maxi task or the Sally-Anne tasks (Baron-Cohen, Leslie & Frith, 1985; Wimmer & Perner, 1983). In the Maxi task for example Maxi stores some chocolate before leaving the room, his mother then moves the chocolate in view of the participant. When Maxi returns the participants are asked a belief question, "Where will Maxi look for the chocolate?" To pass this task, children must answer in accordance with Maxi's false belief, not reality. From the information they have about the actual world, they must make an inference about an alternative world, based on the relevant causal information they possess about how knowledge is accrued and how this knowledge informs subsequent behavior. Children are required to reason not about how the world actually is but the representation of the world that the questioner is interested in, Maxi's world. In one dimension this contrasts with the type of counterfactual reasoning tasks presented in chapter 2, where an element of the actual world has been hypothetically changed, mental state understanding tasks ask about a world different by virtue of the mental state of another person. However, both tasks in a broad sense can be understood as counterfactual in that the representation of the

possible world that must be reasoned about to answer successfully is in tension with the actual subjective reality experienced by the reasoner.

By the age of four or five typically developing children begin to consistently answer false belief questions correctly (e.g. Wellman, Cross & Watson, 2001; Wimmer & Perner, 1983). However, mental state reasoning more broadly extends beyond understanding another's false belief. When children reason about others' minds they need not only represent their beliefs but also other psychological states. The desires, intentions, thoughts, emotions, dreams, pretense and knowledge of other individuals for example can also be represented and reasoned about (Baron-Cohen et al., 1994). Though these psychological states are conceptually different from one another they all fall under the umbrella of mental state reasoning. Given our interest in how children navigate from the actual to the possible it was important for us to examine reasoning ability with a range of psychological states, not simply false beliefs. Further we wanted to find a measure of mental state reasoning that gave us a scale of performance rather than a pass/fail dichotomy.

Wellman and Liu (2004) developed such a battery of scaled tasks in order to provide some way of indexing mental state reasoning performance across childhood. They wished to provide a set of tasks that measured the aforementioned conceptually different elements of mental state understanding, that is not only false belief understanding, while controlling for, as far as possible, the influence of task demands on performance. The tasks were chosen and ordered based on literature about the kinds of mental states that children seemed able to reason about more easily than others at different stages in development. The final, highly replicable scale has been administered cross culturally and to different age ranges. It includes:



- (i) **Diverse Desires:** This task challenges children to indicate which snack a protagonist will choose to eat, a carrot or a cookie. Children are shown pictures of the two snacks and asked which they would prefer to eat. Once they have indicated a preference the protagonist (a puppet) is introduced to the scene. The children are told that he does not like their choice, what he likes most is the other option. To answer correctly children must indicate the snack that the protagonist prefers. This task assesses the ability to infer a character's behaviour based on their diverse desire.
- (ii) **Diverse Belief:** Structured in the same way as the previous item. Children are shown two pictures, a garage/shed and a bush. They are told that Linda is looking for her cat and asked where they think Linda's cat might be hiding. Once they indicate their belief they are told that Linda thinks her cat is in the contrasting location. The belief question is then given "Where will Linda look for her cat, in the garage or in the bushes?" To answer correctly participants must answer in accordance with Linda's belief.
- (iii) **Knowledge Access:** An opaque box is presented to the participant and they are asked to guess what is inside. The contents are then revealed, a plastic ball. The box is closed again and Polly the puppet is introduced to the scene. Polly has never ever seen inside this box. The participants are asked, "Does Polly know what is in the box?" and "Has Polly ever seen inside the box?" To be correct children must answer both elements correctly, that is in line with Polly's available knowledge. This task requires children to infer that Polly will not know what is inside the box, given the fact she has not seen inside.

- (iv) Contents False Belief: A familiar iconic container is shown to participants (e.g. Smarties Tube, Maltesers Box). Children are asked what they think is inside and are then shown the true unexpected contents, a pen. The container is closed and Peter, another puppet is introduced who has never seen inside the container. Participants are asked what Peter thinks is in the container, e.g. Smarties or a pen, and whether he has ever seen inside. Like the knowledge access task children must answer both correctly to pass. This task requires children to infer that seeing the iconic container will cause the protagonist to expect the iconic contents, not a pen.
- (v) Real-apparent Emotion: Children are told a story about a boy who had a mean joke told about him in front of his friends. The boy doesn't want his friends to know how he feels about the joke as they would call him a baby so he tries to hide how he felt. Children are asked to indicate how they boy feels inside and on his face by pointing to emotion faces. If, when asked about how the boy tries to look on his face, they indicate a more positive emotion face than when asked about how he feels inside, then they are credited as having passed the task. This task challenges children to understand that a person may feel one way on the inside but choose to look another way on the outside for some reason.

It is reported that around 80% of children pass these tasks in the order they are presented above (Wellman & Liu, 2004). Diverse desires tasks are the first to be conquered with upwards of 95% of three-five year olds giving the correct response. In contrast only 32% of the same sample were successful on the real-apparent emotion task. Similar trajectories have been reported in Australian and German children

(Peterson & Wellman, 2009; Kristen, Thörmer, Hofer, Aschersleben & Sodian, 2006) (although see differing cultural trends in Shahaeian, Peterson, Slaughter and Wellman (2011) and Wellman, Fang, Liu, Zhu and Liu (2006)). As with other types of conditional reasoning, MSU relies on identifying if-then relationships, in this case between mental states and behaviors. This content specific information can come from the self or from interactions with and observations of other intentional agents. Generating causal dependencies such as when I, or person A, want X, we choose X. If I, or person A believe X, we act in accordance to X and so on. These kind of causal dependencies must come about from experience of attributing actions to mental states. As such many social abilities may be linked to superior MS reasoning.

Generating representations of those causal dependencies driven by others' minds is a more challenging task than generating representations of causal relationships physically observable in the world as the antecedents are often not observable. As discussed briefly in earlier chapters the relationship between an egg falling and an egg smashing is observable: the relationship between a person's desire for a cookie, and them eating a cookie is not. The hierarchical nature of children's proficiency with different types of MSU reflects the commonality and ease of identification of those mental states. People more often act in accordance with their own desires than in accordance with some false belief. As such children have less experience with observing actions undertaken in accordance with false beliefs. The case of real versus apparent emotion also highlights this. Children can only see apparent emotions: the idea of a true emotion hidden behind an apparent one is a complex attribution only possible when children possess very proficient meta-cognitive skills.

### **4.2.3. The Relationship Between Young Children's MSU and Other Developing Abilities**

MSU, as measured through a variety of false belief and other tasks, has been related positively to many other linguistic, social and cognitive skills and abilities throughout childhood. Milligan, Astington & Dack (2007) assert that without a doubt there is a relationship between language and false belief understanding in particular based on their meta-analysis of 324 studies. Use of mental state terms in social interactions, mastery of tenses, complements, syntax, semantic, and general language abilities have all shown positive relationships with mental understanding indices (Bretherton & Beeghly, 1982; Brown, Donelan-McCall & Dunn, 1996; Chiarella, Kristen, Poulin-Dubois & Sodian, 2013; deVilliers & Pyers, 2002; Milligan et al., 2007; Olineck & Poulin-Dubois, 2005). Due to the correlational nature of some studies it has been difficult to identify the cause-effect relationship between language and MSU (Brown et al., 1996; Chiarella et al., 2013). However, in their meta-analysis Milligan et al. in combination with earlier studies (Astington & Jenkins, 1999; de Villiers & Pyers, 2002) judge there to be an increased magnitude of predictive effect from language to theory of mind understanding than the reverse. Earlier measures of language predicted subsequent theory of mind developments whilst the opposite was not the case. Milligan et al. surmise that this direction of effect is due to language skill affording children the opportunity to represent and communicate false beliefs.

Social behaviors such as a shy temperament, moral judgments, observant, non-aggressive temperament and sharing behaviors have shown similar positive relations (Mink, Henning & Aschersleben, 2014; Smetana, Jambon, Conry-Murray & Sturge-Apple, 2012; Wellman, Lane, LaBounty & Olson, 2011; Wu & Su, 2014). Mink et al.

(2014) reported that children with a shy temperament at 18 months had higher MSU scores at age 3 years, as measured by diverse desire and belief tasks, and a knowledge access task. Smetana et al. (2012) investigated the development of children's moral judgments alongside their MSU developments. They concluded the two skills developed in a symbiotic manner.

Finally, attentional and socio-cognitive behaviors such as early attention decrements during habituation, understanding pointing gestures and intention understanding have been reported to predict MSU in young children (Aschersleben, Hofer & Jovanovic, 2008; Colonnesi, Rieffe, Koops & Perucchini, 2008; Wellman & Brandone, 2009; Wellman, Lopez-Duran, LaBounty, Hamilton, 2008). Even when general measures of cognitive development such as language, IQ and EF were controlled, Wellman et al. (2011) reported that social attention, assessed through a decrement of attention paradigm at 10 and 12-months, positively predicted theory of mind at 4-years as measured by the Wellman and Liu (2004) battery. Aschersleben et al. (2008) also used the Wellman and Liu (2004) measure, however, they also inserted an additional false belief item. The authors reported that decrement of attention to goal directed actions specifically at 6-months old related positively to the solving of false belief tasks (but not the other task items) three and a half years later. Both Wellman et al. (2008) and Aschersleben et al. (2008) failed to find a concurrent relationship between language and mental state understanding, this is in contrast to the findings of Milligan et al. (2007). Wellman et al. (2008) and Aschersleben et al. (2008) support a specific social cognition continuity hypothesis. Developmental studies such as these have identified a robust link between indices of social cognition in the early years. They claim developing an understanding of the social agents

around you, and interpreting the actions they undertake as intentional, supports later understanding of their mental states, and how these mental states may affect behavior.

Charman, Baron-Cohen, Swettenham, Baird, Cox and Drew (2000) extend this idea by asserting that a shared social-communicative representational system links abilities such as joint attention, language, imitation, play and mental state understanding. Their longitudinal data provided evidence that infants' joint attention abilities at 20-months were positively related to their mental state understanding at 44-months. Joint attention, in their paradigm, was assessed through the proportion of gaze-switching instances during two experimental tasks designed to be novel and engaging to the infant. In addition, like Wellman et al. (2008) and Aschersleben et al. (2008) Charman et al. (2000) failed to identify concurrent language and mental state relationships. Charman et al. (2000) focus on joint attention in terms of it being in pursuit of some social goal at sharing their representation, or mental state of perception. They too suggest these early social abilities develop in to more sophisticated mental state understanding over the subsequent years.

#### **4.2.4. Structural and Conceptual Similarities Between Executive Functioning Tasks, Counterfactual Reasoning and Mental State Understanding**

For young children to successfully answer mental state understanding questions the first hurdle they must over-come is reasoning on the basis of their *own* desires, beliefs, knowledge and emotions: this requires EF. If they do not succeed in doing so they will select realist answers, those based on their own mental state, the struggle to overcome a “curse of knowledge” (Birch & Bloom, 2004). It is not only young children's performance that suffers in this way as adults given tasks that require the prediction of a naïve reasoner's beliefs about a situation often fail to provide the most objectively logical answers (e.g. Fischhoff, 1975). It has been argued that all

cognitive tasks require EF to some extent (Alexander & Stuss, 2000). Selecting one response over another requires the kind of effortful and goal directed processing characteristic of mature EF. However, as evidenced by the disparity between task performance in the battery, children do not suddenly shift from failure to success in this aspect. Instead different types of mental states are easier to draw inferences from than others and they develop in a relatively uniform pattern (with the aforementioned exception of Chinese and Iranian children, (Shahaeian et al., 2011)). It could be that inhibiting one's own desires and instead selecting an incongruent option is easier than doing the same for conflicting beliefs.

However, it could also be that generating a representation of another person's desires is more straightforward than generating a representation of their false beliefs. The causal representations they have constructed regarding other's desires and beliefs may not be robust enough to support valid inferences about their behavior. Counterfactual reasoning relies on knowledge about the causal relationships between the elements of a given situation. Without this knowledge, or an ability to seek out the required information CFR will falter. Even once children have the ability to represent such concepts as beliefs and desires, they may not yet be able to reliably infer behaviors based upon them (Onishi, Baillargeon & Leslie, 2007). As with CFR and EF, MSU relies heavily on the ability of an individual to select one valid response, over one pre-potent response. Children must answer in terms of a world not yet realized, and/or realized differently, rather than one already assembled, contrasting the actual vs. the possible, based on their knowledge of causal structure and mental states.

Given that EF, CFR and MSU are all forms of conditional reasoning, albeit with different content and rule bases being required, it should come as no surprise that

there are many coalescing links. Executive functions, (both CF and IC) have indeed been shown to have a predictive relationship with MSU (e.g. Benson, Sabbagh, Carlson & Zelazo, 2013; Henning, Spinath & Aschersleben, 2011; Hughes & Ensor, 2007; Muller, Zelazo & Imrisek, 2005). Benson et al. showed an increase in ToM performance following an explicit EF training protocol (Benson et al., 2013). Furthermore a direct link between CFR and MSU has been suggested and was discussed in Chapter 3 (e.g. Guajardo & Turley-Ames, 2004; Riggs et al., 1998). The strength of this link, however, is still contentious with as little as 3% of the variance being explained (Guajardo et al., 2009). It does seem that structurally CFR & MSU tasks have much in common: (i) they require a selection between one more salient and one less salient option, (ii) representations must be generated representing two or more possible causal structures (the actual and the possible), (iii) a correct response must be based on not only the actual world's events but inferred extensions or alterations of them. Where the tasks diverge is that CFR tasks need not necessarily involve mental states of any kind (e.g. Beck et al., 2006). Mental states are more difficult to represent and draw inferences from than concrete and observable causal events where the physical structure embodies the causal one. Therefore the demands of each task will vary based on the content within them, how well that content is understood and how easily and soundly its' causal structure is represented.

MSU in essence is the ability to understand and reason about the causal structure of mental states as MSU requires individuals to make predictions about behavior based on their knowledge of the cause and effect relationships of internal mental-states (e.g. Shahaeian et al., 2001; Wellman & Liu, 2004.) CFR may or may not involve mental states but conceptually the tasks that measure these two abilities are strikingly similar. For a young child predicting the emotional outcome of others



given a situation or set of circumstances can be a challenge. Through their interactions children become increasingly exposed to the causal impacts of their own and others' behaviors on the mental states of both themselves and others. When faced with making real time decisions that are grounded in the current mental state of a given person their ability to incorporate past experiences and select and appropriate future behavior to achieve the desired social outcome is being put to the test. Children who are well able to decide that taking off their shoes will prevent the floor from becoming dirty (based on prior experience) may not yet appreciate that taking off their shoes may also prevent mummy from becoming angry. For both children and adults, the conditional relationships between actions and psychological outcomes are not as readily accessible as the relationships between action and physical outcomes. Adults in some situations still fail to use their mental state understanding to successfully guide their behaviors (Keysar, Lin & Barr, 2003). Never the less the same decision-making processes apply to counterfactual reasoning about physical causal relationships and mental state causal relationships: it is the type of content that differs.

In addition psychological if-then relationships may not be as deterministic and reliable as physical if-then relationships: how your social partner reacts to being teased on one day may differ to their reaction on the next, a lesson I am sure we have all learned the hard way; an egg falling from a counter will break regardless. For mental state understanding individuals stand at the interface between a past psychological reality (for themselves and/or for others) and a new one of their choosing. Having a body of knowledge about people's past and current desires, beliefs, intentions and knowledge states gives children the unique ability to make decisions about how their own actions and behaviors can alter or improve these

mental states to achieve some end. Mental state understanding may not always lead to a change in behavior, it is a two way street. Simply understanding the mental state of another, may influence our own mental state, this alone is a particularly powerful effect.

#### **4.2.5. Potential Benefits of Mental State Understanding for Children**

Mental state understanding has the potential to dramatically increase the amount of causal information children are exposed to, and subsequently retain. Being able to represent the causal motivations for others' behaviors and appraise the causal repercussions of these behaviors is an invaluable resource. Children do not need to personally experience mental states to successfully reason about them in a conditional way. Seeing another child become upset because they dropped their ice cream while running around will be understood, despite not necessarily having found ourselves in the same, or even a similar situation before. We can infer that the consequence of an undesirable event will be an undesirable emotion, perhaps to a level that we have not experienced ourselves, but can understand, such as terror. Children need not have experienced a certain situation to deploy appropriate, or avoid inappropriate, behaviors having seen others do so previously. Conditional reasoning about mental states, as with conditional reasoning about any event in the world is invariably influenced by experience, both personal experience and the experience of observing others.

As we move through the world we see other intentional agents interacting with objects and engaging in social behaviors of many types. Some of these have easily recognizable goals by virtue of the physical manipulations taking place; pressing this lever will open that door. Others are more opaque and abstract posing more of a challenge for young humans to understand such as giving some words of

encouragement will improve a person's mood. However, understanding causal relationships and understanding the intentions and goals of others' affords children a very unique opportunity, imitation. When children see others interacting with objects or one another they observe many behaviors and many outcomes. Many of these outcomes will be desirable to the young human, like getting chocolate out of a tricky container, or initiating a positive and rewarding social interaction. Seeing and understanding which behaviors bring about desirable outcome for others, gives children the opportunity to replicate said behaviors in the quest for bringing about similarly positive results for themselves: without understanding the conditional relationships between actions and outcomes, and possible actions and their resultant possible outcomes this would not be attainable. Furthermore without understanding that others' past actions and their own possible future actions are equally causally adept at bringing about the same or at least similar outcomes, children would not be motivated to imitate. Imitation gives children the perfect tool to stand on the shoulders of giants by using information about the giant's mental states combined with the causal repercussions of the giant's actions to guide their own behavior. In Chapter 5 I shall discuss imitation in more detail.

#### **4.2.6. Aims**

Our first aim for this chapter was to assess the mental state understanding of our sample. We utilized the mental state understanding battery previously detailed (Wellman & Liu, 2004). Children would be given individual scores based on their performance over the five items included in the scale. These five items were interested in diverse desires, diverse beliefs, knowledge access, false-contents and real versus apparent emotions.

Having already collected individual differences data about our sample's counterfactual reasoning and executive functions we inspected the data for evidence of a relationship between these three variables. Children's reasoning, both counterfactual conditional and mental state, is subject to great improvements over the pre-school years and beyond. Processing ease, as assessed via executive functioning too is rapidly developing. Aside from these fortuitous temporal factors there are numerous structural and conceptual similarities that these three skills share.

All three variables we have so far discussed challenge children to reason conditionally: They must apply some 'if-then' rule to the information they have in order to guide their response. Counterfactual reasoning gives children the 'if' antecedent element, while asking them to reason about and signal the 'then' consequent. Executive functioning gives children both the 'if' and the 'then' elements removing the need to formulate a novel response. Mental state understanding, much like CFR also gives the 'if' element whilst the 'then' element must be generated. In all three responses must be made that predict or change the state of reality based on this if-then rule. In CFR and MSU especially, these possible realities are in contrast to the actual one currently being experienced by the reasoner. Counterfactual answers must be based on a reality that once could, but no longer could be: mental state answers must be based on a reality that is perceived by another, be it congruent with the reasoner's representation of reality or otherwise.

Of course, as discussed in the previous chapter, EFs are integrally associated with the planning and execution of goal directed behaviors. Our second aim then was to identify whether CFR and MSU performance in our sample was positively related to EF. It was expected that individual differences in EF would relate positively to individual differences in both CFR and MSU. Although there has been a mixed bag of

evidence with regards to CFR and EF (e.g. Beck et al., 2010; Guajardo et al., 2009; Riggs et al., 1998) we did expect to find some evidence that executive functioning is correlated of reasoning success.

Our third and final aim for this chapter was to ascertain whether mental state understanding and counterfactual reasoning performance were related in our sample as has been reported before (Guajardo et al., 2004; 2009; Riggs et al., 1998). Guajardo and colleagues (2009) reported that when age and verbal ability were taken into consideration the contribution of CFR to MSU, although significant was very small.

There were some fundamental differences to the measurement methods utilized in these previous studies when compared to our own. Guajardo et al. (2004) administered a paradigm that recorded children's ability to alter the antecedent element of a counterfactual situation rather than identify a consequent. They asked how an outcome could have been changed rather than asking how an outcome would have changed given some event. This contrasting measurement decision may be observing a radically different skill in the sample. Riggs et al. (1998) who also reported a strong positive relationship between CFR and MSU used experimental materials that were almost identical for both types of reasoning. Children heard different questions, related to the same scenarios, from which their MSU and CFR scores were obtained. The very nature of the stimuli being virtually identical may have contributed to the shared variance seen.

Additionally both the aforementioned studies that reported a CFR-MSU link relied solely on false belief tasks as the measure of MSU, or ToM as they report it. False belief tasks measure one specific facet of MSU and as we wished to compile a more comprehensive picture of children's mental state reasoning skills we included a wider range of items. Indeed due to the nature of the battery there was no classic false

belief task like the original Sally-Ann task (Wimmer & Perner, 1990). Instead a false contents task was used to set up a false belief. The false contents task does not involve deception in the same way as the false belief task and as such may rely on different cognitive processes.

In counterfactual tasks there is the actual world reality, and then a contrasting counterfactual world reality, however, the counterfactual world must be generated based on a hypothesized, rather than observed change. So, despite having the same number of representations, the false contents and counterfactual tasks have differing levels of demand. The counterfactual reasoning items require the generation of an alternative world due to some change having been made, whereas the information about the false contents world is more readily observable and stays static it: it is only the perception of the world that differs, not the world itself.

If there is a relationship between CFR and MSU and it is driven by having multiple representations of the world alone then we would expect to see a relationships between our chosen MSU and CFR items. If, however, the classic false belief tasks provide an additional layer of similarity, through the need for a change, or updating of the world through the integration of new information, then perhaps we would not expect to find the same positive linear trend between these two developing-forms of conditional reasoning as measured in this study. Riggs et al. (1998) pointed out that ToM and CFR tasks both deal with counter to reality beliefs. This is also true of the false contents task, however, there has been no need to integrate additional information that would change beliefs in the false contents task. In both the false belief and CFR tasks not only are beliefs counter to reality but that have also had to be updated in light of new information, either the event of Ann moving an item, or an alternative antecedent being introduced.

### **4.3. Method**

To assess MSU we utilized a scaled battery of 5-items (as discussed earlier in section 4.5; Wellman & Liu, 2004). This battery has been used in numerous experimental paradigms and proved a reliable and functional tool for assessing the development of children's understanding of others minds (e.g. Wellman et al., 2008; Aschersleben et al., 2008). The items were designed to be as structurally and procedurally similar as possible to reduce additional task demands altering performance. Each item differed only in the type of mental state component it was designed to probe. The items included have been shown to provide variability in children aged 4; with above chance performance on the first item and below chance on the final. We expected to replicate the pattern of success previously reported in our sample of British children.

#### **4.3.1. Participants**

Our sample was the same as that described in Section 2.3.1.

#### **4.3.2. Materials**

The following MSU measure was adapted from the 5-item scale assembled in Wellman & Liu (2004) described initially in section 4.2.2. Item one required a laminated A4 page with a cartoon picture of a cookie on one side and a carrot on the other (counterbalanced across participants with half seeing the carrot on the right, and half the carrot on the left), and a small, male wooden doll. Item two also required a laminated A4 page. This depicted a cartoon picture of a shed/garage on one side and some bushes on the other (again counterbalanced with half seeing the shed on the left and half the shed on the right). A small female wooden doll was also used. For item 3 we used a small, opaque cardboard box (10cm x 10cm x 10cm). Inside this was placed a yellow plastic ball around 7cm in diameter. A third small wooden doll,

female, was also used. For item 4 we used a popular style children's juice bottle, inside of which was a pen. Again a small wooden doll was required, this time a male. Finally for item 5 two laminated sheets were required, one depicting a cartoon boy as viewed from behind, and one presenting three simple, genderless, emotion faces; happy, sad and, neutral. The dolls for items one through four were all different previously unknown characters, Mr. Jones, Linda, Polly and Ben. Two were male and two female. All these stimuli can be found in Appendix D.

#### **4.3.3. Procedure**

Testing took place as part of the schedule described in Chapter 2.3.1. The MSU battery was completed in session 1 at the end of children's first visit to the university. The MSU tasks were administered in a purpose built testing room with inbuilt digital audio and visual recording capabilities. The experimenter and child sat facing one another, either side of a small square table that was covered in a black table-cloth. All the materials were stored under this table and were presented and removed one at a time. Each item corresponded to a set script (see Appendix D). Items were always administered in the same order, with the easier questions before the harder: Diverse desires, diverse beliefs, knowledge access, false belief and real-apparent emotion.

- (i) Diverse Desires: This task challenges children to indicate which snack a protagonist will choose to eat; a carrot or a cookie. Children are shown pictures of the two snacks and asked which they would prefer to eat. Once they have indicated a preference the protagonist (a doll) is introduced to the scene. The children are told that he does not like their choice, what he likes most is the other option. To answer correctly children must indicate



the snack that the protagonist prefers. This task assesses the ability to infer a character's behaviour based on their diverse desire.

- (ii) **Diverse Belief:** Structured in the same way as the previous item. Children are shown two pictures; a garage/shed and a bush. Linda is introduced (a doll). They are told that Linda is looking for her cat and asked why they think Linda's cat might be hiding. Once they indicate their belief they are told that Linda thinks her cat is in the contrasting location. The belief question is then given "Where will Linda look for her cat, in the garage or in the bushes?" To answer correctly participants must answer in accordance with Linda's belief.
- (iii) **Knowledge Access:** An opaque box is presented to the participant and they are asked to guess what is inside. The contents are then revealed: a plastic ball. The box is closed again and Polly the doll is introduced to the scene. Polly has never ever seen inside this box. The participants are asked, "Does Polly know what is in the box?" and "Has Polly ever seen inside the box?" (memory check). To be correct children must answer both elements correctly, that is in line with Polly's available knowledge. This task requires children to infer that Polly will not know what is inside the box, given the fact she has not seen inside.
- (iv) **Contents False Belief:** A familiar iconic juice bottle is shown to the participant. They are asked what they think is inside and are then shown the true unexpected contents; a pen. The container is closed and Peter, another doll is introduced who has never seen inside the container. Participants are asked what Peter thinks is in the container, i.e. juice or a pen, and whether he has ever seen inside (memory check). Like the

knowledge access task children must answer both correctly to pass. This task requires children to infer that seeing the iconic container will cause the protagonist to expect the iconic contents, not a pen.

- (v) Real-apparent Emotion: Children are told a story about a boy who had a mean joke told about him in front of his friends. The boy doesn't want his friends to know how he feels about the joke as they would call him a baby so he tries to hide how he felt. Children are asked two memory check questions about the stories content, (i) "what did the other children do when the mean joke was told?" and (ii) "what would the other children say if they knew how the boy really felt about the joke?" They must answer both correctly to be awarded points for passing the test questions. Children were then asked to indicate how the boy feels inside and on his face by pointing to emotion faces. If, when asked about how the boy tries to look on his face, they indicate a more positive emotion face than when asked about how he feels inside, then they are credited as having passed the task.

Answers were coded by a primary and secondary coder from the recordings after all sample data had been collected. The primary coder was the experimenter. The secondary coder was an undergraduate placement student who was blind to the hypotheses of the experiment. For both tasks coding began once the experimenter had finished asking the memory check and target questions. If participants did not respond the question was asked a second time. Participant responses were coded when they either pointed to or verbalized their choice. Coding reliability was at 100%.

#### 4.4. Results

Each child was administered the five item scale from the original Wellman and Lui (2004) paper. For each item they responded to correctly (including memory checks) they received one point culminating in a 0-5 scale. The percentage of participants who passed the test question for each item can be seen in Figure 4.1. Table 4.1 displays percentage success rates for children in three other studies using the same battery. Performance on Item 1 the Diverse Desires question was lower in our sample than the others. Performance on questions 2 through 4 were in the mid-range when compared to the others. Item 5 was lower in our sample when compared with the others as no child answered both memory check questions correctly and as such were marked as incorrect regardless of their test question responses.

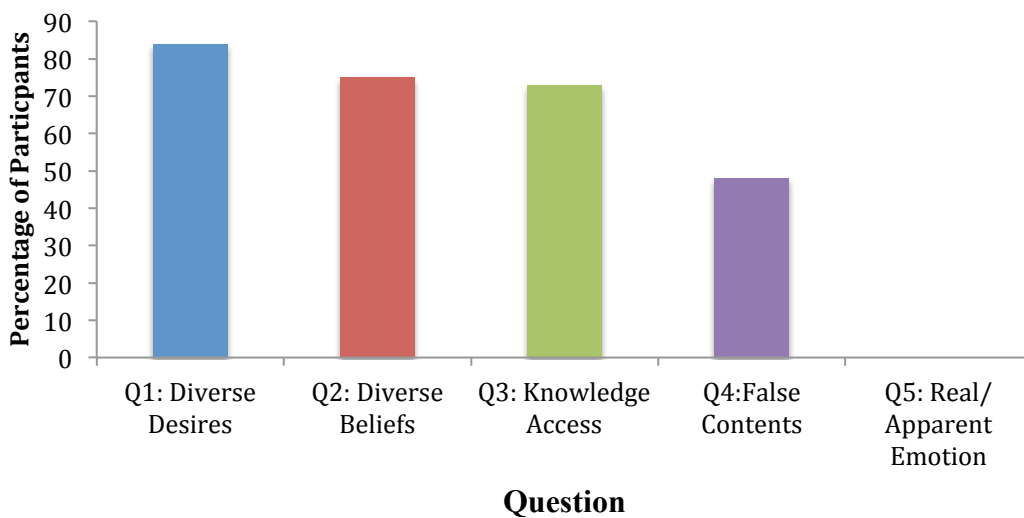


Figure 4.1. Percentage of participants passing each item on the mental state understanding battery.

Performance on items 1 through 3 were significantly above chance set at .5 for each: Q1 mean=.84,  $t(55)=6.9$ ,  $p<.001$ ; Q2 mean=.75,  $t(55)=4.1$ ,  $p<.001$  and; Q3

mean=.73,  $t(56)=3.9$ ,  $p<.001$ . Performance on Q4 (mean=.48) was not significantly different from chance. For question 5 performance was at zero.

As reported in the original paper (Wellman & Liu, 2004) performance on questions 2 and 4 were significantly different,  $t(1,54)=3.25$ ,  $p=.002$ , as were performance on questions 3 and 4,  $t(1,54)=3.23$ ,  $p=.002$ , and 4 and 5,  $t(1,54)=7.2$ ,  $p<.001$ .

For item 5, the real-apparent emotion question, 25 children answered the test questions correctly, however, only 4 children passed both memory check questions. Of the four children who passed the memory checks none answered the test questions correctly. Conversely only seven children gave incorrect answers to the memory question associated with item 3, and two in the case of question 4. In light of the coding instructions given in the original battery, and precedent in the literature, we chose to remove question 5 for the remainder of the analyses and resultant MSU score (Nathanson, Sharp, Aladé, Rasmussen, & Christy, 2013).

Table 4.1. *Table Showing Percentage Pass Rates for the 5-Items on the MSU Battery Scale*

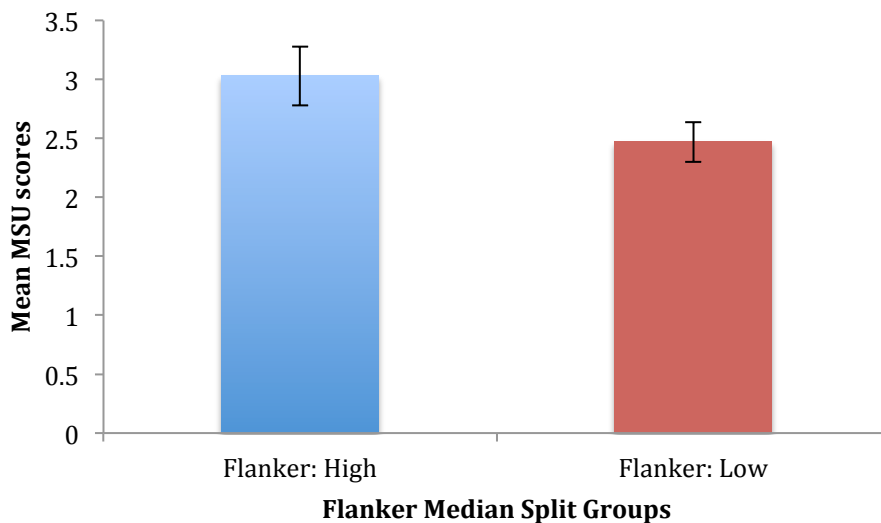
| Study<br>Authors                 | Mean Age &<br>Location | 5-Item Scaled MSU Battery |    |    |    |    |
|----------------------------------|------------------------|---------------------------|----|----|----|----|
|                                  |                        | Q1                        | Q2 | Q3 | Q4 | Q5 |
| Wellman & Liu, 2004              | 4;7 Years<br>N.America | 95                        | 84 | 73 | 59 | 32 |
| Peterson, Wellman &<br>Liu, 2005 | 4;6 Years<br>Australia | 95                        | 85 | 82 | 32 | 19 |
| Shahaeian, Peterson,             | 4;6                    | 95                        | 77 | 68 | 36 | 16 |

|                      |           |    |    |    |    |   |
|----------------------|-----------|----|----|----|----|---|
| Slaughter & Wellman, | Australia |    |    |    |    |   |
| 2008                 |           |    |    |    |    |   |
| Our Sample           | 4;4       | 84 | 75 | 74 | 48 | 0 |
|                      | British   |    |    |    |    |   |

---

#### 4.4.1. Mental State Understanding and Executive Functions

Due to the distribution of the EF scores described in section 3.4 a linear analysis would have been inappropriate. Instead a median split performed on the Flanker and DCCS data was used in a one-way ANOVA to assess whether those in the high group outperformed those in the low group.



*Figure 4.2.* Mean MSU scores of participants in the high and low Flanker median split groups. Error bars represent standard error.

Those in the high Flanker group performed better on the MSU battery, however, this relationship was only approaching significance,  $F(52)=3.5$ ,  $p=.066$  (see Figure 4.2.). No such relationships were seen when DCCS was used as the independent variable  $F(52)=.49$ ,  $p=.49$ . The largest difference between groups (for

both Flanker and DCCS) was seen in Q1 of the battery. Given performance being highest on Q1 this was not surprising, those who failed may have done so due to EF weakness.

Given these weak individual relationships we created a composite EF variable. We split subjects into 3 groups based on their overall EF scores on both items. In this Composite measure those who were in the Low group for both Flanker and DCCS scored 0, those who were High in one and Low in the other scored 1 and a score of 2 was given to those who were High in both. A pairwise comparison was conducted that showed significant difference on question one between the top and bottom EF groups ( $p=.024$  adjusted for multiple comparisons), although a one-way ANOVA was marginal,  $F(52)=2.8$ ,  $p=.07$ . Every child in EF composite group 2 answered Q1 correctly whereas only 71% did so in Group 0 (Figure 4.3.). As such we did additional against chance analysis of the three groups on each of the MSU items to establish if there was some underlying relationship between EF generally and MSU (Table 4.2. and Figure 4.3.).

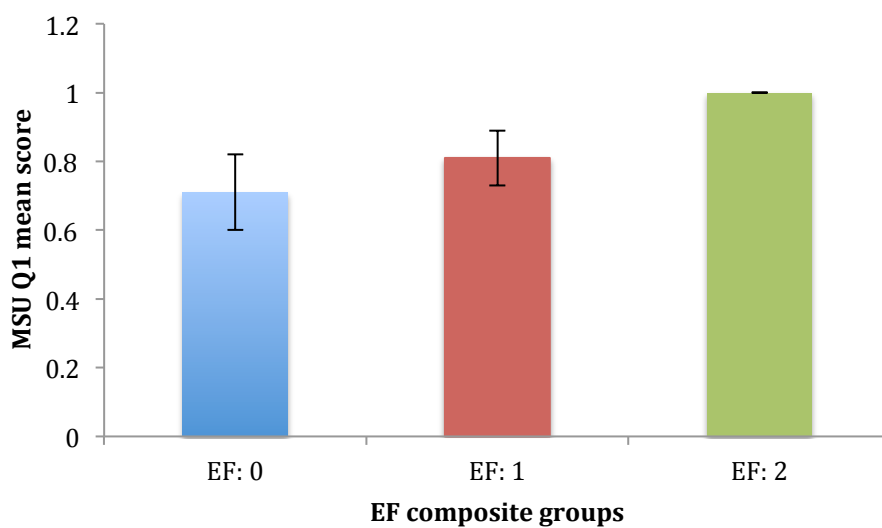


Figure 4.3. Mean performances on MSU question one in each of the EF composite groups. Error bars represent standard error.

Table 4.2.

*Mental State Understanding Performance, Compared to Chance (.5), Based on Composite Executive Functioning Scores.*

| Composite EF | 0          | 1             | 2             |
|--------------|------------|---------------|---------------|
| MSU Q1       | T(18)=2.05 | T(21)=3.53*** | 100%          |
| MSU Q2       | T(18)=1.81 | T(21)=1.58    | T(16)=2.24*** |
| MSU Q3       | T(18)=1.68 | T(21)=3.53*** | T(16)=1.751*  |
| MSU Q4       | T(18)=0    | T(21)=-.65    | T(16)=.49     |

\*p<.05, \*\*p<.01, \*\*\*p<.0085 (Bonferroni correction lowering p to .0085)

Participants in the lowest EF composite group did not perform above chance levels on any of the MSU items (Table 4.2.). However, children in the highest group performed above chance on items 1, 2 and 3. Performance on item 4 was at chance levels therefore as expected there were no significant between group differences.

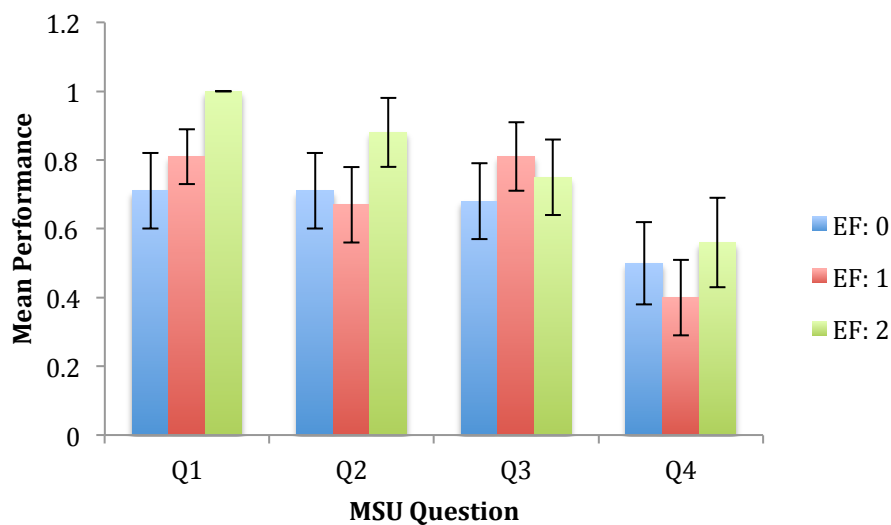


Figure 4.4. Mean performances on the mental state understanding battery items for each of the composite EF groups. Errors bars represent standard error.

#### 4.4.2. Mental State Understanding and Counterfactual Reasoning

There were no significant linear relationships between MSU and composite CFR performance,  $r=.05$ ,  $p>.05$  (see Figure 4.5.) and no between group differences in performance based on either variable. No trends were identifiable with performance showing only very weak positive correlations. As MSU is a form of conditional reasoning we had expected to find that BCR and CFR correlated with it.

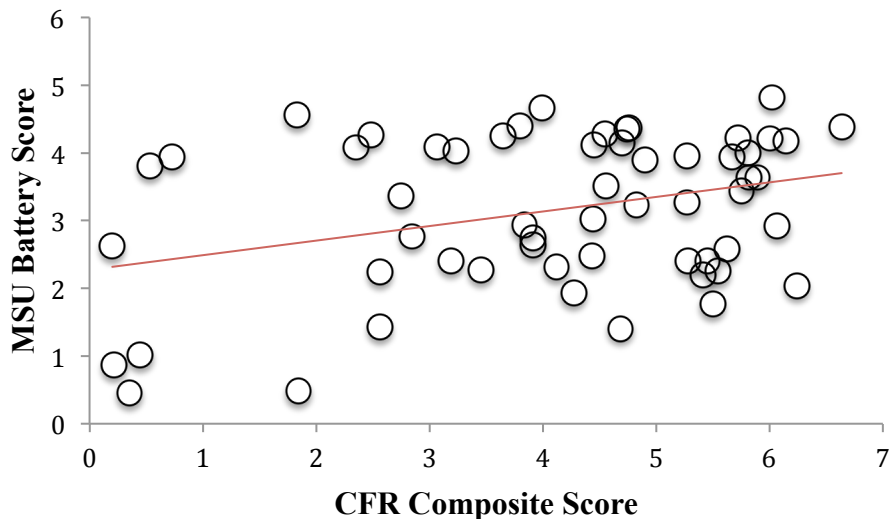


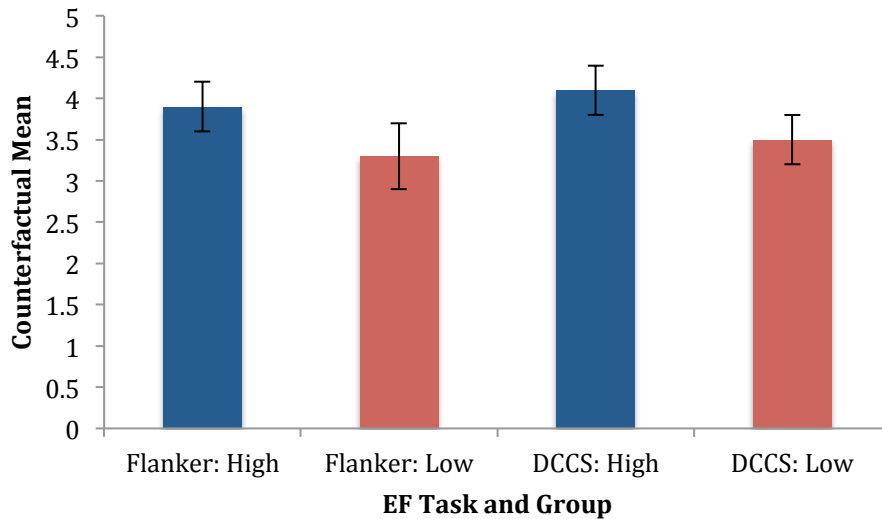
Figure 4.5. Scatterplot (including jitter) showing MSU battery scores and CFR composite scores.

#### 4.4.3. Counterfactual Reasoning and Executive Functions

As mentioned previously due to the distribution of the EF scores a linear analysis would have been inappropriate. Instead a median split performed on the Flanker and DCCS data was used in a one-way ANOVA. We then created a composite score for counterfactuals, aggregating participants' raw scores on the



standard Slide task and Story task. Despite performance being higher for both EF groups there was no significant main effect for the Flanker,  $F(54)=1.7$ ,  $p=.2$ , or the DCCS,  $F(54)=1.7$ ,  $p=.2$  (see Figure 4.6).



*Figure 4.6.* Mean Performances on the Composite Counterfactual Reasoning Measure for the High and Low Median Split Groups on the Flanker and DCCS.

Given these weak individual relationships and in order to better understand the relationship between EF and CFR, we split subjects into 3 groups based on their overall EF performance as in section 4.4.1. This allowed us to rank children in terms of their overall EF skill. Group 1 were on the lower end of the scale and group 3 on the higher. One-sample T-tests were then performed based on this composite variable to ascertain whether performance could be related to it (Table 4.3 and Figure 4.7.). This method of testing was also employed in Beck & Guthrie, (2011). We hypothesised that those in the high group would perform better on the CF tasks.

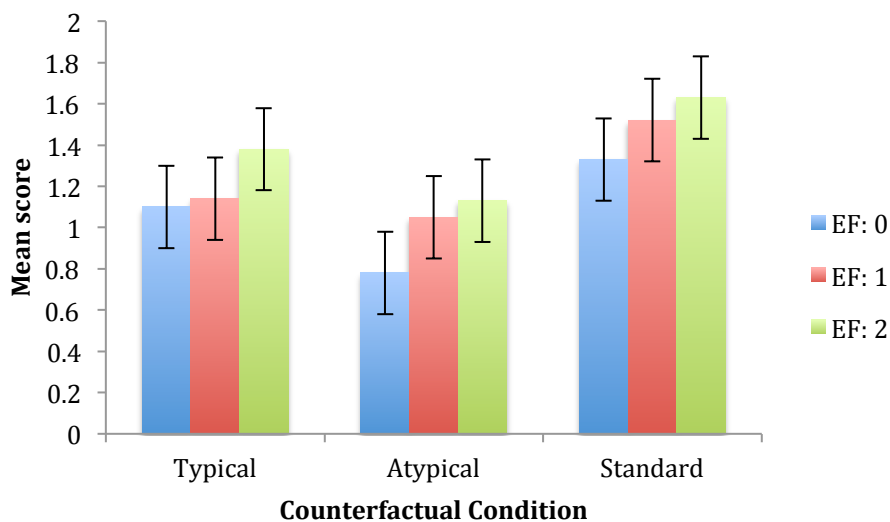
Table 4.3.

*Counterfactual Reasoning Performance, Compared to Chance, Based on Composite Executive Functioning Scores.*

| <b>Composite EF</b> | <b>0</b>    | <b>1</b>      | <b>2</b>      |
|---------------------|-------------|---------------|---------------|
| Stories Total       | T(13)=-.586 | T(20)=-.175   | T(19)=2.698** |
| Typical             | T(13)=.694  | T(20)=.252    | T(19)=2.629*  |
| Atypical            | T(13)=-1    | T(20)=-.623   | T(19)=1.751   |
| Standard            | T(13)=-1    | T(20)=4.24*** | T(19)=3.24*** |

\*p<.05, \*\*p<.01, \*\*\*p<.0085 (Bonferroni correction lowering p to .0085)

Visually the expected pattern emerged with performance being lowest for group 0 and highest for group 2 (see Figure 4.8). Those children in group 1 or 2 on the Composite EF measure scored statistically above chance in the Standard condition of the Slide task,  $t(20)=4.2$ ,  $p<.0085$  and  $t(19)=3.2$ ,  $p<.0085$ . Those in Group 0 performed at chance. Additionally participants in group 2 were the only group who performed above chance in the Typical condition,  $t(19)=2.6$ ,  $p<.05$ .



*Figure 4.7.* Mean performances for each EF composite group on the three counterfactual conditionals. Error bars represent standard error.

#### 4.5. Discussion

The first aim for this chapter was to characterize our sample's mental state understanding abilities. As seen in the original and subsequent papers detailing the battery, performance declined across the items (Peterson et al., 2005; Shahaeian et al., 2008; Wellman & Liu, 2004.) We broadly replicated the findings of Wellman and Liu (2004) with earlier questions being significantly harder than later ones except in the case of items 1 and 2. The majority of participants answered entirely in the expected order (70%). When compared with the other English-speaking samples reported in Table 4.1 our participants were slightly lower than expected on items 1 and 5 in particular. This being said our sample had the youngest mean age of the four reported studies. Furthermore our children were all the same age, the other studies reported data from a sample varying between the ages of three and five years. For our participants over all performance was good on the first three items, Diverse Desires, Diverse Beliefs and Knowledge Access. The False Contents and Real/Apparent Emotions questions, however, were, as expected, significantly more difficult.

The second aim we had set out was to investigate relationships between Executive Functions, Counterfactual Reasoning and Mental State Understanding. Although there was no overall significant main effect between EF and CFR or MSU, the group performance breakdowns (Table 4.2. and Table 4.3.) clearly show a trend for children with higher EF scores to be those performing to a higher standard. Children in the highest EF group performed above chance in MSU items 1-3 while those in the lowest group failed to perform above chance levels in any of the items. A similar story emerged with EF and CFR. Children in the highest performing EF group scored above chance in the typical stories and standard slides tasks whereas those in the lowest EF group did not outperform chance on any of the CFR measures. This

pattern of results suggest that EFs, despite being necessary for the planning and execution of goal directed actions, and although also measured through the application of if-then rules, were not sufficient or wholly responsible for conditional reasoning about the actual versus the possible. This is conditional reasoning measured both through counterfactual situations or mental state reasoning tasks.

Finally speaking to our third aim, and in contrast to previous research in which positive MSU and CFR relationships were reported (e.g., Guajardo & Turley-Ames, 2004; 2009; Riggs, Peterson, Robinson & Mitchell, 1998), our data did not show any positive relationship between performances on the two tasks.

#### **4.5.1. The Measurement of MSU**

In our study we used a battery to assign a MSU score to each child where as in the aforementioned studies only classic location change false belief tasks were used. We chose to measure MSU (as opposed to simply false belief) for several reasons. We wanted to generate an individual differences score that reflected children's mental state understanding as a whole, and was not restricted only to the understanding of false beliefs. We also wished to use a scale or measure that had precedence in the literature and that we could be confident was a reliable measure of MSU. Both Riggs et al. (1998) and Guajardo et al. (2004; 2009) used false belief items exclusively. The Wellman and Liu battery did not contain a false belief item of the same structure as utilized by the other authors. Instead a false contents task that led to the protagonist having a false belief from the outset, was used in our study. The nature of this false contents task is strikingly different to that of the classic false belief due to the lack of new information being integrated to generate the correct answer. The false contents task does not involve deception in the same way as the false belief task and as such may rely on different cognitive processes. It could be that in retaining the extra

deception element the Sally-Ann false belief type tasks become more similar to a counterfactual task in having gone beyond a simple false belief through the addition of another layer or piece of information regarding how the belief came to be false. In the Sally-Ann paradigm three perspectives of one world are involved, the participant's reality, Sally's reality and Ann's reality which all contrast but have the same temporal qualities. In the false belief paradigm some event has changed the situation, namely Ann's moving of an object. In the false contents task there is only the participant's reality and the protagonists' perspectives of reality contrasting, no change or alteration of the world has occurred, therefore it may be more readily solved.

In counterfactual tasks there is a similar need for a change or updating of reality, in this case into a counterfactual alternative. In both CFR and the false belief task the actual world reality, and then updated counterfactual/false belief world realities are contrasting. The counterfactual world, unlike the false belief world, however, must be generated based on a hypothesized, rather than observed change.

So, despite having the same number of representations, the false contents task and counterfactual story task we used have differing levels of demand. The counterfactual reasoning items require the generation of an alternative world due to some change having been made, whereas the information about the false contents world is more readily observable and remains static; it is only the perception of the world that differs, not the world itself. Furthermore there are no additional temporal demands made during MSU tasks: although there are different representations of the world they were all generated at the same time and about the same time. For CFR tasks this added level of representational difference, a temporal difference, may be contributing to the decreased performance

In addition to both of their tasks requiring updated or changed representations Riggs et al. (1998) matched the materials used to test both false belief understanding and counterfactual reasoning in their sample. This meant that they shared the same content and structure while varying only in the type of question asked of the participants. The strong, shared variance reported then may have been aided by these purposefully similar task demands. Riggs et al. (1998) highlighted the shared counter to reality beliefs required for children to be successful. The classic Sally/Anne task is much more similar to a counterfactual task than the MSU questions used in our study for one crucial reason. In the Sally/Anne tasks experimenters are essentially asking children to reason about Sally's belief as if Anne had never been present. However, in the Wellman and Liu battery there was only one MSU question that really required counter to reality beliefs: the false contents task (item 4). Here children had to reason about a protagonist that had never seen inside the distinctive bottle. In this case there was a direct contrast between what a blind protagonist would expect and reality. Item four alone, however, still showed no relationship with CFR performance.

#### **4.5.2. The Measurement of Counterfactual Reasoning**

Guajardo & Turley-Ames (2004; 2009) utilized a very different method of assessing counterfactual thinking ability than the current study. In their tasks children were asked to generate alternate antecedents to change an outcome rather than generate alternate consequents in light of changed antecedents. Children were given scores then based upon how many alternative antecedents they could produce rather than being challenged to answer correctly about the causal outcome of some imposed change. This is a much more creative measurement variable in that children were not being asked to identify one correct causal consequent but instead generate many possible alternative antecedents. This type of generative measure has in the past been

related to primed differences in creativity (Markman, Lindberg, Kray & Galinsky, 2007). The processing mind-set required to generate multiple alternatives is relational as opposed to causal and may rely on largely divergent cognitive abilities.

#### **4.5.3. Limitations**

As previously discussed in section 3.5.1 having EF measures that yield more normally distributed data would have been beneficial for us. Being able to look for a linear relationship between our variables was what we had hoped for rather than reducing the data into high versus low groups for both.

It would have also been beneficial for us to have included a classic false belief item in our investigation into the mental state understanding of our sample (as in Aschersleben et al., 2008). In doing so we would perhaps have been able to draw more conclusions about the relationship between counterfactual reasoning and mental state understanding. The inclusion of a false belief task would have allowed us to explore the possibility that the structure of the question alone may account for previously reported relationships between CFR and MSU.

#### **4.5.4. Conclusions**

In conclusion we did not find the expected relationship between CFR and MSU. This was surprising due to both tasks requiring conditional reasoning abilities and the demands of both being conceptually similar in requiring navigation between the actual and the possible. Both MSU and CFR require children to select an answer not the most salient or personal to them. They must generate and reason from alternative representations of the same, or similar scenarios and they must answer in accordance with a perspective that is not currently their own.

Despite these structural similarities then children must be drawing on some other resources rather than those shared by the tasks from a processing point of view.

MSU and CFR rely on strikingly different types of content knowledge. MSU requires the possession and use of knowledge surrounding the mental states of others and how causal situations may alter this mental state. Counterfactual knowledge also requires causal reasoning knowledge but not in relation to mental states (in this study at least). CFR further challenges children to reason in line with the nearest possible world, that is children must answer CFR questions departing from reality only where necessary. Children often must integrate multiple pieces of information when reasoning counterfactually and answer inline with what could have been. This requires children to disengage from temporal reality. Mental state understanding questions may in some cases be answered in a basic conditional way in that generally people will behave in accordance with their beliefs and desires. MSU questions are asked in a current or future-hypothetical way, so do not place the same temporal demands on the reasoner but they are required to disengage from an egocentric perception of reality. Both CFR and MSU then require disengagement and shift. However, one requires a shift to a contrasting subjective, mental perspective of the same situation and contingent on the same causal rules whilst the other requires a shift to a different temporal representation and therefore potentially also a different set of causal relationships related to the same situation. CFR then could be argued to be more difficult by virtue of the world to be reasoned about being altered and therefore possibly subject to or influenced by different causal rules. During MSU tasks the situation remains the same, it is just the perspective that is altered.



## **Chapter 5: Standing on the Shoulders of Giants: The Development of Imitation and its Roots in Social Cognition**

### **5.1. Abstract**

One of our main sources of information about the world are the other intentional agents who inhabit it: we do not make decisions in a vacuum (Zimmer, 1986). Imitation refers to the replication of another agent's goal directed actions, and is crucial for the dissemination of innovative problem solutions (Mason, Jones & Goldstone, 2008). Observing and adopting the strategies that others deploy to actively change the world around them can prove a cost effective choice. Successful and useful imitation relies on understanding the goals of any behavior, and the intentions behind them: imitation inherently requires mental state understanding. Imitation is another a form of conditional reasoning that is necessarily grounded in the past, but also impacts upon the future. A unique feature of experimental imitation tasks, when compared to the tasks reported earlier in this thesis is that children's responses are unconstrained and they are not asked to indicate a logically reasoned choice: they imitate (or not!) any actions they wish. As such exploring children's imitation will allow us to compare how they deal with reasoning about the actual versus in the possible in such an unconstrained setting with how they did so in earlier tasks where there was a right and a wrong selection to be made. This chapter will focus on the imitation of actions although the same principles can be applied to the imitation of speech as explored in Chapter 6. In this chapter experimental imitation data will be reported to address the following empirical questions: (i) Can we find evidence of faithful imitation in our sample? (ii) Are the patterns of imitation observed sensitive to causal information? (iii) Do imitative strategies vary across objects with different

features? And (iv) Is imitative performance associated with conditional reasoning and mental state understanding?

## **5.2. Introduction**

### **5.2.1. An Introduction to Imitation**

In natural language the terms imitation and mimicry are often used interchangeably. In the psychological literature, however, the distinction between imitation and mimicry is a crucial one: imitation refers to the replication of behaviors interpreted as goal directed, while mimicry can refer to phenomena such as facial expression matching or other automatically replicated behaviors such as gestures. Imitation has been described as one of the “most critical components of socio-cognitive development”, (Zmyj, Daum, Prinz, Nielson & Aschersleben, 2012). As such, a vast amount of research has been undertaken to identify when children first engage in this behavior labeled a “powerful learning mechanism” (Over & Gattis, 2009). The beginning of this development is seen in first few days of life when infants are reportedly able to mimic the facial expressions of their caregiver (e.g., Meltzoff & Moore, 77; 83; 89; Nagy, Pilling, Orvos & Molnar, 2013).

With age, the type and scale of behaviors that children are able to copy increases. Imitation goes beyond simply mimicking those around us. Early facial and vowel sound imitation relies on no understanding about why these behaviors are being executed. For more complex goal directed imitation a recognition that those behaviors being carried out by another person can also be carried out by one’s self is needed: an individual requires self-other equivalence to imitate in a causally meaningful way (e.g., Buttelmann, Carpenter, Call & Tomasello, 2013; Matheson, Akhtar & Moore, 2013; Meltzoff, 2007). Furthermore mental state understanding is crucial to imitation as the goals and intentions of the model are central to the decision

of what to imitate (Meltzoff, 1995). Children will go far beyond what is causally necessary in some circumstances to align with the both the goal, and crucially the intention of a model (Gergely, Bekkering & Kiraly, 2002).

### **5.2.2. The Beneficial Effects of Imitation**

We are interested in the development of imitation and its relationships to other cognitive abilities because imitation is such a highly adaptive strategy for the transmission of knowledge between individuals. Imitation is a smart strategy, in a way similar to counterfactual reasoning, as it reduces the requirement for personal, explicit trial and error experience of the world. Imitation allows children to stand on the shoulders of giants. Children need not be innovative in every endeavor they undertake. In some situations they need only observe, represent and redeploy those motor schemas used by others in the same or similar future situations. When faced with a problem, such as how to break into the chocolate cupboard at nursery, they need only spy on their peers. When challenged to write their first few tentative letters, they need only copy their teacher (not to devalue their achievement!). As they get older and wish to bring about psychological changes they may slam doors as their older siblings do, or take themselves off to their room. Of course not all behaviors that are imitated are so physically large or obvious. Children imitate the voices from their favorite cartoon to garner a laugh, or imitate the annoying intonation or colloquialisms of their peers. Imitation allows children and adults alike to play at being someone else, sometimes in an obvious way when parodying a character, but often in a subtle but functionally useful way. The experiences children have of seeing others “do” make their own attempts more likely to succeed. They get a fast track pass to skills that may otherwise elude them for many minutes, days or longer to

come. Modeling our behaviors on the successes of others is a highly adaptive and freely available resource that more often than not will benefit us more than it costs.

A critical aspect of imitation, however, is that imitation is a learning tool but it also embodies the navigation between the actual and the possible that we are interested in. When imitating children are conceptually behaving as if they were someone else. Children who imitate are effectively reasoning along the lines of ‘If I were you I would do this’ and use this process to bring about a possible world based on an actual one. Imitation contains a duality much like the one that shall be introduced with respect to language later in this thesis. The strategy of imitating actions is a tool for change in the world and requires the navigation between actual and possible worlds, the world of the self and the world of others, but it also embodies this combination of the actual and the possible as imitative acts amalgamate elements of both imitation and innovation.

The act of imitation, combining the actual and the possible, has the potential for providing huge benefits to young children. Children can use imitation to support their own learning about the world and their decisions about who to imitate reflect this learning based benefit. In the case of a behavior or action being performed that is novel to a participant, the characteristics of a model relating to their reliability or perceived knowledgability become important (e.g., Elekes & Kiraly, 2013; Lampinen & Smith, 1995; Jaswal & Neely, 2006; Zymj, Buttelmann, Carpenter & Daum, 2010). Children will imitate different people at different times depending on the type of behavior they are engaging in. When novel actions are being performed children will more often imitate an adult than a peer (Zymj et al. 2010), unless said adult has been previously portrayed as an unreliable source of evidence (Jaswal & Neely, 2006). The objects being explored are also often critical in determining behavior. When young

children are interested in learning about toys they ask their peers, however, when they require information about food they instead seek information from adults (Van der Borgh & Jaswal, 2009). Children are flexible and selective in terms of whom they approach to gain the best information relating to a given situation: children are smart imitators.

Imitated behaviors have the opportunity to become incorporated into their learners' repertoire and have the potential to be redeployed flexibly from then on (Wood, Kendal & Flynn, 2013). Wood et al. (2013) presented evidence that children are able to incorporate both novel and imitated strategies for manipulating a novel piece of apparatus into their behavioral repertoires. Adults too often converge on one type of problem solution even when other equally valid options are available, they jump on the bandwagon (Mason et al., 2008). This further solidifies imitation's position as a powerful, flexible and adaptive learning mechanism, but one with a decidedly social foundation.

From an instrumental point of view the act of imitating gives children the opportunity to generate representations of new behaviors through seeing them performed by someone else. Furthermore through subsequently carrying them out an additional opportunity to generate and strengthen such representations arises. Imitation aids the retention of new information as the imitator generates not only a memory of observing the behavior in question, but also of having carried it out. If they did not imitate, and only observed actions, the strength of the representation would not be as robust, and perhaps not support future use as competently or flexibly. Manually undertaking an imitated action or behavior makes it more likely to be retained and used in the future.

Imitation is not a transient phenomenon that only impacts upon, or is limited to a current state of affairs: imitation is not bound only to the time and place directly following a model (Simpson & Riggs, 2011). Imitating an action once allows it to be used to alter physical and social worlds in future. Skills learnt through imitation may be productively deployed elsewhere if analogies are made between different situations. Having a motor schema in place for opening one type of container, may allow a slightly different container to be opened through using the same or similar pattern of behavior for example. Flexibility of this sort is one of the hallmarks of mature cognition (Zelazo et al., 2003). Imitation allows children to build a repertoire of functional behaviors that they can use flexibly in a variety of situations in the future. Not only this, but children can begin to cognitively simulate or mentally rehearse how these imitated behaviors may be used to manipulate other possible worlds, not only worlds they are currently experiencing. With a greater store of possible behaviors comes a greater number of real world possibilities when it comes to imitative contexts.

### **5.2.3. The Paradox of Faithful Imitation**

Actions involving objects, particularly those that incorporate distinct goals on the part of the demonstrator, allow children to be selective in what they imitate. Infants begin their imitative journeys by imitating the goal of a behavior, choosing to copy the outcome of a behavioral sequence. Selective imitation refers to situations where only those actions that are causally necessary to achieve the end goal are undertaken while any additional intentional acts performed by the model are omitted. This form of selectivity is akin to CFR in that children appear to have reasoned about which aspects of a past sequence were causally necessary (either instrumentally or conventionally) to achieve the goal state, and used this inferred information to plan

and execute their own subsequent behaviors. This ability to imitate rationally is a strikingly intelligent strategy, in terms of providing economy: energy is only expended on causally necessary actions.

Many studies have investigated how sophisticated or selective the imitation that young children engage in is (e.g., Brugger, Lariviere, Mumme & Bushnell, 2007; Hilbrink, Sakkalou, Ellis-Davies, Fowler & Gattis, 2013). In a study of 14 to 16-month-old infants the causal necessity of actions in the context of the desired end states influenced how often they were imitated (Brugger et al., 2007). In a similarly structured study the imitative patterns of 12 to 15-month-old infants were observed (Hilbrink et al., 2013). Participants were shown goal-directed actions using two novel toys (Hilbrink et al., 2013). At both ages the selectivity of the imitation was driven by the physical necessity of the actions, that is, children were more likely to imitate actions on objects that were causally necessary to the outcome than actions on the objects that were not causally instrumental. However, this selectivity decreased with age and a corresponding increase of levels of faithful imitation.

Faithful or over-imitation directly contrasts with selective imitation in that intentional actions are replicated that go beyond those instrumentally necessary to achieve the modeled goal. Both causally necessary and unnecessary actions are imitated. Nielsen (2006) documented children from 18-months onwards becoming more faithful in their imitative styles. Furthermore several publications have reported that children between 3 and 5 years old, and adults, faithfully copy the actions of models even in conditions when causal information is readily available to them (McGuigan, Whiten, Flynn & Horner, 2007; McGuigan, Makinson & Whiten, 2011). This faithful imitation phenomenon is pervasive and readily observable in young children. Hilbrink et al. (2013) found the move from selectivity to faithfulness

occurring sooner than other authors who had reported this intriguing, and intuitively paradoxical phenomenon previously. Where selective imitation is rational and cognitively efficient, faithful imitation seems cognitively wasteful in comparison.

When children faithfully imitate they go beyond imitating solely the instrumental goals of an action sequence. This in some cases appears economically wasteful and rather immature as a learning strategy. However, the imitation of non-causal features of behavior is pervasive in our daily lives and humans have been shown to imitate many different forms of behavior in this way. Literature on adults reveals that posture, gestures, prosody and syntactic structures are often copied in the absence of any concrete causal justification for doing so (Chartrand & Bargh, 1999; Niedenthal, Barsalou, Winkielman, Krauth-Gruber & Ric, 2005).

In experimental settings with a researcher demonstrating a multitude of actions children have been shown to imitate; the operation of puzzle boxes; the search locations of hidden objects; the movement styles of soft toys; the end location of toys, and; hand movements and tool selections amongst others, (McGuigan et al., 2007; Carpenter et al., 1998; Bertenthal & Boyer, 2009; Bekkering et al., 2000). Indeed children seem to be willing and able to imitate, or at least attempt to imitate almost anything, even those things that do not achieve any readily discernable physical, psychological or social causal goal.

The development of faithful imitation in early childhood mirrors the progression of MSU. Children begin to move from understanding only desires i.e. regarding a goal or outcome, to understanding that intentions and beliefs are also involved in the behaviors being undertaken i.e. regarding method or style. Over the following years this tendency to over imitate persists and can be observed in many cultures. 3 to 5-year-old pre-school children in both Western and rural African



communities have been shown to faithfully imitate actions even when there were able to identify which elements were causally unnecessary for goal completion (Lyons, Young & Keil, 2007; Nielsen & Tomaselli, 2010). Children possess the relevant causal knowledge but go beyond it when making decisions about their own behavior.

McGuigan et al., (2007), used puzzle boxes that were either clear or opaque, with both relevant and irrelevant actions being modeled, and models being either live or video-taped. 3 and 5 year olds imitated faithfully both relevant and irrelevant actions regardless of the causal information observed in the live demo. Only 5 year-olds and not 3 year-olds, however, continued with such imitation when a degraded video demonstration was observed. Five-year-old participants showed more faithful imitation than their three-year-old counterparts. The hypothesis that this faithful imitation increases over time obtained more support when a later study found that adults imitated task elements even more faithfully than older children (McGuigan et al., 2011). Two groups of children, (aged 3 and 5-years old) and one group of adults (mean age= 42-years) were given the same puzzle box and modeling procedures. Both causally relevant and clearly irrelevant actions were modeled. The adult group engaged in the highest number of irrelevant action replications, with the 3-year old group engaging in the least.

Evidence from multiple sources then identifies that faithful imitation as a strategy is becoming more entrenched with age (Hilbrink et al., 2013; McGuigan et al., 2007; 2011). The video evidence further suggests that for younger children a stronger social input is required to support faithful imitation of this sort, degraded videos elicited less imitation from the younger children when compared to the older children (McGuigan et al., 2007). Such evidence points to an arrangement of imitative goals emerging along a developmental progression such as the following.

Imitation -> 1<sup>st</sup> physical outcome

-> 2<sup>nd</sup> intentional action

-> 3<sup>rd</sup> social concerns

Children first imitate physical outcomes, then intentional actions and finally any actions deemed socially relevant. But why does faithful imitation emerge at all? The theoretical beneficial effects of imitation were discussed at length in section, 5.2.2. The reasons for this now extensively documented phenomenon, however, are not agreed upon amongst researchers. Some authors have argued that with experience faithful imitation emerges as such an adaptive human learning strategy that it may be employed even at the expense of task efficiency (McGuigan et al., 2007). McGuigan et al. (2007; 2012), suggest that faithful imitation is an example of the over-extension of a “highly adaptive conformist bias” alongside a “model-based bias”. Their account suggests that humans increasingly rely on an automatic strategy of imitation when learning new skills. Evidence that adults robustly over imitate even when the structure and goals of an action sequence are fully transparent supports such automaticity accounts.

Lyons et al. (2007) in contrast suggest that faithful imitation is the result of poor physical understanding on the part of young children. They claim this poor physical understanding results in a proclivity to encode all adults’ actions as causally necessary. However, given the evidence that even adults, with assumed high levels of causal understanding, engage in copious amounts of irrelevant imitation such a claim seems comparatively weak.

Nielsen (2006) emphasized the abundance of evidence supporting the view that social motivation is the key prompt for over-imitation. He views faithful imitation as the result of children becoming less interested with the instrumental or logical

structure of actions but more interested in the social experience they promote. Such a movement from cognitively-efficient selective imitation at 12-months, to an increasingly faithful pattern at 15-months in Hilbrink et al. (2013), lends support to this instrumental -> social explanation. Indeed Hilbrink et al. (2013) found that this emerging social motivation was predicted by the extraversion scores collected from their sample. Such studies link well with evidence pointing to a strong motivation to share emotional states, goals and perspectives from aged 12-months and onward (Tomasello, Carpenter, Call, Behne & Moll, 2005). Infants suddenly become motivated to create, and maintain social co-ordination with a peer rather than just learn new skills: imitation is one simple way to achieve this (Eckerman, Davis & Didow, 1989). Imitation is a highly adaptive tool and can be used in many contrasting social situations. It can be used to communicate likeness, identify with a social partner and repair social experiences (e.g., Over & Carpenter, 2012; Spoor & Williams, 2007; Zmyj & Seehagen, 2013). Retaining the instrumental features of physical actions in a selective way makes sense, as the affordances of an action are universal, whilst the social or conventional affordances of the same actions are not so reliably pervasive: faithful imitation is not always a smart choice. Just because your mother tucks your napkin into your sweater does not mean this is the appropriate way to use a napkin whilst in a Michelin\* establishment! The correct way to hold your knife and fork, however, is more stable and context independent.

What is clear, however, is that over time the actions instrumental to a physical goal's attainment are retained in a more prominent or accessible position within the hierarchy of goals than the intentional, social features of an action. It may be that these conventional actions are forgotten overtime having been encoded in a less robust manner as suggested by Simpson and Riggs (2011). Conversely it could be that

the conventional, irrelevant actions are suppressed: either way these two types of affordances are treated differently by children even at a young age.

Although lower down the hierarchical chain than those actions deemed physically necessary, the fidelity of imitation of non-causal features has been shown to be readily influenced by a variety of social mediators in 18, 24, 48 & 60-month-old participants (Nielsen, 2006; Nielsen & Blank, 2011; Simpson & Riggs, 2011). This is largely unsurprising given the intrinsically social roots of imitation. The types of social manipulations that have been tested experimentally fall broadly into three categories: (1) the model's social behaviors (e.g., Bandura, 1971; Brugger, Lariviere, Mumme & Bushnell, 2007; Hartup & Coates, 1970; Nielsen, 2006; Nielsen, Simcock & Jenkins, 2008); (2) the model's characteristics (e.g., Buttelmann, Zmyj, Daum & Carpenter, 2013; Jaswal & Neely, 2006; Kinzler, Corriveau & Harris, 2011; Ryalls, Gul & Ryalls, 2000; Seehagen & Herbert, 2011; Zmyj et al., 2012) and finally; 3) increasing goal salience through social means (e.g., Brubacher, Roberts & Obhi, 2013; Chen & Waxman, 2013; Elsner & Pfiefer, 2012; Hermann, Legare, Harris & Whitehouse, 2013; Simpson & Riggs, 2011). Globally speaking, increasing social engagement, fostering social affiliation and increasing goal salience through these channels increases faithful imitation. Social processes, however, can only work alongside cognitive processes to produce imitative behaviors: there are a great many cognitive demands placed on young reasoners when engaging in imitation. That being said there are also a great many cognitive rewards.

The distinction between selective and faithful imitation is a crucial one for our purposes: selective imitation sees only actions that are causally necessary elements of a demonstration being repeated in contrast with faithful imitation which describes situations where causally both necessary and unnecessary components are integrated

into a reproduced sequence. This thesis is interested in exploring how children move backwards and forwards between representations of the world, and in time, when reasoning about, and interacting with the world. Imitation then, inclusive of both contrasting forms defined above, has two striking features that make it a particularly appropriate setting from whence to investigate children's navigation through their physical and social worlds. (i) Imitation is a strategy or tool that can be used to manipulate world. Through imitating an action or behavior children can impact upon the world in a similar way to the model from whom they extracted the necessary information. (ii) Imitation is paradoxical in that, although by its very nature there is a replication of the actions and/or behaviors of others, there is also always a greater or lesser extent of innovation or deviation from the model. In the case of selective imitation this independence from the model is very clear through the elimination of some elements. In the case of faithful imitation too, however, individuals will rarely exhaustively copy all aspects and features of the modeled behavior. Examples of unreplicated features could include the speed, motion or precise manner of an action.

Children have an almost limitless number of physical options when it comes to interacting with a new object. However, having the opportunity to extract some salient possible options from the actual world in order to influence future possible worlds is invaluable. Conversely if children were restricted or constrained in their imitative abilities, i.e., unable to innovate and use only previously observed actions, they would not be able to successfully generalize and apply their newly acquired physical skills to novel or unexpected situations. It is clear then that imitation is a multifaceted and highly adaptable behavior whose study affords an intriguing insight into socio-cognitive development.

#### **5.2.4. The Centrality of Goals and Intentions to Imitation**

Imitation occurs across a diverse range of situations and in response to numerous types of model. Mentioned earlier was the distinction between selective and faithful imitation and a move from one to the other being reported early in childhood (Hilbrink et al., 2013). To understand this move toward becoming more faithful imitators we must look how and why we might imitate anything in the first place.

In order to imitate children must parse action sequences, chose what elements to retain or disregard and reconstruct these in order to generate their own attempts (Bekkering et al., 2000). Once children's ability to pick action components from a stream of behaviors is in place (Baldwin & Baird, 1999), and children are able to interpret actions as goal directed around the age of six months (Bertenthal & von Hofsten, 1998) they can begin the process of selecting what to imitate.

The intended goals of the action or actions are suggested to modulate this process: imitation is a fundamentally goal directed pursuit (Bekkering et al., 2000). Behavioral re-enactment paradigms like that of Meltzoff (1995) provide striking evidence for this goal directed theory of imitation. Imitation is not supported only in situations where physical goals are achieved, but also when they are intended. 18-month-old infants imitate the intended goal of animate models even when that goal is not actually achieved by the model (Meltzoff, 1995).

If multiple goals are involved in an action sequence then these are argued to be organized in a hierarchical fashion: the most salient goal is given most importance and is as such more likely to be imitated (Bekkering et al., 2000). In a participant set of 4-6 year olds Bekkering et al. (2000), showed that in an imitation task involving hand movements there were more reproduction errors in situations where the goals were less salient (touching a particular ear with specific unimanual contralateral

movement) than in situations where the action itself was more salient (both hands being used to touch own ears using a contralateral movement). It seems more difficult to satisfy multiple goals when one goal is more salient, or there is competition between many goals. Children do not struggle to remember multiple modeled actions, rather they struggle to identify, represent and replicate multiple goals simultaneously.

Certain types of goal are consistently more salient than others and this supports their preferential imitation over other, less salient options. There are two classic paradigms that exemplify this hierarchical organization of goal theory with young children, the mouse-house (Carpenter, Call & Tomasello, 2005) and, the head-touch (Gergely, Bekkering & Kiraly, 2002). The mouse –house paradigm sees an experimenter hopping a small felt mouse with accompanying sound effects from one location to another. In one condition this end location was marked with a house, in the contrasting condition there was no house marking the end location. 12 to 14-month-old infants copied only the trajectory of movement when a house was marking the end location (Carpenter et al., 2005). They copied both the manner and sound of the movement in addition to the trajectory when no house was present. Children were not struggling to remember the manner of movement when a house location marker was present, they were just failing to incorporate and replicate all intentional goals due to the increased salience of one over the other.

In the head-touch paradigm a model switched on a lamp using their head. At fourteen months infants copied this head touch method more frequently when the model had used their head despite their hands being free to do so (Gergely et al., 2002). In a condition where the experimenter's hands were not physically free during the model participants were almost 50% less likely to use their head when switching on the light. Much like the behavioral re-enactment paradigms mentioned earlier

children copied the intended physical goal of the model, and not their manner or execution unless there was an intentional choice to do so despite another more physically affordant option being present. When combined these paradigms make for dramatic evidence that even in their first year of life children represent actions in terms of goals and intentions with spatial goals being replicated and conventional/non-causal intentional goals being present, but less-salient when in competition.

Additional evidence for this spatial versus conventional goal dependent hierarchy of imitation can be found in older children (3 to 6-years) and using different research methods (Gleissner, Meltzoff & Bekkering, 2000; Perra & Gattis, 2008; Williamson & Markman, 2006).

When children get older, however, their propensity to faithfully imitate both the goal, and the steps demonstrated to achieve it, increases (Hilbrink et al., 2013). Children become better able to replicate multiple goals. In an extension of Gergely et al. (2002), Nielsen (2006) reported that both 18 and 24-month-old participants (unlike the 12-month-olds) faithfully imitated the action sequence of the head touch, regardless of whether a logical reason to do so was provided. In this case it was not only the goal that precipitated imitation, but also those unnecessary elements that preceded its completion: children truly engaged in faithful-imitation. Faithful imitation then, although in some ways appearing to be a rather clumsy, one size fits all tool becomes increasingly common through the pre-school years. Children imitate a whole host of physically unnecessary actions that are undertaken intentionally by a model. They replicate actions in pursuit of more ambiguous, abstract intentional goals in addition to more concretely observable physical ones.



In the case of faithful imitation, unlike selective imitation, both instrumental and seemingly incidental goals are represented and replicated. This requires both causal knowledge and mental state understanding working alongside one another concurrently. The ability of attributing intentions has been firmly established as fundamental to modulating imitation (e.g., Bekkering, Wohlschlager & Gattis, 2000; Carpenter, Akhtar & Tomasello, 1998; Carpenter et al., 2002; Over & Gattis, 2009). This ability is very early to appear in development. When infants are able to attribute intentionality to a demonstrator then they are able to represent their actions and behaviors in terms of goals and intentions as discussed earlier. In this way intentional goals start to be incorporated into the hierarchical structure proposed by Bekkering et al. (2000) alongside the more readily identifiable physical goals. Intentional goals may be social, conventional or even regarded as playful.

In the first year of life infants hone these necessary intention reading abilities and hence their ability to represent intentional goals more robustly. This drives an increase in faithful imitation as an increasing number of goals can be represented concurrently. This ability can be bolstered by increasing the salience of the intentional actions in numerous ways such as vocal cuing (Carpenter et al., 1998), novel word usage (Chen & Waxman, 2013), or intentions being identified prior to actions taking place (Carpenter et al., 2002). An understanding of mental states, in this case the intentions/desires of a model, is intrinsic to goal directed imitation and highlighting these mental states increases their salience and subsequently their likelihood of being replicated.

Such socio-cognitive and attentional skills being involved in imitation as discussed above go beyond the low level associative processes proposed by some to account for imitation (e.g. Heyes, 2001). If we were simply organisms with pre-potent

dispositions to replicate the behaviors we see, then we would be engaged in never ending imitation at the expense of any innovation. This is of course not the case as we have seen in previous sections, even the youngest children are selective in what they imitate. Firstly children replicate only the instrumental, physical goals from a sequence, and later they begin to incorporate other intentional actions as they are better able to identify and represent actions in terms of social or conventional goals. Faithful imitation is much more than an automatic, all encompassing process.

#### **5.2.5. Imitation's Relationship to other Cognitive Skills**

Faithful imitation is inherently similar to reasoning, both counterfactual and mental state, for numerous reasons none less than the causal structure of a scene, and the intention of a model having been strongly and consistently shown to influence children's imitative behaviors. However, another critical component required for counterfactual, mental state and imitative reasoning is the ability to represent and coordinate alternative possibilities in the past, the present and the future. Conditional reasoning in temporal isolation alone is not as challenging as when required to add additional temporal elements. Conditional reasoning about different points in time, from different perceptions of an event, and using the resultant inferences generated to guide one's own behavior is a sophisticated and complex ability.

**Domain general executive functions.** "*All imitative acts are not of the same kind,*" said Meltzoff and Moore (1997). As such each imitative behavior will be the cumulative result of numerous different types or proportions of cognitive processing. Indeed some have suggested that there may be distinct mechanisms for imitation based upon the content to be imitated (Subiaul, Anderson, Brandt & Elkins, 2012).

As we have learnt children can be strikingly selective in what they imitate, inhibiting the replication of some action elements and incorporating other elements

from memory by drawing analogies between current and previously experienced behaviors. This selectivity and flexibility of imitation requires those lower-level domain-general cognitive process discussed in section 3.2.1. Executive functions such as inhibitory control, cognitive flexibility and working memory will be recruited for imitative behavior.

Children must be able to switch between their own and other's perspectives of the world in order to make sense of, and choose to imitate their intentional behaviors: this requires cognitive flexibility. Goals cannot be inferred without the application of some prior knowledge and causal structures of the same, or similar objects. This reasoning feat cannot be undertaken without working memory. If children were unable to inhibit the replication of all aspects of an action, regardless of causal relevance, intentions and goals, then there would be no possibility for selective imitation. Such inhibition of mimicry requires inhibitory control. If children were not engaging EFs then they would replicate failed and successful behaviors equally rather than imitating the inferred goals of an action sequence (Behne, Carpenter, Call & Tomasello, 2005; Meltzoff, 1995; 2005; Olineck & Poulin-Dubois, 2009). Even in the case of faithful imitation, where irrelevant actions are imitated, children will still not imitate every aspect of the behavior they see in an exhaustive manner. Exact movement forms or non-related incidental or accidental aspects of the movements are often omitted as they do not constitute a particular physical, conventional or social goal to be represented in the hierarchy.

Children build representations of a model action based around the perceived goal or goals of the action, and the model undertaking it. Additional information about the modeler's behavior and their characteristics are also incorporated into this cognitive model. Children require inhibitory control, cognitive flexibility and working

memory to successfully imitate based on these representations in a way that fulfills both their instrumental and social goals.

It could be then that faithful imitation comes about due to the executive functioning developments known to be taking place in the pre-school years (Anderson, 2002). For imitation, children's working memory and cognitive flexibility may aid them in committing to memory, generating a representation of, and re-enacting not only those causally necessary elements, but those other, less instrumentally salient features of the scene. They are able to represent more information about the world, beyond that causally necessary for physical goal directed success.

EF tasks themselves, such as those used in this thesis require the same forward and backward mental movement as MSU, CFR and Imitation. EF tasks rely on the application of some rule presented earlier in time to some stimulus presented in the present. The application of this rule constitutes a move forward into a future possible world. In EF the causal relationship that drives the inference is man-made, unlike mental state or true counterfactual reasoning. However, the rule application process is the same. Just like with imitation, children first reason based on physical, causal rules, then intentional rules and finally they can simultaneously reason about a broad range of intentional social features of an action. Still, however, they use the application of causal rules and information about other people and situations to guide their own future behaviors.

**Counterfactual reasoning.** In our imitation paradigm the most salient goal is a physical one. In the counterfactual reasoning paradigm the most salient goal is making a correct inference. We would expect children to behave in a goal directed way in both tasks. As we have seen imitation develops in the pre-school years.

Conditional reasoning, including counterfactual reasoning, also emerges in the pre-school years and continues to develop well into late childhood (Rafetseder & Perner, 2010; Robinson & Beck, 2000; Harris et al., 1996). As seen with MSU, between the ages of three and four children's responses to counterfactual reasoning questions move from being bound to their own perspectives of reality, towards supporting more causally driven inferences based on other possible, counterfactual worlds (Riggs et al., 1998; Guajardo & Turley-Ames, 2004). Children get better at navigation the actual and the possible in these different formats.

Counterfactual reasoning, like imitation requires looking back in time in order to look forward. For CFR this is realized through mentally looking back to a point where some change could have, but didn't take place and then reasoning forward, counterfactually, about the causal repercussions of this change. For imitation this is realized by the child looking to their representation of the actually undertaken elements of an action or behavior and reasoning about the causal relationships and intentional actions held within that representation before selecting which elements to implement in their own impending action or behavior. Around age four children become much better able to use their knowledge in an objective way to guide their inferences and behaviors.

In both MSU and CFR children begin to stop thinking instrumentally and uniformly about their environment and start appreciating the differing perspectives and beliefs of others, as well as the possibilities if a situation was fundamentally different, not just perceived as so. They become less constrained by the actual and more reflective about the possible (e.g., Riggs et al., 1998; Guajardo & Turley-Ames, 2004; Wellman, Cross & Watson, 2001; Wimmer & Perner, 1983). Children begin to more consistently represent the world as others experience it, or how it could have

been and could be. In doing so they are then able to use these representations to guide their own behavior. It follows then that when imitating, children can expand upon the instrumental goals of a model and begin imitating other possible goals. As they begin to experience the world as more than a physical reality, but a social one too, those things they may imitate have the opportunity to expand concurrently.

As we have seen in Chapter 2 children become better at reasoning about the causal structure of abstract properties in addition to concrete properties (Birch & Bloom, 2004). Causal structures exist across both domains, however, it is the type of content that differs: physical causality versus conventional or social causality. If there is some shared underlying ability here, however, supporting causal judgments across these two psychologically distinct domains we could expect to see a positive relationship between counterfactual conditional reasoning and faithful imitation. Correct inferences based on the causal structures laid out in the counterfactual reasoning tasks may correlate with inferences about the conventional structure of an object laid out in the imitation task.

**Mental state understanding.** Being able to imitate any behavior, simple or complex, requires basic perceptual abilities. More specifically, however, imitation requires attention to the actions and behaviors of other intentional agents in the world. Without attending to such agents their goals, intentions and social affordances cannot be represented and replicated. In early childhood, social attention (Olineck & Poulin-Dubois, 2009), point following (Carpenter et al., 1998) and, pointing, reaching, showing and checking back (Bretherton, McNew & Beeghly-Smith, 1987), are all suggested to facilitate imitative behavior. More broadly higher levels social interest, desirability, empathy and surgency also have been related to increased imitation (Chartrand & Bargh, 1999; DiYanni, Nini & Rheel, 2011; Hilbrink et al., 2013;

Uzgiris, 1981). Such evidence suggests that interest in and knowledge about others may have a role to play in subsequent imitation.

The ability to represent behavior, and do so in terms of mental states is vital for imitation (Sakkalou et al., 2012). Being able to interpret the intentions of a model for example, especially in a situation where there has been some form of goal failure requires understanding of space and causality, (Brugger et al., 2007). Children cannot imitate the goal of an incomplete action if they do not have the necessary causal and/or structural knowledge about the world. However, in addition to this causal knowledge they must also have an understanding about mental states that allows them to identify that this instrumental failure was not intentional, and infer that instrumental success was the intended outcome. As we have seen already in this chapter imitation, from an early age, relies on the identification of goals and intentions. Identifying these requires at least a rudimentary insight into the mental states of others. Completing the goal of a failed action for example requires an appreciation that failure was not the intended result of a person's pursuits. As the development of imitation unfolds during the pre-school years great strides are taken in children's mental state understanding. Mental state understanding, as measured through false belief type tasks for example, undergoes radical changes in the pre-school years (e.g., Baron-Cohen, Leslie & Frith, 1985; Wellman, Cross & Watson, 2001; Wimmer & Perner, 1983). Before the age of four the majority of explicit responses given by children to MSU questions do not take into account the perspectives of others. Very young children's appreciation of a scene seems inextricably grounded in the realism of their own perceptions. However, with age the mental states of others, and how they impact upon behavior, become more apparent in children's reasoning. Around age four children's MSU suddenly begins to expand and

judgments about others holding false or differing beliefs emerge. During this time the pattern of children's imitation also evolves with not only goal states being emulated but intentional actions being replicated faithfully. Children are still inferring the goals of a model, but instead of emulating only the causally necessary elements of a behavior, they begin to incorporate smaller irrelevant actions undertaken by the model: they faithfully imitate based on their understanding of the model's goals and intentions. It could be that children who are now able to represent the minds of others and the possibilities of a scenario choose to retain and imitate some elements of behavior they can identify as causally unnecessary for some other, social reason. They reason that a model who is perceived to have a high level of causal prowess may have undertaken a behavior that appeared unnecessary, so there must be some unknown conventional or social reason or goal for them having done so.

#### **5.2.6. Aims**

Our first aim was to assess the faithful imitation of our sample. We expected to find high levels of this often-reported phenomenon in our sample (e.g., Lyons et al., 2007; McGuigan et al., 2007; 2011; Nielsen, 2006; Nielsen & Tomaselli, 2010). We used two novel apparatus, each of which had two distinct moving parts. These moving parts interacted with a ball placed inside the apparatus. Every model the child saw demonstrated both these moving parts being manipulated by the experimenter. In the necessary condition both moving parts had to be manipulated in order for the modeled goal state to be achieved. In the crucial unnecessary condition only one moving part needed to be manipulated for the very same goal state to be achieved: it was only causally necessary to manipulate one of the two available moving parts in the unnecessary condition. Critically the experimenter modeled the same two-action sequence in both conditions, regardless of causal necessity. We wanted to see if, like



other researchers, our sample would readily incorporate both necessary and unnecessary intentional actions into their imitative behaviors. This would provide evidence that those children were willing and able to choose a possible world based on the goals and intentions of a model in the actual world: the model intentionally achieved a goal but also intentionally undertook unnecessary actions. To achieve the goal in the unnecessary condition, however, what happened in the actual world need not also happen in the possible world: children need not replicate all the actions that actually took place in the unnecessary condition to move into a possible world where the instrumental goal state is still met, some intentional actions may be ignored.

Our second aim for this chapter then was to assess the contrasting form of imitation. We expected to find that some children chose a more physically efficient sequence of behaviors to achieve the instrumental goal in the unnecessary condition (e.g., Brugger et al., 2007; Gergely et al., 2002; Gleissner et al., 2000; Hilbrink et al., 2013; Perra & Gattis, 2008). We expected that due to the impact of causal necessity, and the hierarchical structure of goal directed imitation that there would be significantly less faithful imitation in the unnecessary condition. We expected this because a more physically efficient selective imitation response would also result in the modeled goal state. If children were representing the possible world goal state based around the most salient goal identified in the actual world only then we would expect selective imitation.

With our first two aims we were interested in characterizing the imitative behaviors of our sample. We wanted to see how our sample would use what happened in the actual world to reach an unconstrained possible world of their own creation. Faithful imitation sees the actual world be replicated whilst selective imitation creates the same instrumental outcome via innovative means. To speak about our third aim

we must look more specifically at the imitation apparatus to be used. Although on the surface our two experimental apparatus were similar there were several crucial differences that meant we shall be presenting the data both separated by, and collapsed across apparatus (see Table 5.1.). These two apparatus will henceforth be referred to as the Tower and the Rake (see Appendix D).

In both the Tower and the Rake a ball contained within the apparatus could be manipulated and moved along a discrete trajectory. Also in both two actions were consistently modeled by the experimenter. In both the necessary and unnecessary conditions of the Tower there was one possible plane of motion with the ball only able to fall down the tube and trigger the springing sound. In the Rake there was also only one possible plane of motion in the necessary condition with the ball only able to move horizontally toward the end of the box. In the unnecessary condition for the Rake, however, there were two possible planes of motion for the ball: it could move in the same horizontal manner if both actions were imitated or the ball could drop down the small trap door if the first action was not replicated at the right point. This meant that there were more possibilities for motion competing with one another in the unnecessary Rake condition when compared with the same condition using the Tower. As such we expected there to be less faithful imitation using the Rake as children explored these additional affordances.

Furthermore in the necessary condition failure to complete the first modeled action had very different consequences on the potential journey of the ball held within. For the Rake apparatus the omission of the first action (pushing the trapdoor shut) meant that with the completion of the second action the ball would fall down the hole rather than come out of the end of the box as modeled. With the Tower, however, the failure to remove the first and/or second rod meant that the ball could not fall

down the tube fully and no noise would be generated: there was no alternative path for the ball to take. We expected that children would be more likely to faithfully imitate using the Tower during the unnecessary condition as failure to do so would result in a lack of movement and therefore be a boring outcome.

The second distinction between the two apparatus was that upon completion of the modeled goal the Tower apparatus made a comical noise while the Rake did not. This noise being associated with the goal for one and not the other could have altered the salience of the goal, making the Tower's more salient when compared to the Rake's. This salience would make the modeled outcome more attractive to the children and hence they would be more likely to replicate it.

A third difference between the two apparatus was that in the Tower both actions that could be carried out were identical: Pulling a rod horizontally out of the Perspex tower. For the Rake there were two different actions being modeled: the trap door being pushed closed and the rake being pulled horizontally. Having two identical actions may reduce cognitive load and be more easily replicable than having two distinct actions.

A fourth and final distinction between the two was that the set-up of the Tower allowed for both actions to be completed simultaneously: both rods could be pulled from the Tower at the same time. For the Rake apparatus, however, the two modeled actions, closing the trapdoor and pulling the rake, moved on different planes and hence could not be carried out simultaneously.

We expected that these differences would generate different patterns of response most evidently in the necessary condition where there was the most salient divergence in possible outcomes. The actual world and the possible worlds afforded by the Rake could show divergence from the actual instrumental goal. Conversely the

possible sequential worlds of the Tower afforded divergence from the intended sequence in the possible world.

In the critical unnecessary condition for both apparatus the modeled goal was the only possible outcome aside from failure to engage with the toy at all. Given these four crucial distinctions between the Tower and the Rake we were interested in examining the consequent differences in imitative performance they supported. Our aim was to provide evidence that goal salience and object complexity would alter response types and over all levels of faithful imitation (e.g., Brubacher et al., 2012; Chen & Waxman, 2103; Elsner & Pfiefer, 2012; Hermann et al., 2012). Manipulating the Tower led to a more salient, auditory goal, the actions to be carried out were simpler and more swiftly executionable and, there were limited possible movement outcomes. As such we expected there to be higher levels of faithful imitation carried out on the Tower when compared to the Rake, however, we expected the temporal order of this faithful imitation to be more variable.

Table 5.1. *Showing the characteristics of the Tower versus Rake apparatus*

| Feature of the Apparatus   | Tower | Rake |
|--|-------|------|
| Contains a ball which can be manipulated   | ✓     | ✓    |
| Has two identical actions that may be carried out  | ✓     |      |
| Has two different actions that may be carried out  |       | ✓    |
| Modeled actions can be carried out simultaneously  | ✓     |      |
| There is one single modeled movement and outcome to be achieved across conditions                              | ✓     |      |
| In one condition there are two possible movements and outcomes, one that has been modeled and one that has not |       | ✓    |

|  |   |   |
|--|---|---|
| The modeled outcome generated a sound  | ✓ |   |
| The temporal order of actions in the necessary condition is important for the goal     |   | ✓ |
| Two actions are necessary for the goal outcome in the necessary condition              | ✓ | ✓ |
| In the unnecessary condition only one action is causally necessary for goal completion | ✓ | ✓ |

---

Our fourth and final aim for this chapter then, after characterizing the faithful and selective imitation of our sample, was to relate their imitative behaviors to their reasoning skills reported in chapters two through four. We will do this by relating children's imitation performance on both their first trial, and across trials to their performance on the slide, story, executive functioning and mental state understanding tasks. Firstly a solid causal processing base, as investigated by our counterfactual reasoning tasks, could predict one of two things for imitation. In the necessary condition both modeled actions must be carried out in order for the goal state to be reached. As such we would anticipate that robust instrumental causal knowledge would be required for both conditional reasoning and imitation in this condition. In contrast, in the unnecessary condition only one of the modeled actions must be carried out in order for the modeled instrumental goal to be achieved. We might expect then that good conditional reasoning skills would relate to more single action responses in this condition as children reject the undertaking of superfluous actions. However, if children attribute the two action sequence modeled in the unnecessary condition to some unseen, instrumental causal relationship intentionally adhered to by the

experimenter, or identified some additional social or conventional causal structure, then we may in fact expect more faithful imitation to be related to conditional reasoning skill in this condition. The motivations behind how and why children move from the actual to the possible in these tasks are very different.

Secondly we wanted to look for any possible relationship between EF and imitation. As mentioned earlier imitation relies on domain general processing and as such children showing more proficiency in inhibition and cognitive flexibility may display a distinctive pattern of imitation. Much like the concern with conditional reasoning, however, this EF proficiency could be argued to effect imitation in contrasting ways. In both conditions there is an instrumental goal. In the unnecessary condition there is an additional goal, a social one determined by the intentional manipulation of an instrumentally unnecessary fixture. Good inhibitory control may limit the replication of superfluous actions not instrumentally necessary to the intended goal. Conversely good inhibitory control and cognitive flexibility may prevent the abandonment of superfluous action in order to retain features deemed socially necessary despite being instrumentally unnecessary. There being two goals in the unnecessary condition, one social and one instrumental, may add additional work for domain general processes and as such correlate positively with EF. Previous research suggests that we should expect to find a positive relationship between executive functioning and rates of imitation (e.g., Behne et al., 2005; Meltzoff, 1995; 2005; Olineck & Poulin-Dubois, 2009).

From a conceptual standpoint we would expect solid EF to correlate with faithful imitation as developmentally faithful imitation comes later so may in that sense be thought of as a more mature skill. Faithful imitation encompasses multiple goals in a way selective imitation does not. Furthermore how well children navigate

the most simplistic actual/possible tasks (like EF tasks) should impact upon one of the more challenging tasks, converting the intentional goal directed action of another in the actual world to your own decisions in bringing about the possible.

Finally, as has been discussed earlier in this chapter, the attribution of intentions and goals are a key-factor in imitation (Sakkalou et al., 2012). Therefore we expected to find that mental state understanding would relate positively to faithful imitation, most particularly in the unnecessary condition as children attribute intentionality to the model and therefore replicate their modeled two-action sequence. The intended instrumental goal of the model for the Tower is to generate the comical noise: the intended instrumental goal of the Rake is to extricate the ball from the puzzle box. Crucially, however, in addition to the instrumental goal, the unnecessary condition further encapsulates a social goal in that there is a concerted, repeated intentional act of manipulating an instrumentally inconsequential piece of the apparatus. As such those with good MSU may be expected to show higher rates of faithful imitation in the unnecessary condition particularly. Our final aim was to test this mental state understanding hypothesis and identify whether the performances in these two tasks that both require the actual world of someone else influencing the generation of a possible world are related.

### **5.3. Method**

#### **5.3.1. Participants**

Results presented are from the same sample as described in section 2.3.1.

#### **5.3.2. Materials**

Two separate apparatus were constructed for this task. The first was an adaptation of Apparatus G used in Tennie, Greve, Getscher and Call (2010) and

comprised of a transparent vertical tower with two solid rods slotted through the tower horizontally. The second was a rake like contraption based upon one used in Brugger, Lariviere, Mumme and Bushnell (2007).

The tube tower was made from a piece of transparent tubing sitting vertically on a square wooden base (15cm x 15cm x 4cm) and with a wooden lid (Appendix E). The tube was 40 cm in height and had a diameter of 9cm. Three, evenly spaced holes were drilled perpendicular to the tube through which solid, grey plastic rods 30cm long could be slotted to act as stilts. A tennis ball was placed in the tube resting atop one of the rods. When the rod(s) below the tennis ball were removed it could fall to the base where a pressure mechanism was activated and a small sound box affixed to the base would emit a '*boing*' sound akin to a cartoon rabbit or similar. In the necessary condition two rods were placed below the ball therefore both needed to be removed for the sound to be produced. In the unnecessary condition one rod was above and one below the ball, therefore only removal of the lower rod was causally necessary to generated the modeled goal sound.

The rake-box was constructed from wood and topped with transparent plastic so the simple mechanism inside could be seen (Appendix E). The box was 40cm long, 25cm wide and 10cm deep. A T-shaped rake divided the box in two and protruded out one end of the box where there was a small handle. The rake was used to pull a small wooden ball (approximately 3cm in diameter) from one end of the box to the other. Small plastic railings on either side of the rake shaft acted as runners, to provide a path for the ball. At the end of these two railings were small semi-circular openings in the wood, just large enough for the ball to pass through when the rake reached the end. On the right hand side of the box a small sliding portion could be pulled out transecting the runners and creating a hole in the base of the box (4cm x 4cm) down



which the ball could fall. Therefore if the ball was placed on the right hand side of the rake it could only be retrieved via the semi circular holes once the trapdoor had been closed fully. For the necessary action condition the ball was placed on the side of the rake with the trap door, therefore it was necessary for both the door to be closed and the rake to be pulled to retrieve the ball. The unnecessary condition then saw the ball on the left side of the rake with no trap door and hence no causal reason to close the trap door when retrieving the ball.

### 5.3.3. Procedure

Testing took place during the first visit children made to the University as described in section 2.3.1. In this within subject design children saw both the apparatus and both conditions during testing. Children were given one of these four possible apparatus-condition combinations, 1) Necessary Tower 1<sup>st</sup>: Unnecessary Rake 2<sup>nd</sup> 2) Necessary Rake 1<sup>st</sup>: Unnecessary Tower 2<sup>nd</sup> 3) Unnecessary Tower 1<sup>st</sup>: Necessary Rake 2<sup>nd</sup> 4) Unnecessary Rake 1<sup>st</sup>: Necessary Tower 2<sup>nd</sup>. Which one of these four apparatus-condition combination participants undertook was counterbalanced. The experimenter and child sat opposite one another, either side a square testing-table placed in the center of the room and covered in a black table-cloth. The first object was presented from out of the child's sight (retrieved from behind a curtain surrounding the testing area). The object was placed on the center of the table in clear view of the participant. The experimenter then spoke, '*Childs Name: Watch this!*' It was repeated if necessary to make sure the child is watching was fixating the apparatus. The experimenter then demonstrated the two actions in a slow and deliberate manner. For the rake apparatus the actions were the same for both conditions, first the trap door was slid closed and then the rake pulled. For the tower apparatus again the actions were the same for both conditions, pull one rod out, then

pull the next rod out. However, in the necessary condition both rods were inserted in slots below the ball, whilst in the unnecessary condition one was above and the other below. The apparatus was then reset in the same manner out of sight (beneath the testing table). The reset apparatus was then returned to the table and placed in exactly the same position as it had been during the modeling phase. The child then was offered a turn, '*Now it's your turn.*' The response was coded from the point the child first touched the object. This was repeated six times (experimenter model – participant offered a turn) while maintaining a natural social interaction through shifting eye contact with the participant and watching the apparatus. Once six turns had been completed the experimenter replaced the first object with the second (also hidden behind the curtain) and repeated the process in the opposite condition.

#### **5.3.4. Data Analysis**

Responses were coded after the testing visit using the audio-visual recordings of the session. Responses were coded using the Mangold Interact software for coding behavioral data. Responses were coded exhaustively across the response period that was calculated as 60 seconds from either the moment the toy was replaced on the testing table or the child's first contact with the apparatus. This first contact contingency was implemented as many children reached forward to grab the object before it had been placed fully on the table.

We coded any action carried out on the apparatus. An action was defined as the completion of one of the mechanism in questions affordances such as a trap door being closed or a rod being removed. Actions were not coded in instances where the participant completed then replaced or reversed the mechanism unless the instrumental goal was achieved. Action A referred to the first modeled action, Action B referred to the second modeled. In addition to coding the occurrence of actions their

temporal characteristics were also coded as reported in Brugger et al. (2007). In cases where action A and B were carried out simultaneously a third code to reflect this temporal relationship was employed. In all then for each apparatus there were six possible temporal outcome responses that any given participant could produce; completing neither action, completing action A then B, completing action B then A, completing actions A and B simultaneously, completing action A only or, completing action B only. 20% of the primary coders files were secondary coded and agreement was 100%.

We were interested in children's imitation of both the causally necessary and unnecessary actions they saw modeled. In the necessary condition action A and B were both causally necessary for goal completion. In the unnecessary condition only action B was causally necessary: Action A could be deemed redundant for goal completion in the unnecessary condition. Faithful imitation then in the unnecessary condition was defined as the copying of this first modeled action (Action A) regardless of instrumental necessity. This meant that of the six possible temporal outcomes three were considered faithful imitation; completing action A then B; completing action B then A and; completing actions A and B simultaneously. Within faithful imitation then were three distinct temporal patterns that could be further explored. Selective imitation in this condition referred to carrying out only the causally necessary action, just B.

Further, given the differences in our two experimental apparatus we wanted to inspect the patterns of performance on both separately, as well as having them combined. As such all trials were also coded for apparatus in addition to condition.

For each condition/apparatus combination children saw six models and were given six opportunities to interact with the apparatus themselves. We were interested

in both first trial responses for each condition as well as across trial patterns of performance. Our first trial responses allowed us to inspect children's initial response to the novel apparatus and the modeled actions they observed being carried out using it. This would allow us to identify children as faithful or selective imitators.

Finally, by aggregating performance across trials in each condition we generated an imitativity score to be correlated with the other cognitive indices collected throughout the course of the thesis. Each trial showing faithful imitation was given 1 point creating a possible score of 0-6 for each condition.

## 5.4. Results

### 5.4.1. Faithful Imitation and Sensitivity to Causality

Our first aim for this chapter was to replicate the finding of faithful imitation in this age group. Our second was to find evidence that causal necessity influenced imitative performance between conditions. In this section of the results the apparatus will be treated separately given their complimentary, but distinct structural features. Furthermore only data from the first trial completed will be presented.

**Tower apparatus.** In the first experimental trial faithful imitation was highly prevalent with children completing both Action A and B across condition in 95% of cases. Single action, Just B responses constitute selective imitation when carried out in the unnecessary condition. Therefore we compared faithful and selective responses to identify if one was significantly more prevalent than the other in the critical unnecessary condition. Faithful imitation was significantly more prevalent than selective imitation in the unnecessary condition  $\chi^2=25$ ,  $p<.001$  (see Figure 5.1.).

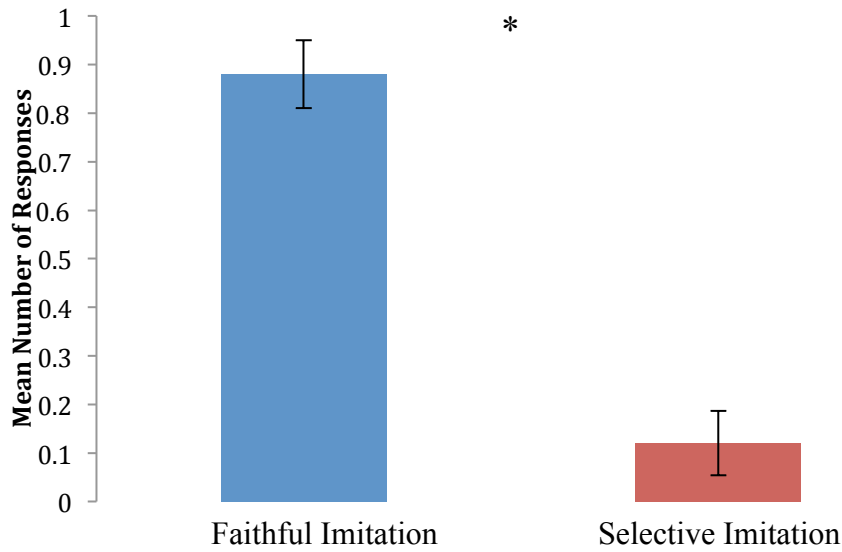


Figure 5.1. Mean number of faithful and selective imitative responses in the unnecessary condition for the Tower apparatus. Error bars represent standard error. \* Represents a significant difference between bars,  $p < .05$ .

When comparing faithful imitation levels in the necessary and unnecessary condition there was a significant difference  $\chi^2=4.05$ ,  $p < .05$ . There was significantly more faithful imitation in the necessary condition, 88% versus 100% (see Figure 5.2.)

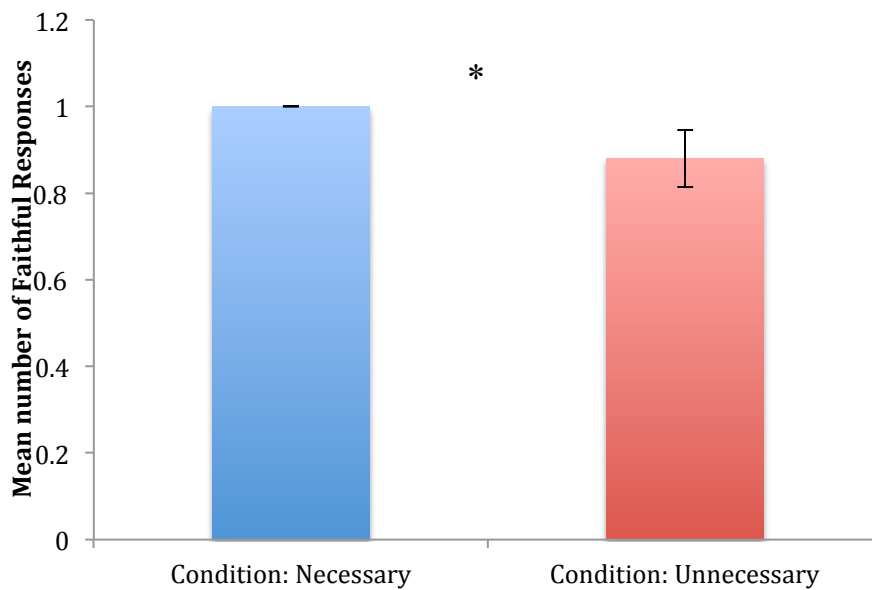


Figure 5.2. Mean number of two-action faithful responses between-condition for the Tower apparatus. Error bars represent standard error. \* Represents a significant difference between bars,  $p < .05$ .

Finally we investigated whether condition order affected children's faithful imitation rates (see Figure 5.3.). Faithful imitation was more common when the unnecessary condition was presented first but the difference between conditions did not reach significance  $\chi^2 = .47$ ,  $p = .49$ .

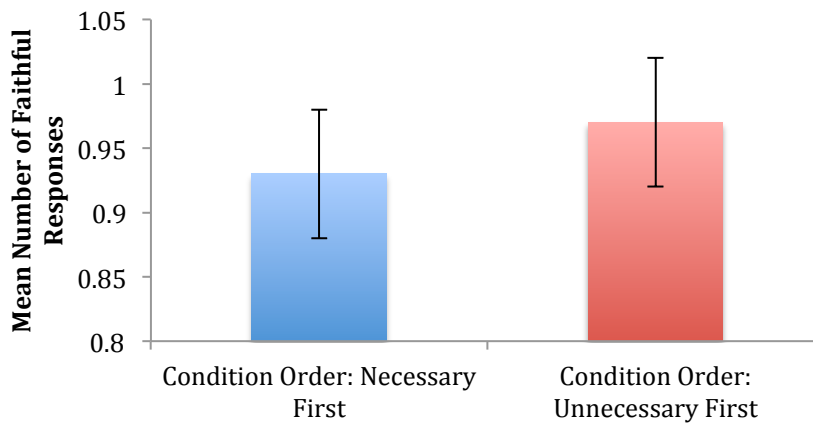


Figure 5.3. Mean number of two action responses, split by condition order, for the Tower Apparatus. Error bars represent standard error.

**Rake apparatus.** As with the Tower more often than not children completed both modeled actions on the Rake. In the first experimental trial children chose to complete both possible actions on the Rake across condition with a 72% completion rate. Again we wished to identify whether faithful imitation was more prevalent than selective imitation in our critical unnecessary condition. Indeed rates of faithful imitation were significantly higher than selective imitation responses in the unnecessary condition using the Rake apparatus,  $\chi^2 = 32$ ,  $p < .001$  (see Figure 5.4.).

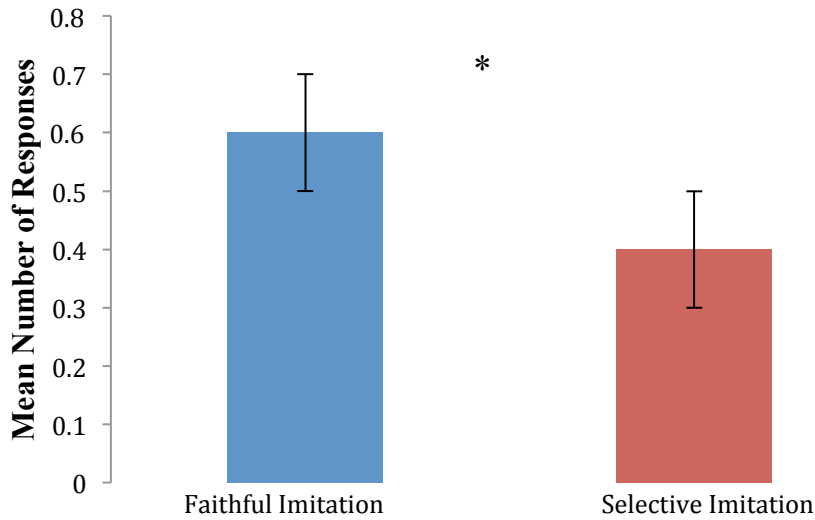


Figure 5.4. Total mean number of faithful and selective imitative responses in the unnecessary condition for the Rake apparatus. Error bars represent standard error. \* Represents a significant difference between bars,  $p < .05$ .

Unlike in the Tower, when comparing conditions there was no significant difference in the amount of faithful imitation  $\chi^2 = 3.65$ ,  $p = .55$  (see Figure 5.5).

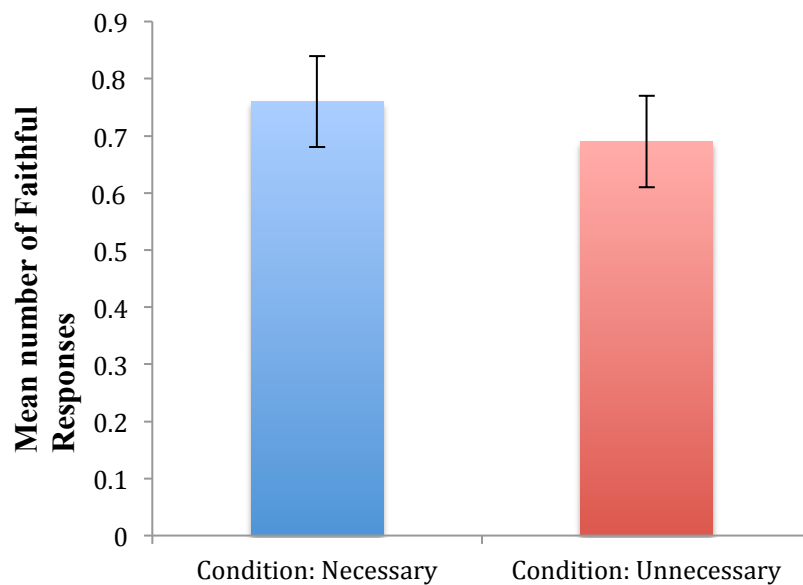


Figure 5.5. Mean number of faithful responses between condition for the Rake apparatus. Error bars represent standard error.

Again we investigated whether condition order affected children's faithful imitation rates (see Figure 5.6.). There was no significant difference in the number of two action responses dependent on condition order although, in contrast to the Tower, imitation was slightly higher after having seen the necessary condition first  $\chi^2=.12$ ,  $p=.72$ .

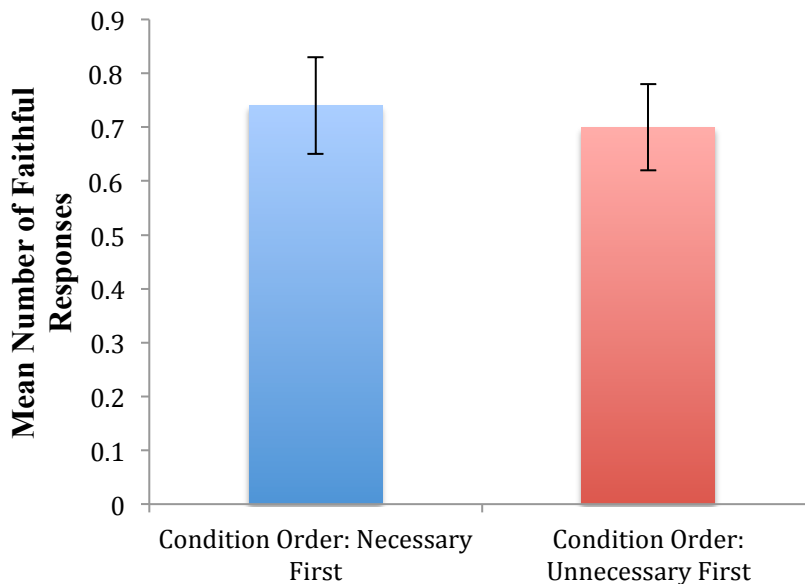


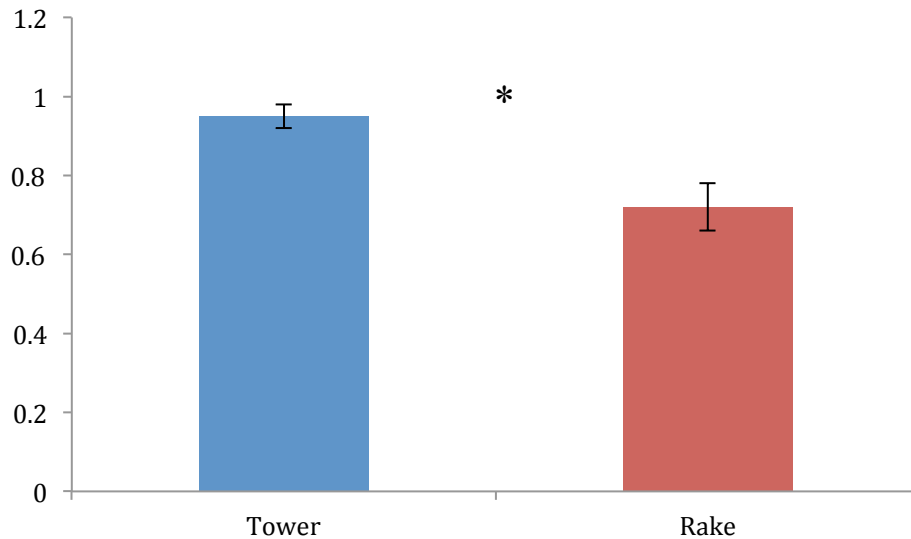
Figure 5.6. Total mean number of two action responses, split by condition order, for the Rake Apparatus. Error bars represent standard error.

#### 5.4.2. Comparing the Tower and Rake

Once we had established that faithful imitation was prevalent in our sample, and was sensitive to causality, we moved to the next aim of comparing imitation between our two apparatus. We expected that given the Tower having only one action

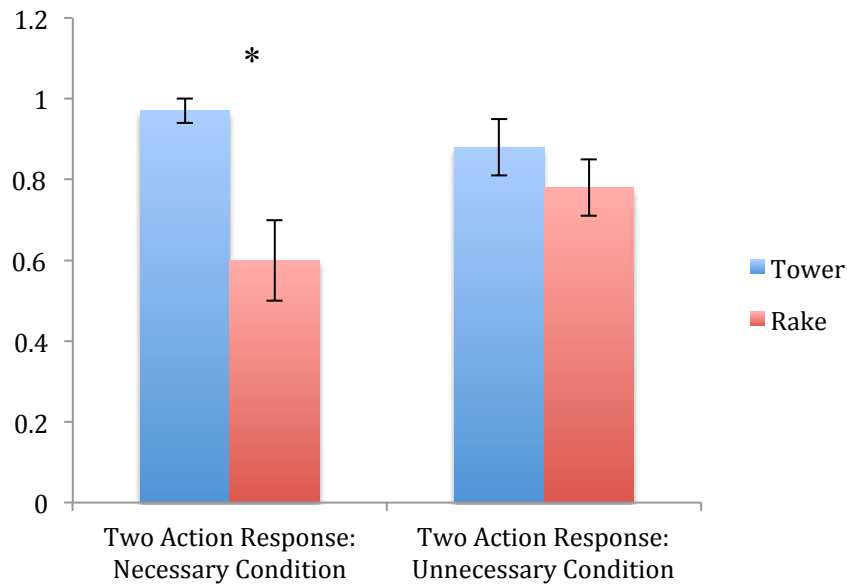


type, one physical outcome and a more attentionally salient goal we would find more faithful imitation than when compared to the Rake. Overall there was significantly more faithful imitation on the first trial following a Tower model than when following a Rake model,  $\chi^2=10.6$ ,  $p<.01$  (see Figure 5.7.).



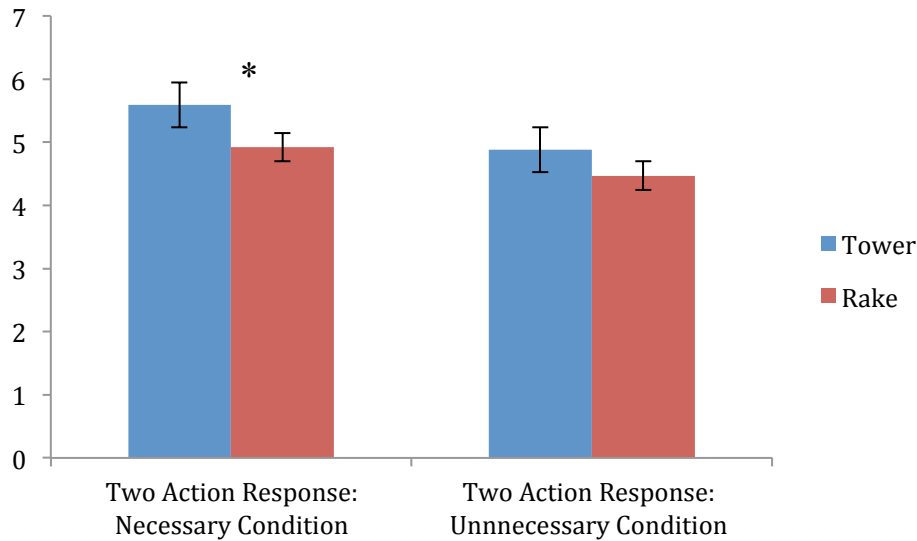
*Figure 5.7.* Contrasting the mean number of two action responses in the Tower and Rake apparatus on the first trial. Error bars represent standard error. \* Represents a significant difference between bars,  $p<.05$ .

In the unnecessary condition there was no significant difference between the number of two action responses carried out on the Tower and Rake,  $\chi^2=2.95$ ,  $p=.08$ . In the necessary condition, however, there were significantly more two action responses performed on the Tower,  $\chi^2=8.54$ ,  $p<.01$ . See Figure 5.8.



*Figure 5.8.* Contrasting the mean number of two action responses in the first trial for the Tower and Rake apparatus in the necessary and unnecessary conditions. Error bars represent standard error. \* Represents a significant difference between bars,  $p < .05$ .

Each child carried out 12 experimental trials, 6 of these in the unnecessary condition and 6 in the necessary. Different apparatus were used for each. When comparing the 6 trials in each condition, split by apparatus, there were significantly more two action responses using the Tower than the Rake in the necessary condition,  $F(1,55)=4.3$ ,  $p < .05$ , but not the unnecessary condition,  $F(1,54)=.92$ ,  $p = .34$  (see Figure 5.9).



*Figure 5.9.* Contrasting the total mean number of two action responses across trial for the Tower and Rake apparatus in the necessary and unnecessary conditions. Error bars represent standard error. \* Represents a significant difference between bars,  $p < .05$ .

**Temporal response types.** We decided to compare children's response types in more detail by examining the temporal order of their actions. To examine the full range of actions more closely we choose to include responses from all six trials.

In the necessary condition there were significantly more B then A,  $F(1,55)=12.42$ ,  $p < .001$  and simultaneous responses,  $F(1,55)=10.24$ ,  $p < .01$ , performed on the Tower than the Rake. The number of A then B responses were not significantly different,  $F(1,55)=2.5$ ,  $p = .12$ . There were significantly more Just B responses,  $F(1,55)=36.43$ ,  $p < .001$ , using the Tower than the Rake.

In the Unnecessary condition there were again significantly more B then A,  $F(1,55)=4.4$ ,  $p < .05$  and simultaneous responses,  $F(1,55)=6.23$ ,  $p < .02$  for the Tower than the Rake. However, A then B,  $F(1,55)=.51$ ,  $p = .48$ , and Just B,  $F(1,55)=.51$ ,  $p = .48$ , responses were comparable across apparatus (see Figure 5.10.)

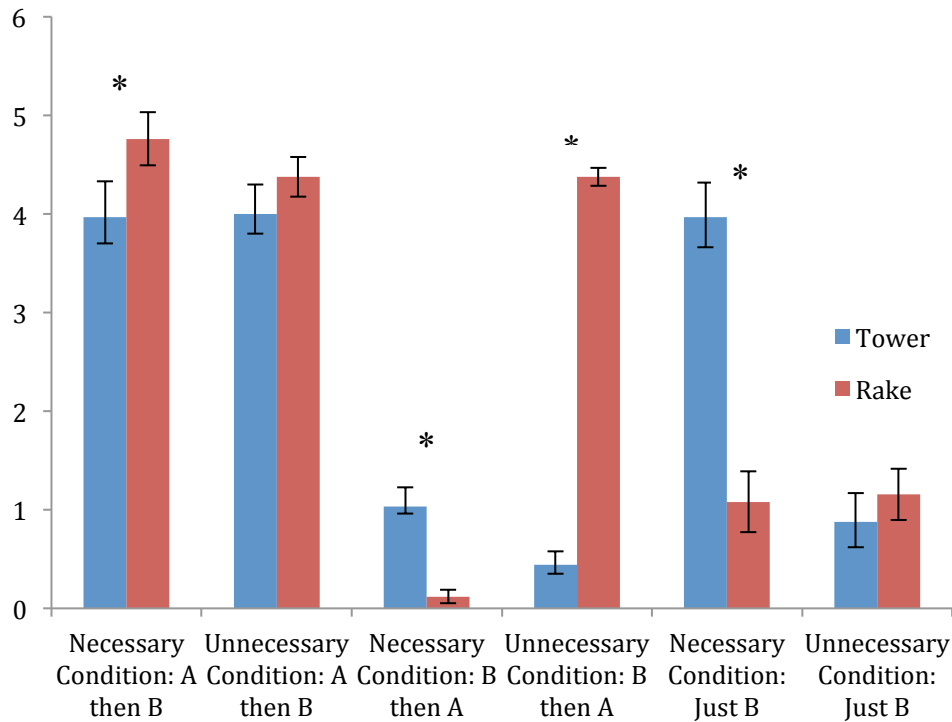


Figure 5.10. Contrasting the number of A then B, B then A and Just B, response types, across trial for the Tower and Rake apparatus in the necessary and unnecessary conditions (simultaneous responses not plotted as in both conditions using the Rake there were no instances of this response). Error bars represent standard error. \* Represents a significant difference between bars,  $p < .05$ .

The necessary condition generated the most variation in terms of one and two action responses, both between apparatus and within apparatus in the case of the rake. As such moving forward, when relating our sample's faithful imitativensess to the other measures of cognitive development that we have generated we used only data from the unnecessary condition where between apparatus faithful imitation differences are not significant. In the unnecessary condition there is no significant difference between the levels of faithful imitation for the Tower and Rake either on the first trial or across trials. Only the temporal order of action completion differs significantly, not action completion itself.

#### **5.4.5. Cross-task Relationships**

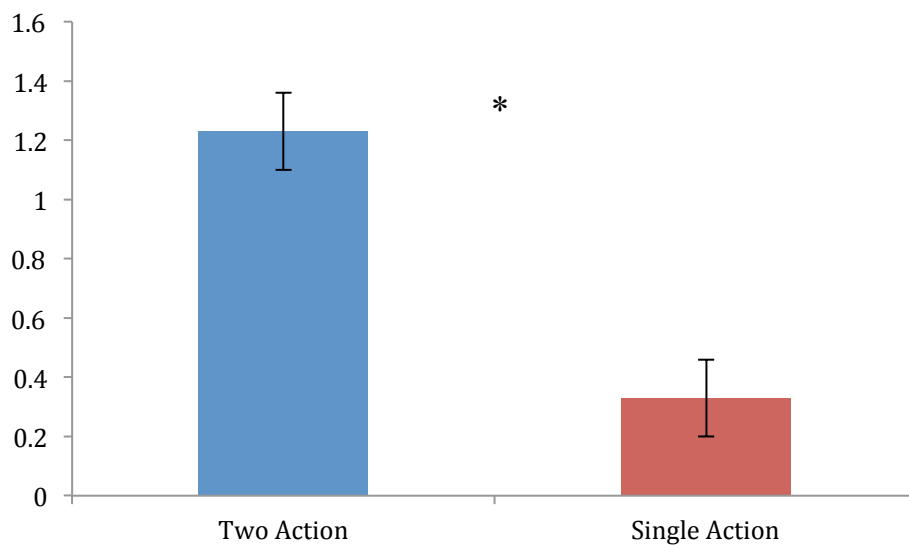
Our final aim for Chapter 5 was to relate the imitative behaviors of our sample to those reasoning skills reported in chapters two through four. We did this by relating children's imitative performance on both their first trial, and across trials, to their performance on the slide, story, executive functioning and mental state understanding tasks. Reasoning scores will be related to performance on each apparatus separately and combined. Firstly we will report the relationship between conditional reasoning and imitation. We had predicted that the reliance on causal knowledge for the completion of goal directed actions would result in a strong positive relationship between reasoning and imitation in the necessary condition. We had then suggested that causal conditional knowledge could predict diverse performance in the unnecessary condition. Selective imitation represents an instrumentally efficient route to goal completion therefore robust causal conditional knowledge may predict selective imitation. In contrast causal conditional knowledge may instead predict more faithful imitation as children are more sensitive to additional implied instrumental or conventional causal relations in the experimenter's intentionally modeled actions.

Secondly we had wished to test the prediction that imitative performance is affected by the domain-general processes of inhibitory control and cognitive flexibility. We will compare the imitative responses of participants high and low on these EFs.

Finally we predicted that those children better at reasoning about the mental states of others' would be more faithful imitators in the unnecessary condition where instrumental causal structure alone is not predictive of the experimenter's goal

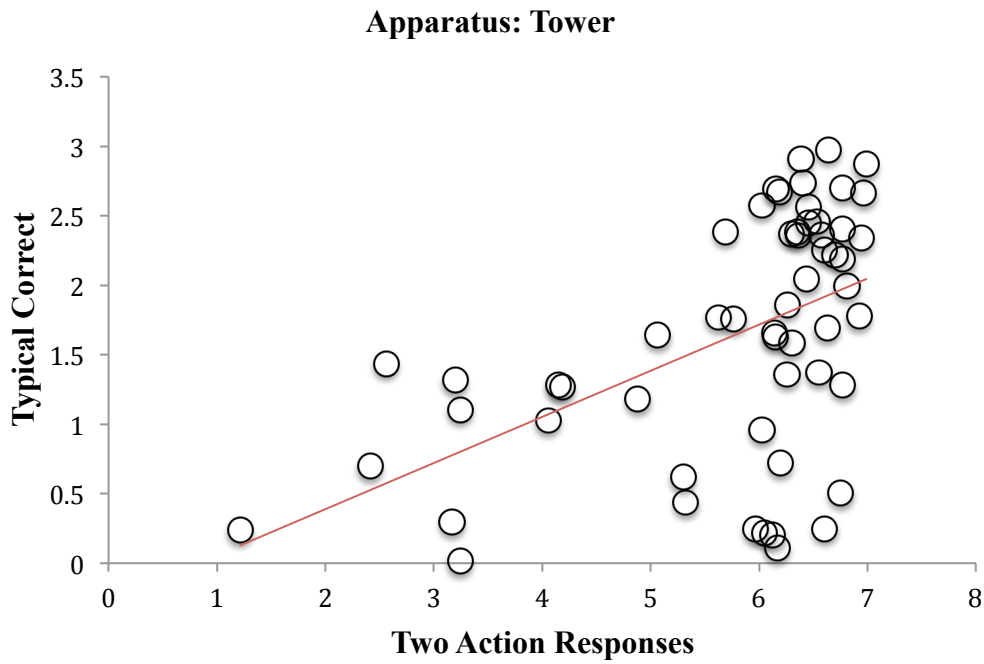
directed actions. We tested whether participant's MSU battery score, or single item responses were predictive of imitative behavior.

**Conditional reasoning. Tower.** For these cross task relationships we used both first trial data and the combined scores collected across all six experimental trials. Children who completed a two action response in their first trial had significantly higher typical story scores than their peers who completed only one action,  $F(1,53)=4.15$ ,  $p<.05$  (See Figure 5.11.). Across trial correlations between typical story and standard slide conditions, and two action responses across trials were marginally significant;  $r=.23$ ,  $p=.088$  (see Figure 5.12) and,  $r=.25$ ,  $p=.071$  (see Figure 5.13). The number of Just B responses correlated negatively with the typical story measure but not the standard slide,  $r=-.47$ ,  $p<.001$  (see Figure 5.14.).



*Figure 5.11.* Mean scores on the Typical Stories task for those completing single and two action responses in their first Tower trial. Error bars represent standard error. \*

Represents a significant difference between bars,  $p<.05$ .



*Figure 5.12.* Scatterplot (including jitter) showing the correlation between two action responses across trials using the Tower and typical story performance.

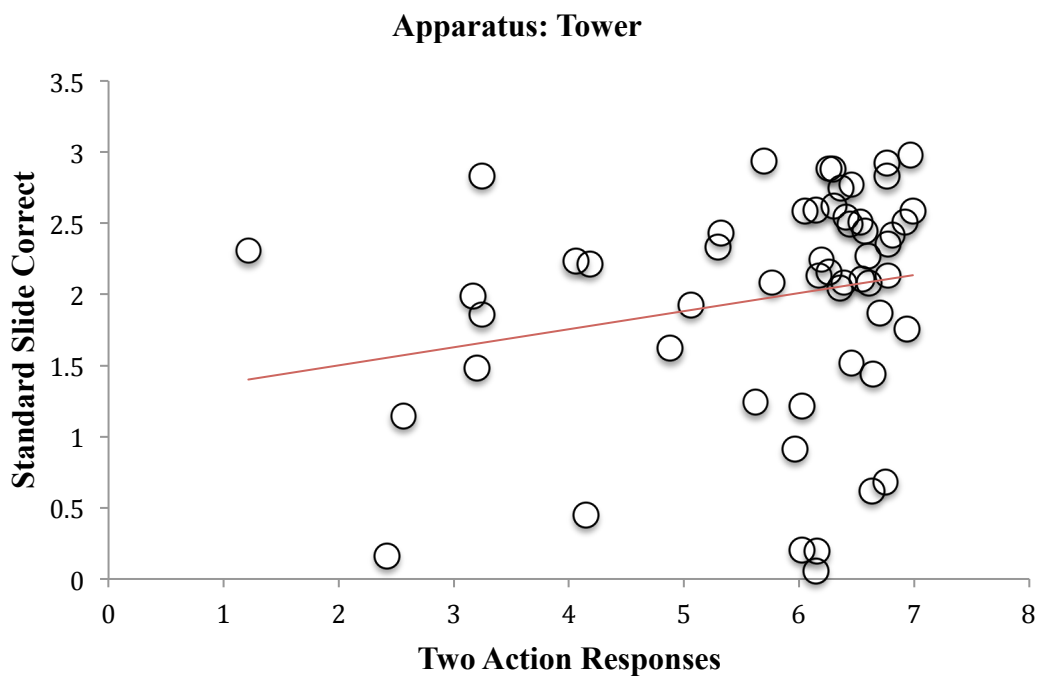


Figure 5.13. Scatterplot (including jitter) showing the correlation between two action responses across trials using the Tower and standard slide performance.

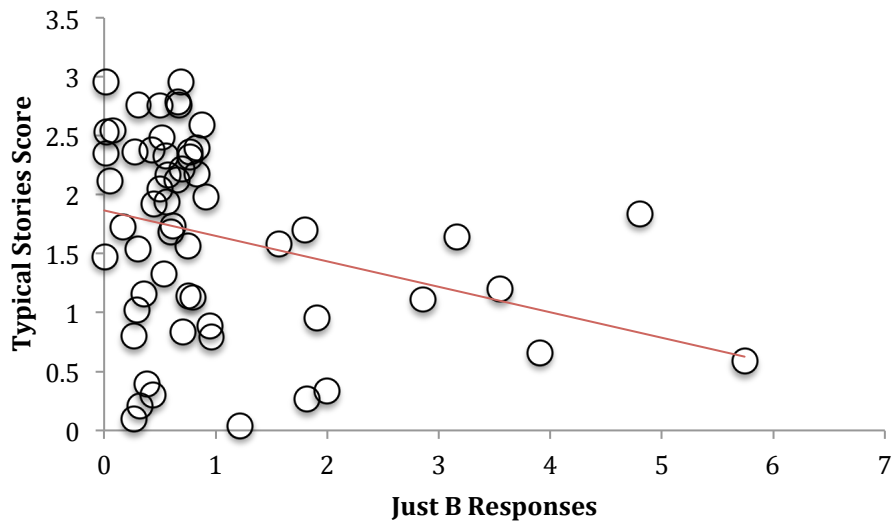
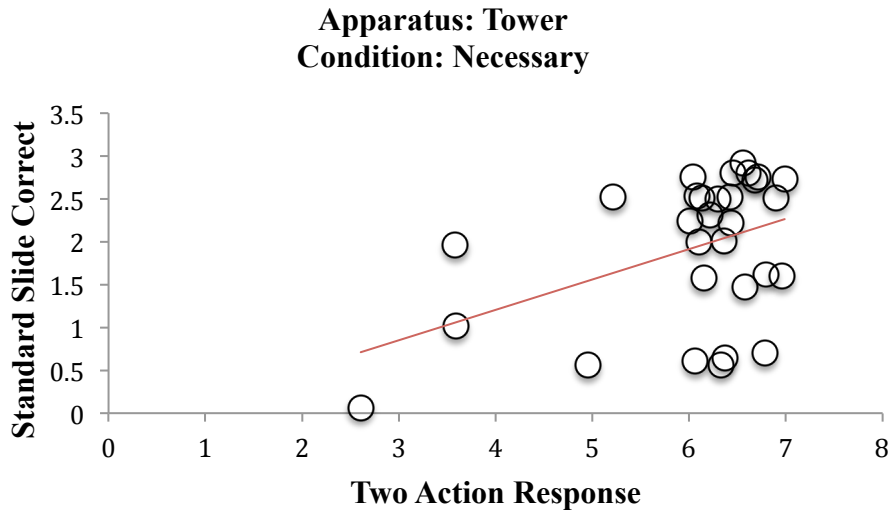


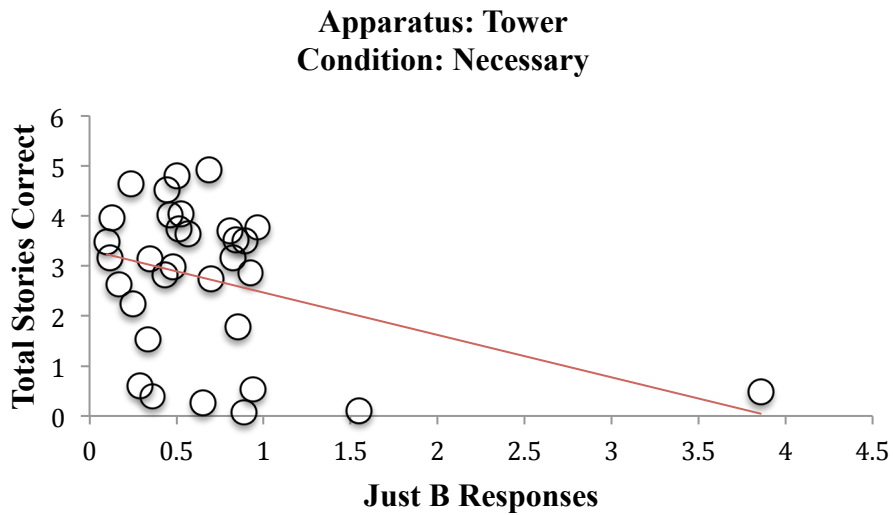
Figure 5.14. Scatterplot (including jitter) showing the correlation between single action, Just B responses across trials using the Tower and Typical Story performance.

When separated by condition the same pattern was evidenced in the necessary condition. Two action responses correlated marginally with standard slide,  $r=.36$ ,  $p=.053$  (see Figure 5.15). Again in the necessary condition there were negative correlations between the number of Just B responses total stories,  $r=-.41$ ,  $p=.025$  (see Figure, 5.16.). However it was clear that this particular relationship was driven by the presence of an outlier. In the unnecessary condition only Just B responses and typical stories retained their negative correlation,  $r=-.46$ ,  $p=.021$  (see Figure 5.17).





*Figure 5.15.* Scatterplot (including jitter) showing the correlation between two action responses across trial using the Tower in the necessary condition and standard slide performance.



*Figure 5.16.* Scatterplot (including jitter) showing the correlation between single action responses across trial using the Tower in the necessary condition and total story performance.

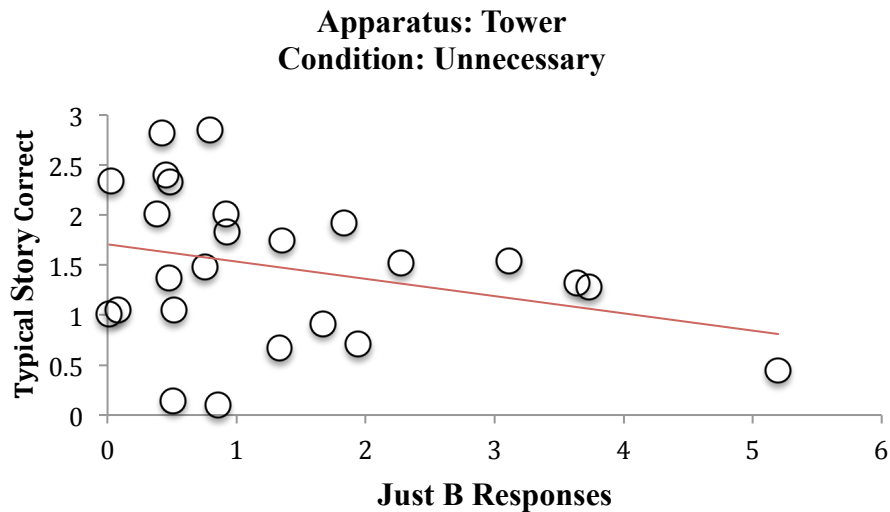


Figure 5.17. Scatterplot showing the correlation between single action responses across trial using the Tower in the unnecessary condition and typical story performance.

*Rake.* Children who completed a two-action response had comparable counterfactual reasoning scores to their single action peers. There were no linear associations seen between two action or Just B responses and counterfactual reasoning. This lack of pattern extended to both the necessary and unnecessary conditions.

*Tower & Rake.* In addition to looking at trends within apparatus we were also interested in imitiveness across apparatus. Only unnecessary trial data was used to this end. There was no linear relationship evident between unnecessary two action responses and either of our conditional reasoning measures. However, there were some cases where the temporal sequence of response was related to conditional

reasoning. B then A responses in the unnecessary condition, that is copying the modeled actions in the reverse order, showed a significant negative correlation with typical story ( $r=-.31$ ,  $p=.02$ ) and atypical story ( $r=-.27$ ,  $p=.043$ ) performance. Indeed B then A responses in this condition were also negatively correlated with total story performance,  $r=-.34$ ,  $p=.01$  (Figure 5.18).

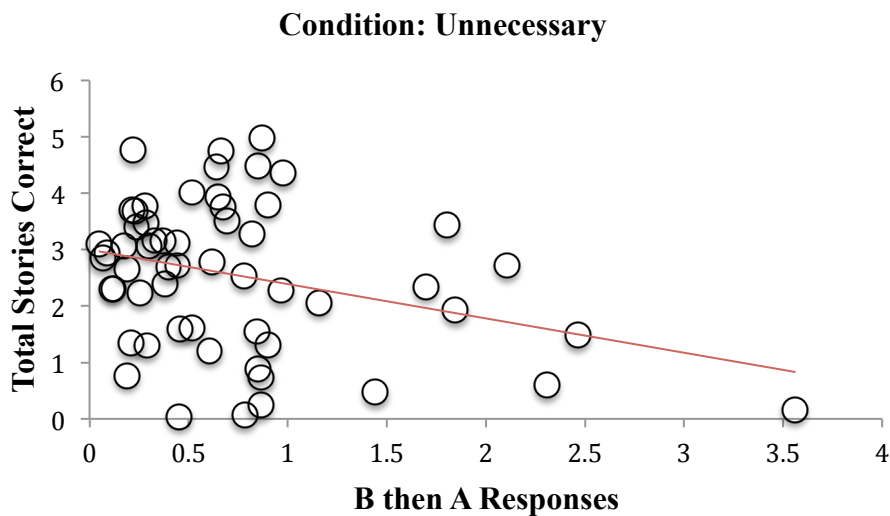


Figure 5.18. Scatterplot (including jitter) showing total story performance and B then A responses for both the Tower and the Rake in the unnecessary condition.

**Executive functions. Tower.** Given the distribution of the EF tasks as discussed in section 3.4 we inspected the between group differences in imitative performance based on a median split of the participants. There were no significant differences in performance on the first trial as a function of DCCS or Flanker performance in the necessary condition as all participants generated a two action response (see Figure 5.19). In the unnecessary condition also there were no significant differences between high and low EF groups.

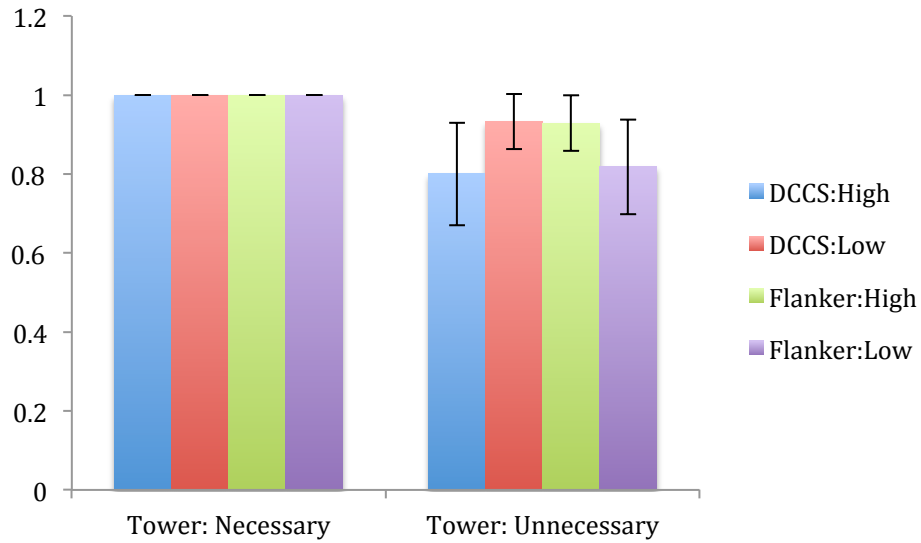


Figure 5.19. Mean number of two action responses performed on the Tower first trials separated by condition and EF task performance. Error bars represent standard error.

*Rake.* In the necessary condition children in the high Flanker group produced significantly more two action responses than their peers in the low flanker group  $\chi^2=4.96, p<.03$  (see Figure 5.20). The same pattern is visually discernable for the DCCS task, however, this difference did not approach significance  $\chi^2=1.79 p=.18$ .

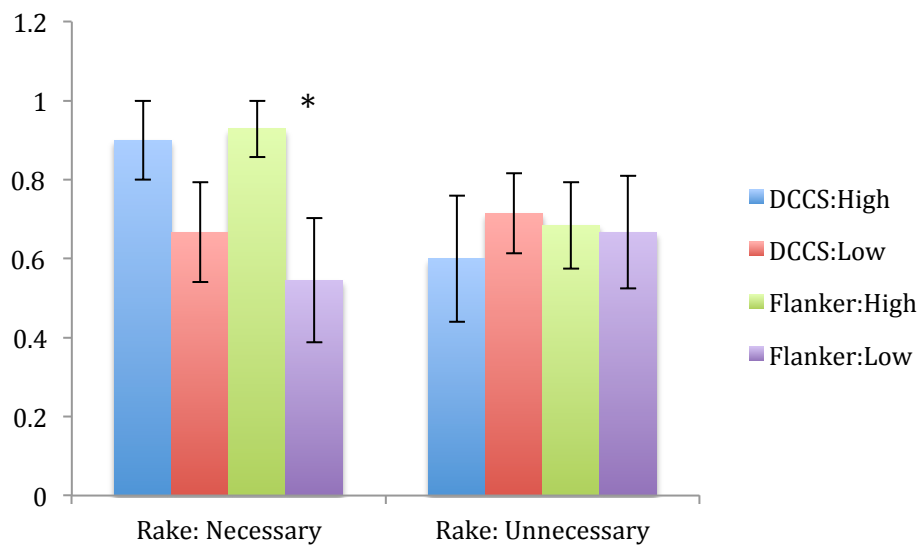


Figure 5.20. Mean number of two action responses performed on the Rake first trials separated by condition and EF task performance. Error bars represent standard error.

\* Represents a significant difference between bars,  $p < .05$ .

*Tower & Rake.* Although in both EF measures the two action imitation scores in the unnecessary condition for the low groups were lower than their high group peers this difference was small and did not reach significance. DCCS:  $F(1,54) = .18$ ,  $p = .74$ . Flanker:  $F(1,54) = .24$ ,  $p = .63$  (see Figure 5.21). Furthermore none of the possible temporal response patterns were related to either Flanker or DCCS performance.

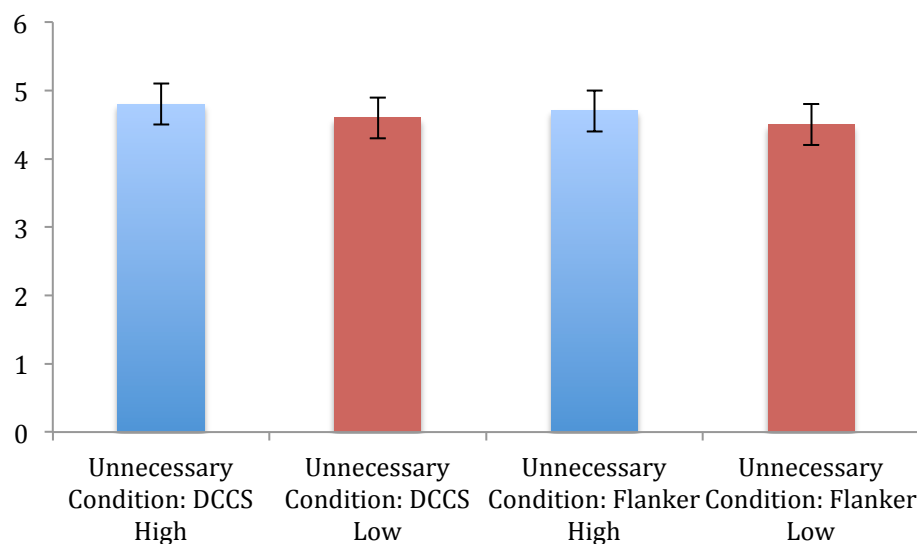


Figure 5.21. Column chart showing the mean number of two action responses in both Flanker and DCCS high and low median split groups. Error bars represent standard error.

**Mental state understanding.** *Tower & Rake.* There were no statistically significant relationships found when looking at each apparatus' first trial responses individually so only the combined data will be presented here. No linear relationship

was evident between imitative responses in the unnecessary condition and mental state understanding battery scores (see Figure 5.22). However, there was one interesting positive relationship between two-action imitation and Question 5, the real apparent emotion item in which none of the children answered the memory check question correctly. Scores for the target questions only on this item correlated positively with the number of two action responses in the unnecessary condition  $r=.29$ ,  $p=.03$ .

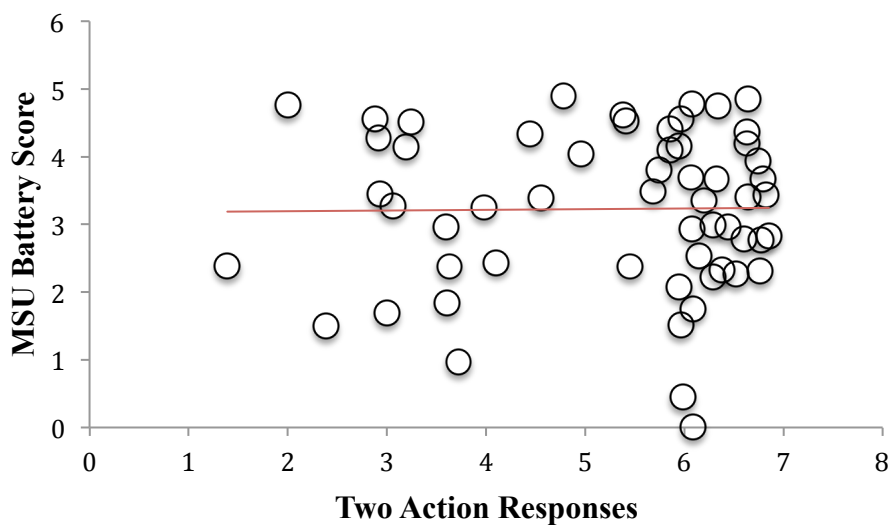


Figure 5.22. Scatterplot (including jitter) showing MSU battery performance and two action responses during the object imitation unnecessary condition.

Table 5.2. Non-parametric correlation table between mental state understanding and two-action responses in the unnecessary condition.

|                                    | Unnecessary Condition |
|------------------------------------|-----------------------|
| Mental state understanding battery | -.08, $p=.57$         |
| Question 5                         | .29, $p=.03^*$        |

## 5.5. Discussion

Imitation is a central tool in children's developmental toolbox. It is both a corner stone of socio-cognitive development and a powerful learning mechanism (Over & Gattis, 2009; Zmyj et al., 2012). Children imitate a variety of modeled actions and behaviors in daily life (e.g., Bertenthal & Boyer, 2009; Carpenter et al., 1998; McGuigan et al., 2007; Wolschlagler & Gattis, 2000). Young children can be sophisticated and discerning imitators (e.g., Brugger et al., 2007; Hilbrink et al., 2013; McGuigan et al., 2013; Nielsen, 2006). Children become increasingly faithful imitators in the pre-school years (e.g., Lyons et al., 2007; Matheson et al., 2013; McGuigan et al., 2007; 2011; Nielsen, 2006). Intentional actions, even those that are physically redundant, are often replicated (e.g., Bekkering et al., 2000; Carpenter et al., 1998; Carpenter et al., 2002; Over & Gattis, 2009). Intentional goals are identified and replicated in addition to those more concretely observable instrumental goals. Our first aim for Chapter 5 was to replicate this established finding of the pervasive faithful imitation of novel actions in this age group (e.g., Lyons et al., 2007; Matheson et al., 2013; McGuigan et al., 2007; 2011; Nielsen, 2006). On the first trial, across condition and apparatus, 83% of responses included both modeled actions. In the critical unnecessary condition this faithful two-action response was the overwhelmingly most popular, with it making up 85% of the total action sequences executed. The critical condition for the purpose of identifying faithful imitation in both apparatus was the unnecessary condition. For both Tower and Rake the unnecessary action was completed by the majority of participants in our sample. In addition, these two action responses, containing one instrumentally unnecessary action, were significantly more common than single action, Just B responses: faithful imitation far outstripped selective imitation regardless of apparatus.

These findings then support the large body of work reporting that children of this age tend to faithfully copy the actions of models even when contradictory causal information is readily available to them (e.g., Lyons et al., 2007; McGuigan et al., 2007; 2011; Nielsen, 2006; Nielsen & Tomaselli, 2010). This strong preference was apparent across both of our experimental apparatus: children are very proficient at incorporating novel actions from their recent past into their future behaviors.

Our second aim for this chapter was to seek out residual evidence supporting the hierarchical goal directed theory of imitation described by previous researchers (Bekkering et al., 2000). Causal necessity has been shown to significantly impact faithful imitation rates in young children with spatial goals being replicated before conventional ones (Bekkering et al., 2000; Brugger et al., 2007; Gergely et al., 2002; Gleissner et al., 2000; Hilbrink et al., 2013; Perra & Gattis, 2008; Williamson & Markman, 2006.) When looking at our apparatus individually there were significantly more Just B responses in the critical unnecessary condition when compared to the necessary condition using the Tower. This pattern remained for the Rake but did not reach statistical significance. Imitation then, although predominantly faithful, was significantly impacted by the causal necessity of the actions involved in one of our apparatus. At the group level children were sensitive to causality to some degree. We can confidently say therefore that faithful imitation is prevalent at age four, and unnecessary actions occur significantly more frequently in our imitative paradigm than required by causal necessity. Moreover, despite this propensity for faithful imitation there is evidence that causal information significantly influences responding.

The patterns of imitation in our sample strongly support the centrality of goals and intentions to imitative behavior. Imitation as we are investigating it is the



replication of goal directed actions (Bekkering et al., 2000). Imitation's goal directed nature has been elucidated through many paradigms (Carpenter et al., 2005; Gergely et al., 2002; Meltzoff, 1995). Children in our sample overwhelmingly chose to complete the modeled instrumental goal in both apparatus. Even in the Rake where there were other possible innovative goals that could be sought children preferentially chose the modeled outcome.

Although our two apparatus generated complimentary data confirming the existence of faithful imitation in our sample there were distinct differences between the Tower and Rake that help us understand further what features impact upon imitative readiness. The Tower apparatus data, as expected, generated significantly more faithful imitation. This pattern was significant only in the necessary condition but the trend also appeared in the unnecessary. We hypothesized three reasons why this pattern would occur. Firstly the two Tower actions were always identical where the Rake actions were of two distinct forms. As such the Rake task could be considered more taxing as two distinct actions must be observed, represented and then selected from replication. This would take more processing capacity than having two identical actions represented in the same way. Additionally the repeated nature of the actions on the Tower would increase their salience to the observer, and hence their likelihood of replication.

Secondly the goal state of the Tower was to allow the ball to fall to the base of the apparatus, triggering a comical sound via a pressure sensor. This addition of sound to the goal was not present in the Rake. Previous studies have presented evidence that vocal cueing and novel labels increase goal salience, and as such increase imitation (e.g., Carpenter et al., 1998; Chen & Waxman, 2013; Elsner & Pfiefer, 2012). This addition of sound to the goal could strengthen its representation and as such its

likelihood of being manually replicated by the subject. In the necessary condition using the Rake apparatus the absence of such a salience cue may have resulted in weaker representations of the instrumental goal leading to children producing an increased number of innovative responses. Children are sensitive to goal salience when interacting with a novel object.

Finally there could be another contributor to this decrease in two action responses in the Rake when compared to the Tower. The Tower's structure meant that the ball could only do one of two things dependent on the participants' actions in the necessary condition: remain static or fall down the tube. For the Rake, however, the ball could do one of three things: remain static, fall down through the trapdoor or come out the end of the apparatus as modeled. There being multiple options or possibilities increases the amount of goal competition present for the children. Where upon manipulating the apparatus the children seeing the Tower can only make the ball move in one direction, the children manipulating the Rake can either have the ball follow the modeled trajectory or have it fall down the trap door: children seeing the Rake in the necessary condition have an opportunity to be innovative and exploratory in a way simply not available with the Tower. Children using this apparatus had the opportunity to reason that an interesting alternative could have occurred had the trapdoor not been closed. As such it is unsurprising then that responses from the Rake in this condition are significantly different from those on the Tower, with faithful imitation being more infrequent in the former, despite remaining at a reasonably high level. Children at age four are sensitive to multiple possibilities when interacting with an apparatus that affords them.

Authors such as McGuigan et al. (2007) suggest that the imitative proclivity seen in young children is a mostly automatic behavior derived from a "highly

adaptive conformist bias”. Lyons et al. (2007), however, claimed that such imitation arises from a poor grasp of physical causality requiring some other strategy to support children’s action. Nielsen (2006) instead identified a strong social motivation for over-imitation in children. Nielsen and colleagues suggest that children become less interested in instrumental goals and more in intentional, social goals. Evidence such as faithful imitation only perseverating when a model remains in close proximity and offers the child the modeled apparatus strengthens this claim (Nielsen & Blank, 2011), as too does evidence from even younger participants showing that a move from selective to faithful imitation styles is predicted by extraversion scores (Hilbrink et al., 2013). Following on from such studies others have investigated the role of early social competence on imitation. A number of researchers have presented evidence for a specific link between nascent socio-cognitive skills such as social attention and cue recognition, and imitation (e.g., Brubacher et al., 2013; Chen & Waxman, 2013; Olineck & Poulin-Dubois, 2009).

Our evidence so far points to children being sensitive to the causal affordances of the apparatus as imitation is reduced in the critical unnecessary condition. This contrasts with Lyons et al. (2007) claiming that imitation is the result of poor causal understanding. Additionally we also agree with there being a strong social motivation to imitate as the majority of our sample imitated above and beyond that which was required by physical necessity alone. In a further attempt to speak to these differing accounts of imitation we then turned to look at how other forms of reasoning were related to children’s imitation. We used indices of both physical and social reasoning to inspect alongside our imitative performance scores.

### **5.5.1. Object Imitation, Inhibitory Control and Cognitive Flexibility**

As we have seen imitation can be both selective and faithful, and influenced by numerous experimental and social factors. As such domain general cognitive abilities such as inhibitory control and cognitive flexibility will be required during performance. Selectivity in imitation would be impossible if children were unable to inhibit the replication of all observed actions (e.g., Behne et al., 2005; Meltzoff, 2005; Olineck & Poulin-Dubois, 2009). Flexibility in imitation is necessary to navigate multiple goals, and types of goals simultaneously. In our sample, although temporally faithful responses were more frequent in the high performance groups of both our EF measures, and Just B responses more frequent in the low performance groups for both measures these differences did not reach statistical significance across the data set as a whole.

When looking only at the Rake apparatus necessary condition, however, there was a significant positive relationship between Flanker performance and faithful imitation. Those children in the high Flanker group carried out significantly more two-action responses during the first trial of the necessary condition than their low group peers. Recall that this apparatus condition combination allowed children to select one of three possible outcomes; the ball remains still, the ball falls down the trapdoor or the ball comes out the end of the runners as modeled in the actual world. Children who opted to achieve the modeled outcome inhibited the desire to explore the object's possible, non-modeled affordances and instead chose to take the road more travelled. This evidence could be used to refute McGuigan et al.'s (2012) claim that imitation is a largely automatic behavior. Children with better inhibitory control were MORE likely to imitate faithfully. If this behavior were highly automatized then

we would expect the opposite pattern with little or no influence of processing limitations.

Another possible reason for this relationship being significant in the Rake and not the Tower is that the Rake model incorporated two different actions on the object in contrast to the same action being repeated during Tower models. Perhaps then this inclusion of two distinct actions was more taxing to domain general cognitive functions. A third and final possible reason is that the physical goal in the Rake condition was less salient due to the absence of an auditory cue. Goal salience has reportedly been increased by such auditory means in the past (e.g., Carpenter et al., 1998; Carpenter et al., 2002; Chen & Waxman, 2013). As such identifying, retaining and replicating the modeled goal may have been more cognitively challenging for those seeing the Rake when compared to the Tower in the necessary condition.

Executive functions here seem necessary but not sufficient for object imitation. There were generally non-significant results except in a small subsection of our sample. This was not entirely surprising in light of the type of actions we had chosen to use and the complexity of the objects. Only two actions were demonstrated on each object and when combined with the causal transparency of the objects this meant a very simple action sequence. They did not need to inhibit or switch between representations in the same way as demanded by CFR and MSU. The model and the participant shared the exact same set of circumstances, the causal structure remained the same (unlike in CFR) and the viewpoint for each was the same as the objects were symmetrical and entirely visible to both. Only the person acting on the object changed across trials. Although novel objects were used in a novel way the actions themselves were familiar, either a pull or a push. There were no hidden mechanisms or ambiguous parts. Had the children been given the opportunity to manipulate the

objects prior to modeling it may have been expected that the majority would have rapidly solved the mechanisms to achieve the goal state: the object itself was not enigmatic. As such we would not have expected to a high EF demand when all that children needed to do in order to solve the apparatus was to copy some or all of the actions they were being shown. Had a more complex task been used with an opaque or counterintuitive causal structure then perhaps EF would have begun to show a predictive effect.

### **5.5.2. Object Imitation and Counterfactual Conditional Reasoning**

We had hypothesized that due to the similar conceptual demands of imitation and counterfactual reasoning that there would be a positive relationship between these two skills. Both imitation and counterfactual reasoning require a person to use information about a recently passed reality to inform their decisions or behaviors in the near future: they require looking back in time to the actual in order to look forward to the possible. Counterfactual reasoning involves thinking about possible alternate outcomes while imitation involves actually choosing a modeled possibility over some innovative alternative. Furthermore both counterfactual reasoning and faithful imitation require the identification of, inferences about, and adherence to some recently presented causal structure.

When investigated in our sample, however, there was no linear relationship found across apparatus between faithful imitation responses and counterfactual reasoning performance. However, this was somewhat unsurprising given the prevalence of faithful responses across participants and conditions: there was much greater variation in response type during counterfactual reasoning than during imitation trials.

When looking at each apparatus separately, however, it was clear that responses from the Tower apparatus were related to CFR in a way that responses on the Rake were not in the necessary condition. The amount of selective imitation, particularly in the necessary condition correlated negatively with CFR story and slide performance. Children who exhibited a causally inappropriate response in the necessary condition were poorer at reasoning causally about the location of a protagonist or the path of motion of a ball. This pattern was not seen in the necessary condition of the Rake, perhaps because here the failure to carry out the first action still led to an interesting and innovative spatial outcome, i.e., the ball disappearing through a trap door.

Looking back to the cross apparatus data then, when the specific temporal order response types were investigated there were some interesting results. B then A responses in the unnecessary condition correlated negatively with counterfactual story performance. Remember in the story task: those who performed poorly on the story task were those who failed to reason exactly in line with the causal structure as set out in the story. In the imitation task then: those who chose the B then A response also more often failed to reason in line with the implied temporal causal structure of the imitation apparatus. Although they carried out both modeled actions they did not do so in the same modeled temporal sequence. There was some contiguity based on reasoning about and acting upon some newly presented causal structure.

There could be multiple reasons as to why children reasoned in this way across the two tasks; (i) Children may have been unable to remember the specifics of a given causal structure, verbally presented in the story or visually implied with the apparatus, i.e., there may have been a bottleneck in working memory; (ii) Children may have chosen to disregard, or been unable to retain specifically the temporal

information presented to them; (iii) Children may have struggled to integrate the information being presented and their subsequent actions, i.e., there may have been a bottleneck in cognitive flexibility although this seems unlikely given the absence of a CF relationship reported earlier; (iv) Children may have had a more general problem with recognizing and utilizing causal structures; (v) Children may have understood the presented causal structures but been unable to use this knowledge to guide their goal directed actions. The unnecessary condition may have been particularly ambiguous as the modeled actions were not both necessary, and as such the order of the actions were not as critical. The subsequent representation generated then may not have put enough weight on the temporal sequence, only the occurrence of the actions. The contrast between causal necessity and intentional action may have caused those children with a more tenuous ability to reason about causal structure to be less confident with their own subsequent actions.

It is key to remember that the counterfactual story task relied almost completely on a linguistic presentation of a set of causal events embellished by props: the object imitation task was a purely observational task with no details about causal structure being verbalized. Despite these very different presentations of a causal sequence participants' response patterns across the tasks were still correlated to some degree. Four-year-olds who perform poorly on counterfactual conditional reasoning tasks that require the integration of their own causal knowledge and newly presented causal knowledge are less likely to copy the exact action sequence of a model during an simple object imitation task using a novel, but causally transparent apparatus. Such evidence points to some participants having a prevailing difficulty with executing behaviors and/or selecting responses based on their representations of causal structure at this age. This is regardless of the medium through which causal structure was



generated or the method and type of responding required. From a functional perspective imitation and counterfactual reasoning are similar as their deployment reduces the requirement for personal, explicit trial and error experience of the world for learning. They both offer a learning proxy that reduces the risk of acting out the entirely unknown.

### **5.5.3. Object Imitation and Mental State Understanding**

Understanding behaviors in terms of the mental states supporting them is crucial for imitation (Sakkalou et al., 2012). Given evidence in the literature that early socio-cognitive skills are related to early imitation we had hypothesized that the older children in our sample who had better MSU would show increased levels of faithful imitation (e.g., Bretherton et al., 1987; Brubacher et al., 2013; Chen & Waxman, 2013; Olineck & Poulin-Dubois, 2009). Furthermore general levels of social interest and other social indices have been related positively to imitation in children (e.g., Chartrand & Bargh, 1999; DiYanni et al., 2011, Hilbrink et al., 2013; Uzgiris, 1981). Faithful imitation, however, was the most dominant response form by a significant margin in our sample as a whole. It was not entirely surprising then, as seen initially with our CFR results, no significant linear relationship was evident when looking at the MSU battery scores and the different imitative responses, although some small positive trends were visually evident. However, in this case even looking at each apparatus separately did not reveal any of the expected underlying relations.

Only the last of our five individual items in the MSU battery showed a statistically significant positive relationship with our target imitation condition. This was despite all of the children answering the memory check portion of this question incorrectly. Item five, the real-apparent emotion question, required children to answer two test questions: how the protagonist felt inside and how they looked on their face.

This was the only item on the battery that challenged children to not only represent the mind of another but also represent an additional layer of conflict by contrasting how they really felt, and how they wished to appear. All the earlier MSU questions comprised of only one dimension of a protagonist's mental state, what they desired, what they believed, what they knew and what they expected. Here the psychological contrast came between the participant and the protagonist. Item five, however, asked children not only how the protagonist felt (their mental state) but also how the protagonist acted (which contrasted with their mental state). The contrast occurred within the same person, not between two different individuals. Children heard information about a past event, teasing, and a present event and then had to identify how the protagonist would be feeling and acting: the integration of multiple pieces of causally inferred information here is crucial. The processes involved in faithful imitation are then more akin to this item 5. Participants watch a model then must choose how to act themselves moving forward. Children watch a model behave in a way that contrasts with the most salient goal, the instrumental goal. As such children must attribute some additional layer of intention to the model: there must be some additional social goal to be achieved. Here then is the same discord or tension as the model or protagonist does not externally behave in a manner entirely congruent with their initially perceived mental state. Item 5 challenges children to reason about the agent's same psychological reality in two contrasting ways in much the same way as imitation does.

That the battery score did not relate to faithful imitation was initially surprising, however, it could be that at this age the majority of children are able to represent others' minds to a sufficient extent ensuring no real challenge is imposed during these relatively simple, two action sequences. Conversely it could be that the

MSU questions used with the exception of Item 5 were not sufficiently complex to highlight any significant individual differences.

Other earlier emerging socio-cognitive indices pertaining to mental state understanding have been related to imitation rates therefore our lack of global battery predictive effects by no means precludes the importance of MSU in imitation in this age group. Rather the types of imitation and MSU tasks themselves perhaps could be modified in order to find more variability in the sample and perhaps allow any predictive effect of mental state understanding at this age to be observed.

Several options to this end present themselves. Regarding the imitation paradigm it would be possible to make the critical first action, deemed causally unnecessary in the critical condition, more distinct. Children are much less likely to copy an action that is “off object”, that is it does not physically manipulate the target object during imitation tasks (Brugger et al., 2007). Other options for increasing imitation task difficulty include increasing the number of modeled actions, making the causal structure more ambiguous or adding additional goals. Additionally MSU questions that probe the multiple possibilities inherent within a single protagonist could be created, rather than only the contrast between self and other. By increasing the complexity of the MSU task, to include past, present and future information more analogous results may be obtained.

#### **5.5.4. Limitations**

There were two main limitations associated with the imitation tasks chosen for this chapter. Firstly, as with the counterfactual stories task it would have been better for us to choose two tasks that were identical in structure so as to simplify the results for our later analyses with other variables. Being able to treat the Rake and the Tower data as a single body of evidence in both conditions would have been beneficial for us

rather than having between apparatus differences. Another consideration was that perhaps these apparatus were too simple to find meaningful variation between children. Having a more complicated apparatus structure, through having a more opaque structure or through an increased number of actions may have strengthened our results and the conclusions we can draw from them.

### **5.5.5. Conclusions**

Faithful imitation is highly prevalent at fifty-two months of age: four-year-olds are very good at incorporating actions performed by others in the recent past into their own repertoire. Children are good at incorporating the actual into the possible. Regardless of the causal necessity of the actions involved faithful imitation is always the preferred option at the group level in this age group. In addition to this, however, there is evidence to suggest that the causal affordances of actions are not ignored completely. Action sequences that diverge from faithful imitation are more frequent in the condition where all modeled actions are not in fact causally necessary in order for some physical goal state to be reached. Children imitate in line with a hierarchical organization of goals with conventional goals being slightly less replicable than their physical counterparts.

The salience of a goal as altered through action repetitiveness, auditory cues and the number of possible outcomes significantly effects faithful imitation. The more salient a goal, the more likely it is to be imitated. The more salient an action or outcome is in the actual world influences the likelihood of its transmission into a possible one.

Faithful imitation is pervasive and robust, and does not appear to be globally reliant on executive functioning skill in this age group. EF relations are restricted to one distinct condition and apparatus combination suggesting that domain general

processing skills are necessary but not sufficient for imitation as measured here. There are some significant relationships between counterfactual conditional reasoning and imitative success. Those who reason more successfully also act in a more imitative manner across the domains. In both CFR and imitation moving forward requires looking back, looking to the actual as you form the possible.

Finally MSU does appear to be related to faithful imitation, albeit in a more subtle way than predicted. MSU, like CFR an imitation requires the selection of one among a set of possibilities. In all of these still developing socio-cognitive skills children must look back at the actions, thoughts, or causal landscapes of others in the actual world in order to generate and their own representations of the world from which to derive their own possible paths.

## **Chapter 6. Imitation of Language and Linguistic Competence: Grammaticality, Novel Lexical Items and Complex Syntax**

### **6.1. Abstract**

In the previous chapters we have discussed children's reasoning skills and patterns of imitation, highlighting how these skills function to move children from actual to possible realities. Children's actions on objects can have highly constrained causal consequences as seen in their imitation of actions in the previous chapter. Language, however, is an arguably even more potent force for change in a child's life. Comprehending and producing language catapults a child from experiencing reality based on observation alone to being able to represent situations based on the words of others. Not only can children better understand the world through comprehending language, they are more able to change it. In the context of this thesis being interested in children's navigation of actual versus possible words, language is an invaluable resource for our investigations based on two unique features: language sits astride the concepts of imitation and innovation and language is *the* tool for entertaining differences between the actual and the possible, and it is no surprise that the majority of tasks examined in this thesis so far are unthinkable without language.

In Experiments 1 and 2 we shall focus on this first feature of language, its paradoxical nature as being both highly imitative and highly innovative. We shall look at children's willingness to imitate different features of language (sentences, novel verbs and syntactic structure) and some influences on this willingness (grammaticality, novelty and familiarity). Experiment 3 shall further investigate whether imitation of these linguistic features is positively correlated with the action imitation behaviors described in the previous chapter: is imitation as a strategy deployed similarly across both action and language tasks?

In the Chapter 7 then we will turn to look at language's role as a tool for the manipulation of the world, and whether competence in this predicts any of the cognitive and behavioral indices collected in Chapters 2 through 5.

## **6.2. General Introduction**

*“As far as acquisition of language is concerned, it seems clear that the reinforcement, casual observation, and natural inquisitiveness (coupled with a strong tendency to imitate) are important factors, as is the remarkable capacity of the child to generalize, hypothesize, and ‘process information’ in a variety of complex ways. These may be largely innate or may develop through some sort of learning or through maturation of the nervous system” Noam Chomsky (2002, pp 137).*

When children learn to speak they are learning to use the lexical tokens of their native tongue. They do this within the grammatical and syntactic environment of this native language. However, to become truly competent language users children must go above simply regurgitating the language they have already heard, they must be generative and productive in new and flexible ways. Tomasello (2000) called this interplay between imitation and innovation, the actual and the possible, a “paradox of language”. Children must be conventional in order to successfully convey information using the correct lexical items, arranged in a coherent syntactic structure. However, children must also be innovative in rejecting previously heard utterance constructions in favor of the one that is most appropriate for the task in hand. Within language production, then, is a distinct tension between the actual and the possible. The linguistic input we hear constitutes linguistic reality, how language has actually been used. Conversely, generative utterances we produce, or could produce on a daily basis

embody a possible range beyond that what has been heard before. One of the hallmarks of human language is this limitless number of possible coherent re-combinations that, although novel, can be easily comprehended by other language users.

How we navigate these actual and possible linguistic worlds has indeed proven a fruitful ground for research into, and theorization about, language acquisition. Children's productivity with language has been held as evidence against language acquisition being the result of exclusively imitative processes. Linguistic output that features over-regularization errors i.e., "she hitted me" or "there were three sheeps," have similarly been used to generate hypotheses about the mechanisms by which children acquire productive language (Tomasello, 1992). Some researchers have attributed such errors to children possessing innate, or rapidly formed syntactic categories and linguistic rules that they can flexibly apply to newly learnt lexical items. Chomsky (e.g. 1968) put forward the thesis that children's rapid acquisition of language, in the face of "the poverty of the stimulus" must be resultant of them having some innate "universal grammar" embodying key abstract principles about language. Chomsky's work began a popular movement of research based around the principle of language acquisition being grounded in underlying linguistic representations (Tomasello, 2009, p.3).

More recently, however, such accounts have attracted criticism from scholars characterizing young language learners as more conservative and imitative than previously suggested. Tomasello's (1998) usage based account, for example, highlights the imitative nature of young children's language. Here 'imitative' is taken to mean exactly what it means in the context of the imitation work of the previous chapter, namely a 'copying' that is not mere mimicry, but rather sensitive to the



functions, goals, and intended meaning of the utterances that children are copying. These more recent theories benefit from advances in other branches of developmental psychology that highlight the breadth and variety of learning mechanisms that young children have available to them when learning language. Socio-cognitive skills such as mental state understanding, joint attention and imitation, and domain general categorization skills, all support language above beyond the simple inductive or associative processes favoured in the mid-20<sup>th</sup> century (Tomasello, 2009, p3-4). Tomasello argues that alongside learning specific lexical items children are accumulating syntactic knowledge through specific, concrete items that then, only gradually, facilitate the acquisition of more abstract rules. Tomasello (2009) proposed that familiar early word combinations could contain “slots” into which different words could be inserted. These so called “pivot schemas” allow children to be generative in their communications whilst retaining the same structural qualities they are familiar with. Tomasello et al. (1997) showed that children can incorporate entirely novel lexical items into such slots e.g. “Wug gone” or “more Wug.”

In the context of verb learning Tomasello (e.g. 1992; 2000) proposed that item-based constructions are generated before children can become flexible in their use of verbs. The “verb island hypothesis” puts forward an account where children under the age of four exclusively use verbs in the construction that they have been exposed to, and, in that sense, are limited to ‘imitation’ (Tomasello, 1992). Evidence from a corpus of speech collected from one infant showed that almost all of her utterances were organized around single verbs (1992). Further evidence from a corpus for 12 pre-school children (Lieven, Pine & Baldwin, 1997) emphasized the restricted nature of children’s productions, with the majority of verbs only being used in one construction. According to this account, then, young children’s language is wholly

fixed upon verbs in set constructions around which various nouns/pronouns can be inserted. Children may be able to generalize a particular verb in a particular construction, but not yet have mapped that verb as belonging to a particular class of verbs. This results in young children being resistant to using verbs in a flexible way. With increased exposure to language children can gradually build upon these verb islands through a process of structure mapping and by drawing analogies between single items over time, hence beginning to generate verb general schemas. Marchman and Bates (1994) suggest that a critical mass is reached after which children are able to utilize their lexical knowledge of specific verbs to form semantic subclasses such as transitive versus intransitive verbs. Once these subclasses are formed children can use them to support not only comprehension of new words but also to produce novel utterances incorporating the rules governing the subclass. Children move from context level representations of verbs to broader sentence level constructions that can be viewed as invariant across numerous verb items of a certain class (Brooks & Tomasello, 1999b).

As children become more competent users of language they become better able to understand the world, past, present, and future. In addition, they become better able to reason about, and make changes to the world, past, present and future due to this newly acquired ability to learn about things temporally and physically removed from them. Furthermore children become able to represent new possibilities based upon their linguistic abilities due to being privy to worlds only accessible through language. Language supplies children the systems of thought necessary to reason, interpret, organize action and engage in other mental acts they could not have achieved before (Chomsky, 2007).

### 6.2.1. Research Questions and Aims

We then have set up two reasons as to why language is an important setting in which to explore children's developing navigation of actual versus possible worlds.

(i) Language is a tool that can be manipulated and used to change their future world and, (ii) language is a paradoxical skill that relies on the integration of imitation and innovation. Specifically in this chapter we will use data from two experimental tasks that allow us to explore this second feature in greater detail.

The first experimental task we shall report simply investigated whether children would imitate or alter different types of repetition in a simple sentence constructed with familiar words. This task was adapted from Over and Gattis (2010). In earlier work relating to the imitation of repetitions by young children Dan Slobin and colleague (e.g., Slobin, 1968; Slobin & Welsh, 1973) suggested that language imitation was reliant on memory and grammatical knowledge rather than intention understanding. Their conclusion was based upon children being more likely to replicate the semantic content of a sentence rather than its' specific surface features. This in many cases led to repetitions being omitted in favor of adding stress to the singular form (Slobin & Welsh, 1973). This pattern of imitation mirrors the selective imitation of actions discussed in the previous chapter: children replicate the modeled communicative goal but not the exact means used to achieve it. Furthermore children were more likely to correct ungrammatical utterances with age, again, arguably a selective imitation strategy. The authors argued that this was an adaptive strategy as otherwise children would struggle to navigate the accidental "stutterings and false starts" inevitably observed in adult speech from time to time (Slobin & Welsh, 1973). In contrast to Slobin, who ruled-out intention understanding as a critical component of verbal imitation, Over and Gattis (2010) reported that the perceived intentionality of a

model did indeed impact upon imitation patterns in young children, aged three-years. Over and Gattis concluded that children were interpreting ungrammatical repetitions as a failed attempt of sorts much like those documented in behavioral re-enactment paradigms (Meltzoff, 1995). Children then repair the ungrammatical breakdown during the course of their own utterance. Language imitation appears to rely not only on memory processes and linguistic competence, but also intention understanding. Over and Gattis' work was based upon numerous observations in the imitation literature focusing on the role of intentions on action imitation discussed at length in Chapter 5 (e.g., Bekkering et al., 2000; Sakkalou et al., 2013).

The same task as used by Over and Gattis was chosen for this thesis to chime with the kind of object and tool use imitation paradigms discussed in the previous chapter, where intentional actions could be both necessary and unnecessary, but for the most part are faithfully imitated regardless. As language, like any other behavior, can be represented as intentional and goal directed, we wished to document how our sample chose to imitate a modeled sentence in addition to a modeled action. Some of these sentences contained repetitions. The repetitions were either grammatical or ungrammatical. Over and Gattis (2010) showed that three year-old children who completed this task often removed ungrammatical repetitions but retained grammatical ones. In addition children removed ungrammatical repetitions more often when the model was an animate agent, when compared with an inanimate one: this shows a marked social component and influence of intentionality on imitation. We wished to compare older children's imitative behavior in this linguistic context with their imitative behavior in the action imitation contexts described in Chapter 5. This task would help draw comparisons between how children chose to manipulate the

world with their actions in the action imitation tasks, and how they chose to manipulate the world with their utterances in the verbal imitation task.

For the second task we chose to implement a verb-learning paradigm designed to assess comprehension and production of a form of syntax that remains relatively rare in the spontaneous speech of four year-old children. The passive verb construction that we chose to use will be discussed later in more detail. We wished to use an experimental design, as with the reasoning and object imitation tasks, to most closely mirror the setting in which the imitation data was collected. Through using a novel verb we could determine whether children would again faithfully imitate, through replication of this verb or selectively imitate by reverting to a semantically appropriate but familiar alternative. The same could be said for the novel verb form, would children faithfully imitate and retain the same verb construction or would they, where possible selectively revert to a more familiar form. We aimed to generate data that could help elucidate the question of how imitative and how innovative children at age four are with their language, and how their imitation of language relates to their navigation of these actual and possible worlds in other contexts.

In addition we hoped to speak to the question of how language was used as a tool. We specifically were interested in collecting data about how readily children could flexibly alter their utterances in response to communicative demands.

In the final analysis we wished to compare our experimental measures of imitation in both the action and linguistic domains. During experiments one and two we collected measures of repetition imitation (grammatical versus ungrammatical), measures of novel verb imitation and information about transitive structure replication. All these would be related to action imitation performance to try and identify whether imitative proclivity in one was seen also in the others. We expected

that imitation in all three areas would be influenced by the goal-directed intentional behaviors of the model and the causal necessity of these behaviors to the perceived goal. As such we expected there to be strong cross-domain relationships with children's imitative proclivity being consistent in both action and verbal domains. Imitation, of both actions and language, embodies a duality of function. Imitation embodies the process of moving from the actual to the possible. Imitation is also a tool for navigating between the actual and the possible. Imitation is the navigation between the actual and possible but is also an invaluable tool for learning about such navigation.

Experiment 1 focuses on the first task, the imitation of grammatical versus ungrammatical repetitions. Experiment 2 reports the result of our novel verb and varied transitive structure task. Here participants' language comprehension, as well as their imitation of novel lexical items and transitive structure will be reported. In the third section, titled Analysis 3 the relationship between the imitative behaviors recorded in these two tasks will be explored. In addition the imitation data reported in Chapter 5 will also be included as we look for contiguity between imitation behaviors in the linguistic and action domains.

### **6.3.1. Experiment 1: Introduction**

As outlined, acquiring language is a fundamentally imitative endeavor. Although small nuances and generational differences exist within speakers of the same regional accent, the vast majority of a language's conventional features are retained. If this were not the case language would lose its prime cause for existence: the ease and specificity of communication. This is the paradox of language, its' tokens and grammar are imitated, but the way in which they are specifically delivered or combined need not be. It is both highly imitative and highly innovative. However,

imitation is also an integral facet of language development (Bannard, Klinger & Tomasello, 2013). Simple sound effects demonstrated alongside actions are imitated by infants as young as 12-months old (Carpenter, Call & Tomasello, 2005). The prosody of language can be imitated by 3-5 year-olds (Loeb & Allen, 1993). Children can imitate novel nonce verbs (Brooks & Tomasello, 1999). In an experimental setting Over and Gattis (2010), showed that 3.5-year-olds would imitate the grammatical and ungrammatical speech of both an intentional (stuffed animal) and unintentional (coloured cardboard box) model, albeit in distinctly different ways. The authors likened this behavior to the kind of imitation seen in object based behavioral re-enactment paradigms where children imitate the models goals, rather than their actions (e.g. Meltzoff, 1995). Moreover, Kuczaj and Maratsos (1975) reported that children often imitate the grammatical structures used by adults before they appear in spontaneous speech. Indeed children readily adopt the syntax heard from a modeler in structural priming studies (e.g. Vasilyeva, Huttenlocher & Waterfall, 2006). It seems clear then that children's imitative abilities are pervasive and wide ranging from simple actions to complex, abstract features of language. In Study 1 we decided to implement an experimental imitation paradigm to gauge how children would choose to imitate sentences composed of familiar lexical items. We gave them the opportunity to imitate grammatical and ungrammatical repetitions of familiar words, as well as control sentences containing no repetitions.

We replicated the methods used in the aforementioned study by Over and Gattis (2010). In their initial study children aged three and a half imitated the ungrammatical speech of both intentional and unintentional agents (Over & Gattis, 2010). Ungrammatical sentences such as “the cat was too big big for the chair” were

less often repeated verbatim than grammatical sentences such as “Sam was a big big cat.” Children tended to correct the ungrammatical sentences.

Furthermore Over and Gattis (2010) identified that when the “speaker” could have had intentions attributed to it (i.e. a stuffed toy) children were more likely to correct these ungrammatical utterances when compared to the sentences being produced by an inanimate box like apparatus. In this case children were more likely to omit a repetition to retain the intended message whilst removing the excess, “incorrect” element. This provides striking evidence that much like in paradigms involving goal directed action imitation in toddlers, some young language learners were selectively emulating the intended goal sentence in an efficient way, i.e. were producing a semantically identical grammatical sentence when the speaker was an intentional agent: children were engaging in selective imitation. Of course in this case the ungrammatical sentence is conventionally incorrect rather than just causally unnecessary as with the action imitation paradigms. It highlights an important idea, however: children treat both physical and verbal behaviors in a goal directed manner. As we have reported in our object imitation study in the previous chapter, however, children move from an emulative to an imitative strategy in the action domain. Children as early as eighteen months begin to selectively copy not only instrumentally necessary actions but also those intentionally undertaken. Data collected from our four year-old sample would help us determine if a similar shift takes place in language imitation. If the imitation of language were to follow the same trajectory as the imitation of actions then we would expect to see several key features emerge:

(i) Children would more readily selectively imitate causally relevant features of the language i.e. those that achieve an intended communicative goal.



(ii) Children would more readily imitate intentionally produced features of the language.

(iii) Children would faithfully imitate more incidental or non-critical features of language with age.

(iv) Children would imitate a live social model more faithfully than an inanimate or distant model.

(v) Familiarity and or affiliation with a model would influence rates of faithful imitation.

(vi) Attentionally cueing or framing some feature of the language would increase its' imitation.

(vii) Instrumental features of the language that are imitated would be more readily retained than features not intrinsically related to the functionality of the item (e.g. a new word is retained but not necessarily in the construction it was initially encountered in).

Some of these features of language imitation have been observed before. For example recall in Over and Gattis (2010) that: three year olds were more likely to alter grammatically incorrect sentences than matched grammatically correct sentences and; children were more likely to correct such ungrammatical sentences when they were produced by an intentional agent compared to an inanimate box. In another study with three year olds, context determined whether a novel word was beneficial or not to the communicative intention of an utterance (Bannard, Klinger & Tomasello, 2013). Participants imitated novel adjectives significantly more often when the context made them communicatively useful. Children were more likely to use novel adjectives when in doing so they provided a social partner with useful information. These results both support point (i). Bannard et al. (Exp. 3: 2013) also found that

children were less likely to imitate novel adjectives produced “accidentally” by an experimenter. This provides evidence for point (ii) and the imitation of intentional features of language. It is important to remember, however, that despite between condition differences in imitation rates, faithful imitation at high levels was evidenced across condition. This mirrors our, and others’ results, from the action domain in Chapter 5: children at this age most often chose to faithfully imitate actions (e.g. Brugger et al., 2007; Lyons et al., 2007; McGuigan et al., 2011).

Other studies, from quite a different theoretical background have looked at other features of utterances that children may imitate, in addition to familiar lexical items. These include novel verbs, syntactic structure and prosody. In Shimpi et al.’s 2007 study an experimenter described pictures of a dynamic scene to 3 and 4 year-old children. The experimenter used a particular transitive syntactic structure, either active or passive. In some trials children were asked to repeat the model’s sentence before being presented with a new picture that they could describe freely. The type of syntax modeled was more likely to be imitated as children got older. This supports point (iii). Shimpi et al. also reported that asking children to repeat the model exactly, before being given a target describe increased levels of imitation, when compared to trials where no such repetition was prompted. This support point (vi) as repeating the modeled syntactic structure serves to highlight its salience, thereby increasing its imitation.

#### **6.3.1.1. Aims**

For the purpose of Study 1 we chose to focus on the imitation of sentences, some of which contained grammatical and ungrammatical repetitions. We adapted the paradigm from Experiment 2 in Over and Gattis (2010). Each child heard thirteen sentences: eight contained grammatical and ungrammatical repetitions; five contained

no repetitions. Sentences were verbally produced by an experimenter in the context of a game called “Can you say what I say?”

Our first aim was to assess the difference in imitative behavior between the experimental conditions. We expected to find significantly less verbatim imitation and significantly more repetitions removed in the ungrammatical condition than the grammatical condition due to the increased likelihood of selective imitation omitting the non-functional ungrammatical repetition. This would be a similar pattern of performance to the unnecessary condition in Chapter 5 where there were significantly more Just B responses in the necessary condition: the unnecessary element was omitted.

Using this data, and the data collected in Experiment 2 of this chapter we aimed to identify any commonalities between imitative performance in the object and linguistic domains. In Chapter 5, regardless of the causal necessity of actions, children’s most frequent imitative response was faithful to the model. Similarly we expected that at the group level children in both the grammatical and ungrammatical repetition conditions would imitate faithfully, that is repeated the sentence verbatim.

Of course the linguistic and object tasks cannot be directly compared as causal necessity is a different concept than grammaticality: however, there are enough similarities to warrant comparison being drawn in a more general sense. Of course in our verbal imitation task the communicative goal can be achieved with or without the repetitions, particularly in the grammatical condition. In the object imitation task the omission of one of the actions in the necessary condition would result in a goal failure. This would presumably reduce the likelihood of omission when compared to the verbal task where communication could proceed successfully regardless.

#### **6.3.1.2. Methods**

#### **6.3.1.2.1. Participants**

The same participants as presented in section 2.3.1 took part.

#### **6.3.1.2.2. Materials**

To overcome shyness and introduce children to the task six practice sentences were included. The first two sentences contained one word, the second pair contained two words and the final pair contained three words. For the experiment proper thirteen sentences were taken from Over and Gattis (2010), Experiment 2. Eight of these were test sentences that contained repeated words; four sentences containing grammatical repetitions, e.g. ‘Sam was a big big cat’, and four containing ungrammatical repetitions, e.g. ‘The cat was too big big for the chair’. Also included were five grammatically correct filler sentences each containing five words but no repetitions, e.g. ‘I like to eat cake’. Examples of the practice, filler and test sentences can be found in Appendix F.

#### **6.3.1.2.3. Procedure**

Parents and children were invited to our on-campus lab for their second visit. These visits began with a short warm up time before the experimental tasks were introduced in a separate room. Tasks were administered in a purpose built testing room with inbuilt digital audio and visual recording capabilities. The complete verbal imitation procedure was the first task of the session. The participant and experimenter sat opposite each other, either side of a square testing table covered in a black table cloth.

Children were introduced to the game ‘Can you say what I say?’ The single word practice utterances were produced one at a time by the experimenter for the participant to repeat, followed by the two and three word utterances. Following the practice children were reminded of the game ‘Can you say what I say?’ and the

thirteen experimental sentences were introduced one at a time. If after three repetitions children did not respond the experimenter moved on. Children were praised for each response they gave, regardless of what the response was. The sentences were presented in a pseudo-random order: filler, grammatical and then ungrammatical so that similarly structured sentences were not produced in too close proximity to one another. There were five different scripts with differing combinations of sentences but retaining the same type order.

#### **6.3.1.2.4. Data Analysis**

Verbal responses were transcribed from the audio recordings and then coded. 20% of videos were secondary coded and inter-rater reliability was at 100%. Participant's were then given three scores to reflect how imitative they were overall, in response to grammatical repetitions and, in response to ungrammatical repetitions. Non-verbatim responses were then further coded to reflect the type of alterations made; *repetitions removed* (removal of the target repetition only), *repetition additions* (addition of repeated words over and above those in the original utterances), *block omissions* (removal of one or more words other than the repeated item alone), or *other alterations* (changes to the content).

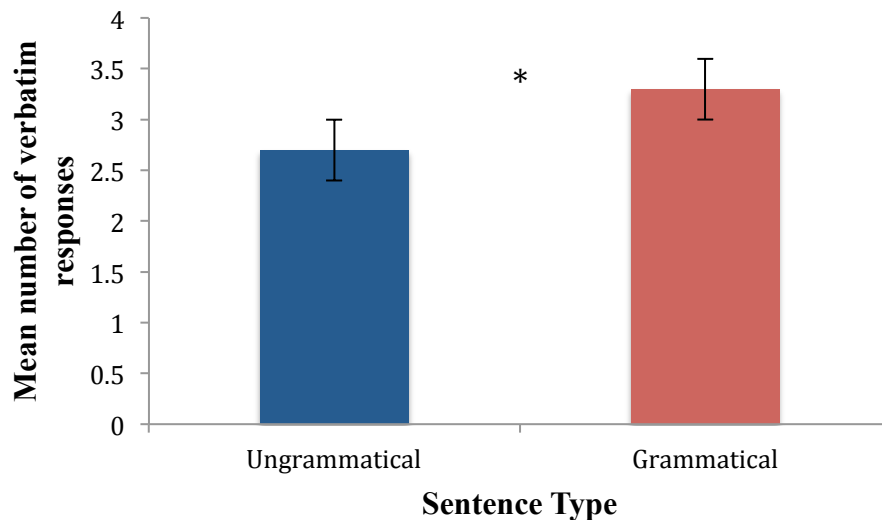
#### **6.3.1.3. Results**

To begin with we examined the levels of verbatim imitation in our sample (see Figure 6.1.). Verbatim responses for the test conditions containing repetitions were scored between 0-4. There was a significant main effect of condition on imitation rates using a one-way ANOVA,  $F(2,159)=47.75$ ,  $p<.001$ . Mann-Whitney U comparisons were carried out between the target conditions and grammatical sentences were significantly more often recounted verbatim than ungrammatical sentences ( $U=929.5$ ,  $p<.001$ ). The four sentences containing grammatical repetitions

were recounted verbatim at a mean of 3.3 whilst the mean for sentences containing ungrammatical repetitions was the lowest at 2.7. Grammatical sentences and ungrammatical sentences were indeed treated differently with higher levels of imitative fidelity in the former when compared to the latter. In addition to a difference in verbatim imitation between the experimental conditions, repetitions were significantly more often removed if utterances were ungrammatical when compared to grammatical ( $U=1020$ ,  $p<.001$ ) (see Table 6.1. And Figure. 6.2.).

Table 6.1.  
*Table Showing Mean Imitation Scores and Alteration Types in the Grammatical (G) and Ungrammatical (U) Conditions.*

|                   | Condition | Mean | Standard Dev. |
|-------------------|-----------|------|---------------|
| Total Imitation   | G         | 3.3  | 1             |
|                   | U         | 2.7  | .98           |
| Reps Removed      | G         | .20  | .66           |
|                   | U         | .56  | .74           |
| Reps Added        | G         | .06  | .236          |
|                   | U         | .39  | .566          |
| Block Omission    | G         | .59  | .926          |
|                   | U         | .52  | .846          |
| Other Alterations | G         | .24  | .586          |
|                   | U         | .26  | .48           |

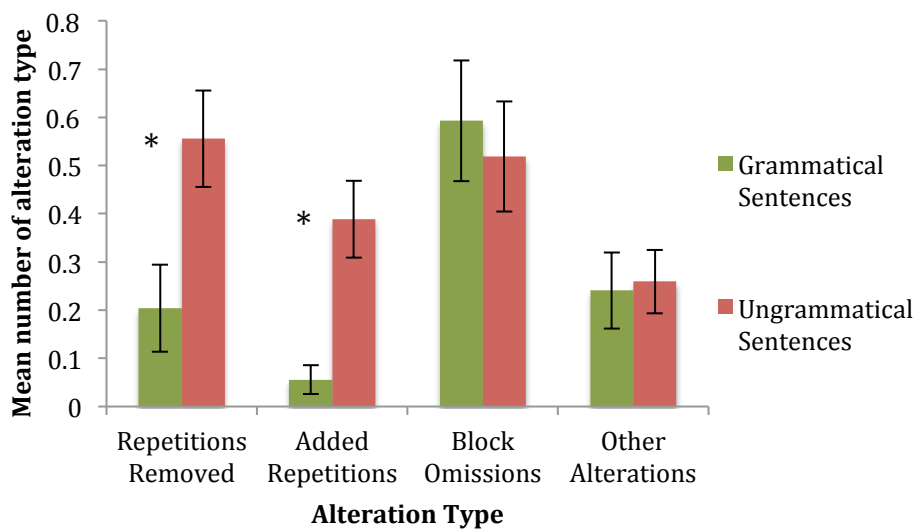


*Figure 6.1.* Mean number of sentences recounted verbatim on the vertical axis for each of the target model types on the horizontal axis: grammatical and; ungrammatical. Error bars represent standard errors. \* Represents a statistically significant difference between the bars  $p < .05$ .

Next we wanted to briefly explore the other types of sentence alterations that the children produced during the experimental procedure (see Table 6.1). Interestingly in addition to there being an increased number of repetitions removed in the ungrammatical condition there were also significantly more repetitions added than in the grammatical items ( $U=1188$ ,  $p < .01$ ) (see Figure 6.2.).

In the grammatical condition there were more repetitions removed than added with the difference being marginally significant,  $t(53)=1.74$ ,  $p=.09$  In the ungrammatical condition there were comparable numbers of added and removed repetitions,  $t(53)=1.2$ ,  $p > .05$ . This points to the grammaticality of the repetitions, rather than simply the presence of repetitions being meaningful and influences how participants respond.

Block omissions and other alterations were at a comparable level for grammatical and ungrammatical sentences ( $U=1289$ ,  $p=.21$ ). This supports the theory that it is not the length or complexity of these sentences that are different, it is something unique about the grammaticality of the repetitions and how those repetitions are dealt with that is altered.



*Figure 6.2.* Mean number of each alteration type made during grammatical and ungrammatical condition repetition. Error bars represent standard errors. \* Represents a significant difference between bars  $p < .05$ .

#### 6.3.1.4. Experiment 1: Discussion

Our first aim for this experiment was to characterize the patterns of imitation in our sample and to test the impact of grammatical versus ungrammatical repetitions on children's imitative responses. As we had predicted at the group level verbatim imitation was the preferred strategy for the target grammatical and ungrammatical sentences. This result mirrors children's preferred imitative strategies in the object



imitation task where again faithful imitation (A then B responses) were the most common across condition.

We had further expected to record fewer verbatim repetitions in the ungrammatical condition when compared to the grammatical condition. This group level pattern highlights the importance of communicative goals in imitation: omitting an ungrammatical repetition would retain the communicative goal of the utterance while selectively removing communicatively redundant material. As predicted, a between condition difference in verbatim repetition was found, with there being significantly more faithful imitation of the sentences containing grammatical repetitions when compared to their ungrammatical counterparts (See Figure. 7.1.). Over and Gattis (2010) also reported this difference, though of a greater magnitude. This relates to the first feature of imitation we remarked on in the introduction: children more readily selectively imitated causally relevant features of the (behavior) language. They more often altered sentences that were ungrammatical than they altered perfectly grammatical sentences, by doing this they retained the semantic, communicative goal whilst omitting intentional ungrammatical element.

In addition we showed that repetitions were more often removed in the ungrammatical condition than the grammatical condition by our sample. This result replicates the findings of Over and Gattis (2010) who also reported increased levels of repetition removal in the ungrammatical condition. However, the additional types of alterations made contrast with the results reported in Over and Gattis (2010). In the coding scheme applied to their sample of three year-old children if an ungrammatical sentence was altered it was in order to be corrected to a grammatical form. In our sample of four year-olds the children made a variety of alterations including

additional repetitions that were significantly more common in the ungrammatical condition.

We expected that the most common type of alteration in the ungrammatical condition (as in the necessary condition of the object task) would be to omit the redundant/causally unnecessary element in the pursuit of communicative (instrumental) clarity. We expected children to behave in a way consistent with only omitting the repetition while retaining all the other functional and semantic elements of the utterance. This would be a similar pattern of performance to the unnecessary condition of the action imitation task where there were significantly more Just B responses than in the necessary condition: the unnecessary element was omitted whilst those functional elements were retained.

However, we did not find the expected significant difference between the incidences of repetitions being removed from the utterance and the incidences other types of repetition. The expected removal of repetitions was at a comparable level to the unexpected addition of repetitions. When the number of repetitions removed and added were compared between the grammatical and ungrammatical condition, however, the results were striking (see Figure. 6.2.). As expected there were significantly more repetitions removed in the ungrammatical condition. In addition there were significantly more repetitions added in the ungrammatical condition when compared to the grammatical. These two poles of behavior were both more frequent in the ungrammatical condition. Block omissions and other alterations, however, were at a comparable level across conditions.

As mentioned earlier the increased number of repetitions being removed in the ungrammatical condition was expected and replicated earlier results (Over & Gattis, 2010). The additional repetitions in this same condition were at first surprising.

However, when reflecting on the importance of goals and intentions for imitation a potential explanation became apparent.

The usual function of language is as a communicative tool. Grice's Maxims set out a useful set of assumptions that language learners expect and adhere to when communicating with one another (Grice, 1975). When speaking about the manner of utterances in particular Grice stated that they be brief, being only as informative as is needed in the situation: repetitions and certainly ungrammatical ones would intuitively seem to go against this maxim. However, language can also be a tool for play and humor within a larger communicative context. In the context of our games children may have become more sensitive to the potential elements of playfulness in our language and embellish their linguistic choices accordingly. Recall that children's imitation relates to the hierarchical organization of goals (Bekkering et al., 2000). Further the intention of a social partner is crucial for imitation as highlighted most elegantly in behavioral re-enactment paradigms (e.g., Meltzoff, 2005). If these two ideas are combined it may be that children: (a) understanding the adult model to be a competent, or even expert, language user, infer that they meant to produce the ungrammatical repetition and; (b) having inferred that the repetition was intentionally produced, further infer that the goal of the utterance was to produce a repetition in the pursuit of play. As a result children may have organized their representation of the utterance around the production of a repetition, rather than the conveyance of some semantic information. Having reached this point then children may engage with this goal and extend the model repetition further, then perpetuating or engaging with the inferred game. Children made inferences about the actual words the experimenter chose and based upon those inferences selected from a number of possible utterances they themselves could make in response. Children are using information about the a

speakers passed intentions and goals to inform their own behavior, and their own goals in the future. Such an opportunity would not have been available in the context of the object imitation task described in Chapter 5, as there were only two possible actions for participants to carry out.

### **6.3.2.1 Experiment 2: Introduction**

In Experiment 2 we chose a more challenging linguistic paradigm in order to probe the limits of our sample's imitative abilities. Based on Experiment 1 we can conclude that children can indeed imitate simple sentences of a certain length made up of familiar lexical items and containing repetitions, grammatical or otherwise. However, for the purpose of our investigation into language imitation we chose to focus not only on the imitation of simple sentences, but also on the imitation of novel words and syntactic structure. By using a paradigm that combined an unfamiliar structure and novel words we had two main aims to use an experimental method of assessing linguistic competence, both receptively and productively and to have additional measures of language imitation. In addition, using novel verbs would allow us to observe how children dealt with a new lexical item, would they use it flexibly through the application of general abstract rules (e.g., Chomsky, 1968) or be conservative in its usage, preserving the original syntactic frame as per the verb island hypothesis (Tomasello, 1992)? The syntactic structure we chose was the transitive passive voice and an introduction to it follows.

#### **6.3.2.1.1. Introduction to the Passive Verb Construction**

Transitive verb constructions broadly involve an action being carried out by one object on another. The active transitive construction focuses the listener's attention on the agent carrying out an action (see example. 1a). Conversely the passive construction directs attention toward the patient of an action (see example.

1b). Tomasello (1998) has highlighted the appreciation of how the order of an utterance's constituents can manipulate attention in such ways as the essence of truly understanding communicative intentions.

1a. Mark designed that building

1b. That building was designed by Mark

In the context of syntactic transformations, the passive has been described as the “grand-daddy of them all,” (Ross, 1974). Passives are structurally more complex than their active counterparts. They constitute the marked form of an expression, with word order being reversed. The most common passive form in everyday speech is the truncated passive, (Horgan, 1976; Jespersen, 1924, p.168; Slobin, 1963; Svartvik, 1996). These truncated, or short-passives, have no active subject, containing only one noun or pronoun explicitly (see example. 2a). Full passives then have both subject and object present in the utterance (see example. 2b). Not only are truncated passives more common, they are also easier to comprehend (Harris, 1976) and are recalled more successfully (Slobin, 1963).

2a. The window was broken

2b. The window was broken by the tree

Full passives, those with both subject and object included, can be semantically reversible or non-reversible (Horgan, 1976). Reversible passives often include two animate actors (see example. 3a) where either could logically be in the agent or patient position. Non-reversible passives involve scenarios where it would be semantically implausible to reverse the noun positions (see example. 3b).

3a. The boy was hugged by the girl

3b. The tree was hugged by the girl

In addition to the inclusion and arrangement of nouns and/or pronouns the characteristics of the verbs themselves can differ. The most studied dichotomy is that between actional and non-actional (or experiential) verbs. Actional verbs result in an altered end state such as kicked, pushed or broken. Non-actional verbs are not similarly perceptible to an observer, they are experienced by the patient, or describe a state of affairs, for example seen, heard or ruined. Children have been reported to be more productive with actional verb passives than non-actional (Pinker, LeBeux & Frost, 1987). Non-actional or experiential verbs can be further categorized into; perceptual (e.g., noticed, ignored), cognitive (e.g., believed, understood), or affective (e.g., trusted, admired) (Sudhalter & Braine, 1985). Another type of distinction within non-actional verbal passives recently investigated was the contrast between theme-experiencer and experiencer-theme passives (Messenger, Branigan, McLean & Sorace, 2012). Experiencer-theme passives have the experiencer in the object position (4a), where as theme-experiencer have the opposite construction (4b.) They pose differing levels of difficulty for comprehension.

4a. The fairy was seen by the girl

4b. The fairy was frightened by the girl

#### **6.3.2.1.2. Developmental Patterns with the English Passive Verb Construction**

English speakers are not typically seen to produce spontaneous full passive utterances until aged 4 or 5 years (Harris & Flora, 1982; de Villiers & de Villiers, 1978). This poverty is mirrored in child directed speech with only .36% of utterances being classified as passive in one natural language study (Gordon & Chafetz, 1990). The occurrence of full reversible passives is particularly restricted with none reported in a corpus of 12'000 utterances from five year-old participants (Harwood, 1959). Lovell and Dixon (1967) reported low levels of comprehension and production of full

English passive verb constructions in 4.5-year-old subjects. Similarly Turner and Rommetveit (1967a), found that the full passive was comprehended only 25% of the time in nursery aged children, and produced in only 20% of these same children using an imitation paradigm. Bever (1970) has suggested that pre-school children tend to interpret all noun-verb-noun statements as agent-action-patient relationships.

Truncated passives, however, with one explicit noun or pronoun, and non-reversible passives seem to appear earlier in speech, and have been elicited from the majority of 3-year-olds (Slobin, 1963). Furthermore truncated passives are also easier for children under 6-years of age to understand (Slobin, 1963). Of those passives recorded in child directed speech the vast majority were indeed truncated forms (Gordon & Chafetz, 1990)

Some verb types are also more difficult to comprehend in the passive construction. Non-actional verbs have been shown to result in considerably more difficulty for children and even teens (Maratsos, Kuczaj, Fox & Chalkey, 1979; Sudhalter & Braine, 1985; Becker & Maratsos, in Maratsos & Chalkey, 1980). This effect of non-actional verbs appears not to be due to the verbs regularity or frequency in speech, it is the nature of the transitive action itself that interferes with performance (Sudhalter & Braine, 1985).

It is clear from the many studies investigating the passive that although evidence of productivity and comprehension can be found in the pre-school years, there are many factors influencing performance at the group level. In addition there is a large range of individual differences within tasks and age groups. As such there have been extensive studies on passive voice comprehension and production at the group level that have served to highlight some of the ways task contents and demands can mediate performance. This shall be discussed further in section 8.1.4. Language

production tasks are generally harder than comprehension tasks (Hirsh-Pasek & Golinkoff, 1996). This seems to indeed be the case for the passive construction (Flavell, Beach & Chinsky, 1966; Turner & Rommetveit, 1967; however, see Messenger et al., 2012 for a counterexample). In the following section I will detail further experimental studies elucidating how task demands and differing paradigms change young children's performance in comprehension and production of the passive construction.

#### **6.3.2.1.3. Abstract Schemas vs. Verb Islands: Assessing Productivity with the Passive Experimentally**

Baseline measures of how children choose to freely describe verb-action sequences consistently identify the preference they hold for utilizing active verb forms over passive forms (e.g., Kidd, 2012; Savage, Lieven, Theakston & Tomasello, 2006; Shimpi, Gamez, Huttenlocher & Vasilyeva, 2007, and Whitehurst et al., 1975). These unconstrained opportunities for children to describe pictures have been used as comparisons for studies in which the use of novel, less frequent or unfamiliar verb forms is elevated. Structural priming is described as the tendency for individuals to generate new utterances while retaining the structural features of other recently heard speech (Bock & Griffin, 2000). Priming studies are one of the linguistic world's equivalents of imitation paradigms. Priming studies use models of a particular syntactic structure to investigate whether individuals or groups of participants hearing them will subsequently incorporate some, or all of the structural features into their own utterances. This feature of conversational language, incorporating and imitating structural elements heard in the speech of others, is both abundant and inadvertent in normal speech (Bock & Griffin, 2000). Indeed Bock and Griffin argued that the act of structural repetition following another speaker is implicit, independent from intention,



effortless and without need of explicit attention to form in adults (Bock & Griffin, 2000). In essence then priming is considered by some as an automatic feature of natural, social language. Further investigation of the features and hypotheses of structural priming more generally will occur later in this chapter.

Developmental productivity with the passive has previously been assessed through the lens of such paradigms. One hypothesis derived from their use is that if children hear a passive model they will only readily adopt it in a different semantic context if an abstract representation of the construction is already in place. For example if a child hears a passive sentence like “*The lemon is being cut by the knife,*” then sees a picture of a knife and a loaf of bread, they will only describe the picture as, “*The loaf of bread is being cut by the knife,*” if they have already developed an abstract representation of the passive construction. Resultant evidence that young children do indeed represent syntactic structure abstractly has been put forward by numerous authors (e.g. Bencini & Valian, 2008; Gentner, Fisher & Eisengart, 2000; Shimpi et al., 2007; Thothathiri & Snedeker, 2008).

In contrast to these early abstraction accounts, item-based accounts suppose that young children are not able to make the abstract connection between verbs and instead are incredibly conservative with their verb construction use. The verb island hypothesis introduced earlier is an example of such item-based accounts (Tomasello, 2000). In the subsequent paragraphs I will outline a sample of the experimental works, many using priming paradigms, which can speak to these accounts of early grammatical development.

In one of the first experimental studies of passive verb acquisition Whitehurst et al. (1974) modeled English passives over five sessions. Four to five year old children listened to these models that described drawings of actors and actions. The

participants were prevented from overtly imitating the modelled speech. Following these models the children were shown probe drawings of different actions, some they were asked to describe and others simply identify. This procedure generated both production and comprehension scores for the children. In comparison to performance baselines prior to the modeling sessions both production and comprehension of passive verb constructions improved. All participants in the experimental condition produced at least once passive utterance to describe the probes. No children in the control group that had received no passive models produced any passives, although their comprehension was still above chance. Whitehurst et al. attributed this increase in production to selective imitation (also characterized as a form of abstraction in their paper) that is imitating the grammatical structure whilst excluding the content (Whitehurst et al., 1974). They claimed that exposure to the structure also increased comprehension. In a study of younger, three-year-old children, Shimpi et al. (2007) used an altered paradigm to produce similar priming results. Shimpi et al. used fewer passive verb models, had participants repeat the modeled sentences and inserted target pictures after each model rather than after blocks of models. The participants showed increased passive productivity following models. They concluded that three-year-olds also had an abstract representation of the passive construction, however, it was less easily accessible than for slightly older children (Shimpi et al., 2007).

Thatcher et al. (2007) extended these findings with evidence of cross verb priming in younger children (aged 3), and with varying semantic classes of verb. Both actional and non-actional verbs modeled in the passive construction increased participants' use of the passive structure in subsequent utterances. Thatcher et al. also attribute these results to children possessing an abstract representation of the passive that was not tied to specific verbs, or even semantic classes (Thatcher et al., 2007).

Different forms of the passive, for example Be passives vs., Get passives, also show cross item priming, although stronger from Be to Get than vice versa in six-year-olds (Messenger, 2009). Additionally truncated passives can prime full passives and vice versa (Messenger, 2009).

There are several key aspects of such paradigms which could bolster the effects seen – namely the increased production of passive verb forms. Firstly the experimental setting itself dictates that children and experimenters are interacting in a dyadic manner. Priming has previously been conceptualized as a communicative, imitative and social process (Pickering & Garrod, 2006). Children will be particularly sensitive to, and therefore more likely to imitate, the features of speech that are directed to them (Branigan, Pickering, McLean & Cleland, 2007). Being the member of a group that is addressed during storytelling also increases productivity of the passive verb form (Vasilyeva, Huttenlocher & Waterfall, 2006). The number of models heard can influence the likelihood of adoption also (Bencini & Valian, 2008; Pickering & Branigan, 1998). If, as Brooks and Braine (1995) argue, children build models of syntax through extracting commonalities across numerous exemplars of a construction, the more exemplars present, the more opportunities for syntactic learning (see also structure mapping processes hypothesized by Tomasello, 2000). Children begin to categorize verbs and as such draw analogies between newly encountered verbs and more familiar entrenched forms, potentially supporting future generalizations. With the use of familiar items then there may be individual differences in the number of exemplars of its' use already in long-term memory. Additionally with familiar lexical items participants are not being as taxed in terms of their understanding of the actions itself, therefore have more working memory to assign toward processing features of syntax. If children are familiar with the verb, in

any form, they will be more able to attend to the structural rather than semantic features involved.

Studies such as those described above were interested in the circumstances under which young children would adopt an unfamiliar syntactic structure they had heard in the speech of an experimenter. However, many of these studies used familiar verbs in these structures. Indeed the majority of passive production studies with children utilize familiar English verbs (e.g., Bencini & Valian, 2008; Bock & Griffin, 2000; Branigan, Pickering, McClean & Cleland, 2007; Huttenlocher, Vasilyeva & Shimpi, 2004; Kashack & Glenberg, 2004; McClure, Pine & Lieven, 2006; Kidd, 2012; Messenger, Branigan & McClean, 2012; Pinker, LeBeaux & Frost, 1967; Savage, Lieven, Theakston & Tomasello, 2003; 2006; Thatcher, Branigan, McClean & Sorace, 2007; Vasilyeva, Huttenlocher & Waterfall, 2006; Whitehurst & Novak, 1973.)

Several attempts have been made to further investigate the abstract nature of children's grammar and their proclivity with unfamiliar items, by upping the task demands (e.g., Brooks & Tomasello, 1999; Childers & Tomasello, 2001). One option for doing this is to use an unfamiliar construction in conjunction with a completely novel nonce verb. Nonce verbs give the researcher a glimpse into how participants interact with an entirely new lexical item. To be used flexibly the nonce item must be recognized as coming from the transitive verb class, and applied as such in a different context. The first productive use then could be to retain the verb construction but describe different nominals. For example the verb meek being modeled in an active transitive to describe the motion of a trolley being pulled across a ramp by a lever: "*The hammer is meeking the boy.*" This could be used productively in another situation where "*the elephant is meeking an aeroplane.*" In addition the verb itself

could be modified according to grammatical rules already in place for such verbs. Going even further then by altering the construction used to describe the event could see a passive sentence like “*the aeroplane is being meeked by the elephant,*” being produced.

Children can incorporate novel verbs learnt in experimental settings into their own novel grammatical utterances (Childers & Tomasello, 2001). Between two and four-years-old this ability develops to a point where children can begin to flexibly use novel verbs in a transitive way at the group level (Abbot-Smith, Lieven & Tomasello, 2004). Brooks and Tomasello (1999) gave two and three-year-old children a large volume of passive or active construction models with a novel verb and action. Participants most often attempted to retain the structure of the verb that had been initially presented, 75% of those who had been trained with a verb in the passive construction subsequently used it in that construction. However, around 1/3 could be productive, that is move independently between constructions under certain task demands. The authors concluded that these productive participants had general construction knowledge about the active and passive robust enough to support flexibility. However, in their set of experiments children were exposed to huge numbers of active and passive utterances. Despite this Brooks and Tomasello argued that young children are still extremely conservative and vary their use of verbs in line with the adult models they encounter. They imitate the speech they hear in a clever way in that they often rely on the linguistic context to facilitate their productivity (Tomasello, 2003, p75-76).

The move children make, from extreme conservativeness towards productive flexibility is suggested to take place between three and five years of age (Brooks & Tomasello, 1999). Children in their earliest stages of multiword utterances retain new

verbs in the structures that they were initially presented. After this stage then children begin to represent the form-meaning similarities held within the many utterances they have been exposed to through the aforementioned process of structure-mapping (Tomasello, 2003). The relational structure between different verbs and the constructions in which they appear can be extended to novel verbs discerned to be in a similar semantic category (Brooks & Tomasello, 1999b). Children suddenly are able to imitate a novel item in an appropriate and innovative manner when in an altered context as opposed to mimicking the verb and its structure in an inappropriate way given the new discourse demands.

#### **6.3.2.1.4. Imitation versus Innovation: How Language Sits Astride the Actual and the Possible**

An interesting contradiction can be evidenced here in situations where the goals of language are manipulated. Suppose children do have a robust and easily accessible abstract structural representation of the passive transitive form. Suppose they can readily imitate both passive and active utterances, be they containing English or nonce verbs: how far might this desire/automatic propagation to retain the structure previously heard extend into maintaining it in the presence of incongruent discourse pressure? In a second study in which both active and passive constructions were modeled some children showed evidence of being able to alter the structure of their descriptions in response to the experimenter's phrasing of questions: they responded to discourse pressure (Tomasello & Brooks, 1999a).

In the active construction speakers and listeners are attending to the agent of the action, therefore in unconstrained questioning will have no reason to alter this emphasis. With agent-focused questions too then, the question type is congruent with the initial construction and the representation of the action. If an interlocutor,

however, asks, what is happening to the patient of the action, the focus of attention switches: so too can the language used to describe it switch. Patient focused questions, with their emphasis on the patient lend themselves to a response using the passive construction. Here the goal of language has clearly been mapped for the listener, therefore potentially changing the hierarchically organized representation of the task. For clarity let us look at an example below (Figure 6.3.).



Figure 6.3. Illustration to show the effect of question focus on syntactic constructions.

A neutral unconstrained question would ask, “What is going to happen now?” The most likely response would be an active transitive utterance such as “Minnie is going to hit the ball.” An agent-focused question asking what Minnie is going to do would increase the likelihood of this form of answer. Asking a patient focused question, however, “What is going to happen to the ball now?” would be expected to elicit an increased number of passive utterances like “*the ball will be hit/get hit by Minnie.*” Language being modified in this way eases fluency between speakers by retaining the focus of attention. It allows the intentionally salient feature often the agent in an action sequence to be identified and information about it exchanged in a parsimonious way. This is the goal of language in the current situation. This prevents inferences having to be made about what happened to the ball based upon an

utterance focused on Minnie, who for whatever reason, may not be the questioner's priority.

Not only verbal discourse pressure can produce such results. The presentation of prompts in pictorial form of the patient of an action has been shown to trigger passive voice recall (Turner & Rometveit, 1968). Additionally pre-action direction of attention through language can also increase the likelihood of passive utterances being produced in an experimental setting (Turner & Rometveit, 1968). In this way linguistic behavior emulates object imitation. Behavior is based upon the salience of different goals of the interaction and these goals can be made more or less salient for participants as explored in Chapter 5.

Altering their language in response to modulation of attention in this way, although possible, can be difficult for majority of children under 5-years-old using English verbs and novel verbs alike (Brooks & Tomasello, 1999; Brooks, Tomasello, Dodson & Lewis, 1999; Vasilyeva, Huttenlocher & Waterfall, 2006). However, Brooks and Tomasello (1999) showed that some children younger than three are capable of being productive and flexible with novel verbs. 35-40% of children could produce the verb learnt in one construction, in the opposite construction under discourse pressure with 90% of those passives in response to neutral and patient focused questions (see also Pinker et al. 1987, who used a teaching paradigm).

Responding to discourse pressure, however, comes at the cost of imitation: Or does it? If a person has been speaking about agents, listeners will form an expectation of action agents being focused upon, based on the statistical evidence from these previous utterances. This expectation would in-turn encourage the listener to attend to agents and subsequently form speech focusing on agents. Such expectations, based on the subject-verb-order of prior utterances, have been identified in young children



through preferential looking paradigms (Thothathiri & Snedeker, 2008). In addition imitation is often used as means to initiate, maintain and repair social interaction. If, however, a social partner signals that some other behavior is desirable then injecting some innovation is a positive response.

Responding to discourse pressure displays linguistic, conversational and even social sophistication: it indicates that you can quickly and effortlessly track and respond to attentional-focus changes or any other subtle cues generated by a conversational partner. Tomasello (1998) describes understanding communicative attention as being that very ability, understanding how a social partner means to manipulate your attention. Such a skill, being able to over-ride your initial automatic willingness to retain structural parity, may in itself be an indicator of great things. Remember that priming of this sort is often considered an automatic process (Bock & Griffin, 2000). Rejecting the modeled construction in favor of allying with discourse pressure requires inhibition of the pre-potent form and subsequent flexibility in generating a productive utterance based on the application of abstract rules about grammatical verb transformations. Indeed, some researchers have suggested priming itself is a result of a lack of executive control (Pickering & Ferreira, 2008). Tomasello (2000) called this interplay between imitation and innovation a paradox of language, as mentioned earlier. Children must be conventional in order to successfully convey the information being requested of them, using the correct lexical items, arranged in a coherent syntactic structure. However, children must also be innovative in rejecting previously heard utterance constructions in favor of the one that is most appropriate for the task in hand.

Perhaps responding to such instances of switching focus, has benefits not only for language but other cognitive processes too such as cognitive flexibility.

Huttenlocher, Vasilyeva, Cymerman & Levine (2002) reported that the complexity of caregiver speech, using multi-clause sentences for example, relates positively to their infant's syntax. It stands to reason then that there will be individual differences in the strategy children take, in terms of what, when and the extent to which they imitate. It may be individual differences in basic cognitive functions such as inhibition or cognitive flexibility that determine how able a child, who possesses the necessary abstract grammatical rules, can apply them in situations of discourse pressure.

A speaker has made numerous intentional decisions that a listener may or may not choose to imitate; single words, word order, syntax, what to talk about and even to commence speaking at all. Again, however, children cannot be flexible with rules that they do not possess. Free from linguistic constraints children may have greater or lesser implicit social desire to imitate that extends not only to the use of a new verb but also the construction it was used in. They are socially motivated to use a novel verb in a passive form regardless of the question focus. Children may do this regardless of their ability to change the structure, relying only on the pattern previously exhibited by the modeler. They may, however, correctly identify that the focus of attention has shifted for whatever reason and instead of retaining the passive construction apply their abstract grammatical knowledge to this newly learned verb and produce the contrasting form. Children who are perhaps less able to do this may still respond to discourse pressure but being unable to apply rules (either they are not there to begin with or are difficult to access and deploy) may answer in a linguistically congruent form but using a familiar English verb. Childers & Tomasello (2001), suggest that a verb schema may be in place in children that is strong enough to support comprehension processes but not production. Such hypotheses are supported by the work of Thothathiri and Snedeker (2008) who provided evidence of

comprehension competence ahead of productive abilities. Regardless of social motivation then some children may struggle to imitate certain features of speech, preferring instead familiar lexical and syntactic options.

Children may not copy the novel verb at all, but retain the construction. The key here would be seeing English verbs used in passive constructions. Perhaps retaining properties of syntax, is for some kids, more salient than a new verb? This option seems unlikely given children's rapid uptake of new words in natural language. Aside from all these possible patterns of response there may also be increased numbers of incomplete or intransitive utterances, as children struggle to manage the many working memory demands being placed upon them.

The combination then of priming studies and the application of discourse pressure allows for multiple facets of language production to be assessed concurrently. Additionally the use of a novel verb allows for researchers to investigate, in real time, how children incorporate a new lexical item into their subsequent utterances, across contexts and under differing discourse demands.

#### **6.3.2.1.5. What Facilitates the Imitation of Syntactic Structure?**

It is clear that children struggle with the spontaneous production of passives until late childhood and performance can be facilitated by a number of factors. At this point it would be useful to remind ourselves of the factors that influence the imitation of actions. Object imitation has been shown to respond to experimental manipulations of: (i) social behavior: aloof or socially unresponsive models reduce imitation whereas warm experimenters facilitate it, social cues such as eye-gaze also increase imitation (Bandura, 1971; Brugger, Lariviere, Mumme & Bushnell, 2007; Hartup & Coates, 1970; Nielsen, 2006; Nielsen, Simcock & Jenkins, 2008); (ii) model characteristics: the model's age, social background and even language spoken can

influence imitation rates (Buttelmann, Zmyj, Daum & Carpenter, 2013; Jaswal & Neely, 2006; Kinzler, Corriveau & Harris, 2011; Ryalls, Gul & Ryalls, 2000; Seehagen and Herbert, 2011; Zmyj et al., 2012); (iii) increasing goal salience: looking behaviors, novel language, repetition, verbal framing, synchronicity and timing can all alter imitation rates (Brubacher et al., 2013; Chen & Waxman, 2013; Elsner & Pfeifer, 2012; Hermann et al., 2013; Simpson & Riggs, 2011). Such factors would presumably also impact upon the imitation of syntactic structure.

Indeed experimental paradigms that manipulate the volume of passives present in the input, have been shown to impact upon the rate at which they appear in the participants' subsequent utterances (Messenger, 2009; Vasilyeva et al., 2006). Such repetition of the structure would serve to increase its salience, and consequently the likelihood of imitation in much the same way as happens in the object domain. Furthermore being instructed to imitate a modeled passive increases subsequent production for three year-olds. Having the children use the structure increases its salience and availability (Shimpi et al., 2007).

Features of the agents and patients used by experimenters modeling the construction have also been reported to influence performance. In experimental settings utilizing pronouns consistently aids four year olds to produce passive utterances when compared with nouns and pro-nouns (Savage et al., 2003). Pronouns seem to scaffold performance from model to production. A mix of both pronouns and nouns in a series of examples has also been shown to benefit production (Childers & Tomasello, 2001). This facilitatory effect of pronouns has been used as evidence to support the verb island hypothesis and by having one stable element of the utterance children can more readily substitute other elements.

Model utterances with animate patients produce poorer comprehension and production outcomes for children (Lempert, 1990). Animacy of the agent is a more common feature of passive utterances for children therefore more readily learnt (when compared to an inanimate object acting on something). Moreover common or familiar agents and patients act to scaffold performance through a suggested form of lexical scaffolding for children in much the same way as seen with the use of pronouns (McClure, Pine & Lieven, 2006).

It is not only features of the agent and patient that can influence production performance. The type of verbs themselves being described can also boost or hinder production in children, with actional being more readily produced than non-actional passives (Sudhalter & Braine, 1985). Having a variety of individual verbs modeled for participants further provides a performance boost through highlighting the common structural features rather than semantic features (Savage et al., 2006).

#### **6.3.2.1.6. Why Might the Imitation of Language Structures Itself be Useful for Young Language Learners?**

In Chapter 5 we introduced the idea that through imitation children are able to stand on the shoulders of giants. Children need not be innovative in every situation they find themselves in because they can model their behaviors on the behaviors of those who have gone before: an efficient strategy. Of course for language production each token is imitated due to its inclusion in said language's lexicon. However, there are many features of language that children can imitate: speaking at all; which of a selection of lexical items to choose; prosody; pronunciation or; syntax -- to name a selection. Hearing others use language in a certain way in a sense invites us to use it in a similar way. Observing language, how it is used and the subsequent physical or social effects that it has, is an invaluable tool for young language learners. Children

can learn about language in a vicarious way that they can then emulate or imitate immediately or on a later occasion.

There are several papers investigating why structural imitation in particular is so prevalent in human speech. Suggestions include it being the result of a basic phenomenon such as short-term cortical activation, where hearing a certain syntactic structure activates the associated production neurons (Chang, Dell & Bock, 2006). This spreading activation would then potentiate the structure being reproduced in the listeners' speech above a competing form. Crucially, however, this effect would be short lived and not impact upon processing outside of the situation. Another possibility is that priming of this type is an indicator of longer-term implicit learning. The cognitive processes that generate utterances adapt as a result of linguistic input and subsequent productive utterances reflect this adaptation. There is evidence from both adults and children that imitation effects do persist for longer than the experimental session itself (Bock & Griffin, 2000; Boyland & Anderson, 1998; Savage et al., 2006). Pickering and Garrod (2004) also emphasized a social motivation for such behavior.

Pickering and Ferreira (2008) have set out three main reasons why the behavior of priming may be beneficial for conversational situations. (i) It promotes alignment between speakers both in terms of the language used but also through what features of a situation are being attended to and the mental/informational states of both. Pickering and Garrod (2007) in line with a spreading activation account suggest that comprehension of certain syntactic construction covertly activates production systems in readiness for generated an aligned response. (ii) There is also evidence that being primed through conversation promotes the ease of production fluency of that form. Processing effort is seen to decrease making subsequent speech more effortless,

Smith & Wheeldon (2001). (iii) Finally, priming is deemed to be a product of, and propagator for implicit learning. Information about syntactic/semantic mappings is repeated and stored for future use.

With imitation in a linguistic context, however, it can be difficult to disentangle whether children understand the item or feature they are imitating, or whether simple mimicry is taking place. Young children have indeed been shown to imitate linguistic forms they do not understand (Fraser, Bellugi & Brown, 1963), and nonsense material (Shipley, Smith & Gleitman, 1969).

Grammar specifically can be positively influenced by the usage of words as yet not semantically grounded in the lexicon. Slobin (1968) noted that when engaged in imitation language learners are able to produce grammatical features that are not seen in their spontaneous utterances. This work and later work by Ruth Clark (1976) identify that imitation has “pervasive effects on speech” as the context in which the imitative behaviors take place serve as conceptual markers to infants about what sounds, words or phrases are semantically linked to the experience. With time and repetition, the abundant phonetic, syntactic and semantic information held within the imitated sequences can be digested and gradually incorporated into ones knowledge structure to be flexibly redeployed elsewhere. After initial imitation without understanding patterns are in time recognized and the initially “liberal” semantic contexts begin to be narrowed down and promote the formation of productive rules (Cazden, 1968). Another salient reason why imitation is a positive activity for a language learner to undertake is that familiarity of the content can foster more rapid and discriminatory recognition of similar forms in the speech stream of adults encountered in the future (Clark, 1976). By manipulating the perceptual sensitivity of

the child more opportunities for experiential and contextual learning will be made available in future, thanks to the initial imitation of a related form.

Additionally the action, (or perhaps strategy) of imitating some feature of language which was previously unfamiliar is itself a pervasive quality of human communication. The frequent use of phrases, idioms and clichés by competent language users, which are cued by contextual or normative linguistic features and “bypass normal sentence processing routes,” aid communication and serve their purpose even if not fully analyzed by all those who utilize them (Ervin-Tripp, 1973 as cited in Clark, 1976). This communication alone is a desirable social outcome.

#### **6.3.2.1.7. Why Explore the use of Passive Verb Competence at Age 4?**

Using novel lexical items allowed us to assess how willing and able young children are to comprehend and incorporate such items into their own lexicon. We first looked at how often children incorporated these newly introduced lexical items into their own utterances. In addition we probed how flexible children could be in terms of switching between constructions in response to discourse pressure.

Contrasting children’s willingness to produce two novel items, presented in two constructions, one familiar and one unfamiliar, will further allow us to assess their linguistic representations of syntax. At age four children’s spontaneous transitive constructions are overwhelmingly active, with passive, particularly full passive utterances being scarce. However, there is evidence that children can produce novel passive utterances through structural priming (Brooks & Tomasello, 1999a). One paper specifically has reported that truncated passives can be elicited from the majority of 3-year olds using discourse pressure (Marchman, Bates, Burkhardt & Good, 1991). Furthermore children at this age have also been shown, albeit in small numbers, to flexibly produce a passive using a verb only previously heard in and



active construction (Brooks & Tomasello, 1999). Given this evidence we chose to streamline the training procedure implemented by (Brooks & Tomasello, 1999) modelling two novel verbs for the participants, one in an active and one in a passive voice. We applied discourse pressure in the production phase in an attempt to elicit the production of these novel verbs in both active and passive constructions. Such flexible use would further point to imitation as opposed to mimicry. Furthermore we would assess the comprehension of these modelled verbs to contrast whether the construction in which a verb was learnt effected its' likelihood of being understood. Children were asked to show the experimenter an example of the action in the active voice, regardless of the construction initially used. Previous studies have shown passively introduced verbs to be more poorly comprehended (Lovell & Dixon, 1967; Turner & Rometveit, 1967). We could also probe how successfully and flexibly these same participants could comprehend the novel verbs being presented in the two contrasting transitive frames.

To our knowledge this is the first study to experimentally combine both the comprehension and production of novel transitive verbs in the same set of participants. This design was chosen to elucidate whether children who used the novel verbs or unfamiliar syntactic structure were doing so in a truly imitative way vs., simple mimicry. Children who produced the target utterances but failed the comprehension questions could be identified as having mimicked the utterances modeled while not understanding them. Conversely participants who both produced and comprehended the items could be labeled as imitating the model, that is replicated the communicative goal and the means by which it was realized. Our second question then was centered round whether our sample appeared to be imitating the use of novel verbs or mimicking form only. For both production and comprehension elements of

this study we expected to see a range of individual differences in the following aspects of performance: how often a participant used a novel verb; how often they produced a passive utterance; how sensitive they were to discourse pressure and; how well they comprehended the verbs.

How frequently participants use a novel word could be mediated by several factors. Children love to imitate intention actions so may simply mimic the use a novel lexical item, without fully understanding its semantic or syntactic features but wishing to achieve the same goals. Children may use the verb and do so flexibly, indicating the item has been incorporated into the lexicon and become subject to the grammatical rules associated with other items of this class. Children may have taken these lexical addition steps but be unable or unwilling to translate this new knowledge into novel productive utterances. The newly formed representation of the verb *mee* for example may not be strong enough to support productive utterances in anything other than the construction in which it has been modelled. Or the representation may be strong enough but children have not picked up upon the cues inherent in the discourse of the experimenter, they may not correctly identify that discourse pressure is present.

When compared to Experiment 1, Experiment 2 affords children many more options as to how they choose to respond to the elicited production questions presented to them. How children organize the hierarchical goals of the situation consequently many differ along several dimension. Children could choose to copy all the features that the model produces, novel word, syntax and beyond. Some children might reject imitation in favour of answering in a way they feel most competent, i.e., with familiar English verbs and constructions. In such a task there are less clearly defined intentional features when looking at the literature. With object imitation the

salience of contrasting goals (such as in the mouse house task) has been explored. Spatial goals are replicated first, then process goals, then intentional goals etc. In relation to the features of language there are no such defined orders of importance. One might intuitively suggest that at the top of the hierarchy is replicating the ultimate goal of effective communication, replicating lexical items may come after this, followed by the replication of syntax. However, such goal directed distinctions have, as yet, not been discussed in the context of language development. As such predictions about what, and why, children imitate from the speech of others' in this experimental paradigm are difficult to make.

Broadly speaking, however, we expected many children to imitate the novel verb due to its salience as a shiny and new lexical item that has been repeated many times by the experimenter. In addition we expected that following passive models there would be an increased number of passive utterances produced by the participants as seen in previous experimental studies (e.g., Bencini & Valian, 2008; Bock & Griffin, 2000; Branigan, Pickering, McClean & Cleland, 2007; Brooks & Tomasello, 1999; Huttenlocher, Vasilyeva & Shimpi, 2004; Kashack & Glenberg, 2004; McClure, Pine & Lieven, 2006; Kidd, 2012; Messenger, Branigan & McClean, 2012; Pinker, LeBeaux & Frost, 1967; Savage, Lieven, Theakston & Tomasello, 2003; 2006; Thatcher, Branigan, McClean & Sorace, 2007; Vasilyeva, Huttenlocher & Waterfall, 2006; Whitehurst & Novak, 1973). Further we expected comprehension of the passively introduced verbs to be poorer than those actively introduced.

If children showed high levels of novel verb and syntax use coupled with poor comprehension, however, they could be identified as engaging in mimicry, the replication of surface level features of a behavior without understanding the means and the goals of the model. If children showed imitation and flexibility with novel

verbs combined with good comprehension, however, it would be clear that they not only replicated the linguistic behavior of the model but further understood its' function.

### **6.3.2.3. Method**

#### **6.3.2.3.1. Participants**

Data was collected from the same participants presented in section 2.3.1.

#### **6.3.2.3.2. Materials**

Replicating Brooks and Tomasello (1999), two novel verbs, *meek* and *tam*, were chosen to refer to two novel transitive actions the children would see being modeled by an experimenter. Each action had an agent and a patient. In total 36 objects were used in these roles. 18 were animate e.g., stuffed animals or dolls; and 18 were inanimate common household objects or toys e.g., a toothbrush or a plastic car.

Two pieces of apparatus were designed and constructed to enact our two transitive verbs (see Appendix G), one for each verb. Apparatus 1 was a rectangular trolley and track (60cm x 20cm) with a small ramp in the middle and a thin, vertical piece of tubing at one end (40cm). A wooden ball sat atop this tubing. The ball attached to chord that ran down the tube and affixed to a small wooden trolley with wheels (15cm x 20cm). When the ball was pulled upwards vertically the trolley would move horizontally along the track and across the ramp. Apparatus 2 consisted of a rectangular wooden frame (30cm wide and 60cm high) attached to a rectangular wooden base for stability. Suspended from the center of the frame by two pieces of thin rope affixed to either side was a transparent cylindrical container 10 cm in diameter and 15cm tall. An upturned plastic cup formed a pedestal placed 10-15cm from the base of the frame. When swung gently the container knocked off an object

placed upon the pedestal. For both Apparatus 1 and 2 both an agent and a patient are required to model the transitive action (meeking/taming). In Apparatus 1 the patient is placed on the trolley whilst the agent “pulls” the wooden ball thereby moving it along the tracks and over the ramp. In Apparatus 2 the patient is placed on the pedestal whilst the agent inside the plastic container swings toward it thereby knocking it over.

#### **6.3.2.3.3. Procedure**

Parents and children were invited to the on-campus lab for their session. Visits began with a short warm up time before the experimental tasks were introduced in a separate room. Tasks were administered in a purpose built testing room with inbuilt digital audio and visual recording capabilities. All actions were modeled using an apparatus placed in the centre of the black testing mat between the experimenter and participant. The position of the apparatus remained constant throughout the trials. First training, production and comprehension tasks were completed for one novel verb using one apparatus. Once these three phases were concluded, training, production and comprehension elements for the second verb and second apparatus were undertaken. Following are details for these three phases each of which children went through. There were four possible combinations of novel verb and voice in this within subjects design with children seeing both verbs and both verb constructions once. 1) Meek 1<sup>st</sup> passive/Tam 2<sup>nd</sup> active. 2) Meek 1<sup>st</sup> active/ Tam 2<sup>nd</sup> passive. 3) Tam 1<sup>st</sup> passive/Meek 2<sup>nd</sup> active. 4) Tam 1<sup>st</sup> active/Meek 2<sup>nd</sup> passive.

**Training phase.** Prior to seeing the apparatus children were told by the experimenter; “I am going to show you something, something I am pretty sure you will have never seen before.” The first apparatus and supporting objects (agents and patients) were then removed from behind a neutrally shaded curtain that skirted the testing room. The participants were shown the 18 items to be used with the apparatus

and were asked to identify them. Errors were immediately corrected by the experimenter. The experimenter then randomly selected one animate and one inanimate object and proceeded to model the novel verb eight times. Two types of modeling procedure were used, active or passive. Each verb was only modeled in one construction condition. Which condition and verb combination came first was counterbalanced across participants. In the active condition the novel verb was only used in the active voice with the focus on the agent of the transitive action, e.g., “Look! The toothbrush is going to meek/tam something.” In the passive condition the focus of the sentence was reversed to highlight the patient of the transitive action, e.g., “Look! The toothbrush is going to get meeked/tammed.” Three of the eight constructions were truncated: the remaining five were full. See Appendix G for the full script. Throughout the process the participant was asked questions about the patient or agent to engage their attention during the procedure. After the initial eight models children were asked to say the novel verb. Two new items were chosen for agent and patient, reversing whether the animate or inanimate was the agent/patient. The verb was then modeled a further eight times in the same voice and using the same frames as before. Again the participant was asked to say the verb following this phase.

**Elicited production phase.** Three forms of production question were utilized designed to exert differing discourse pressures on the participants; neutral, agent focused and, patient focused (adapted from Ahktar & Tomasello, 1997). These production questions were asked, one at a time by the experimenter following each of the three new combinations of agent/patient being used to carry out the action that had been previously modeled in the training phase. These questions were, neutral, agent and patient focused. The neutral question did not highlight either agent or patient of

the transitive action e.g., “What just happened?” The agent-focused question asked specifically about the action’s agent e.g., “What did the Cellotape just do?” Finally the patient focused question asked about the patient of the action e.g., “What just happened to the duck?” If children failed to respond or were non-compliant the action was repeated once and the question repeated. All participants were asked all three questions once: Neutral, then agent and finally patient focused. Each time they witnessed a different agent and patient pair. This was counterbalanced with alternating animate/inanimate items.

**Comprehension phase.** This directly followed the elicited production phase for the same novel verb. Eight objects, four inanimate and four animate, were ensured to be within reach and view of the participants. The experimenter indicated that it was the child’s turn to play with the toys. Two forms of comprehension question were asked: “1. Can you show me the cow meeking/tamming the train?” or, “2. This time, can you make the cow tam/meek the train?” Question 1 was asked twice, and then question two was asked twice. Each time the question involved one animate and one inanimate object. Again an alternating system was used so that the agent’s animacy was different from one question to the next. If the child failed to respond or was non-compliant the request was repeated once. Following completion of the first comprehension phase the second novel verb is trained, then its’ production is elicited and comprehension assessed.

#### **6.3.2.3.4. Data Analysis**

**Verb production.** All utterances were coded to identify the effects of training condition and discourse pressure on production. 20% of videos were secondary coded and inter-rater reliability was at 100%. Multiword utterances were categorized on several dimensions. Was the utterance grammatically correct or incorrect, active or

passive, transitive or intransitive, full or truncated and, was a novel or familiar verb used. If the child altered or repaired their construction in an unsolicited manner the revised version was retained. On N=4 occasions children failed to respond or stated they did not know the answer. These were coded as incorrect.

**Verb comprehension.** Responses were coded only as correct or incorrect. A correct response was defined as the placing of an agent and patient in their correct respective positions on the apparatus and subsequently carrying out the transitive action requested to completion. If a child initially placed the protagonists in the incorrect position and then corrected their mistake independently, before carrying out the action, this was coded as correct. Any other responses and non-compliance were coded as incorrect.

#### **6.3.2.4 Results**

##### **6.3.2.4.1. Novel Verb Utterances**

We first examined the utterances children produced looking at the number of English versus novel utterances. We were interested in assessing which verb type was being favored in the productive utterances in our sample (Figure. 6.4.). Furthermore we were interested in the transitive structure of the utterances participants' produced utilizing the novel verb, active, passive or incomplete (see Figure. 6.4.)



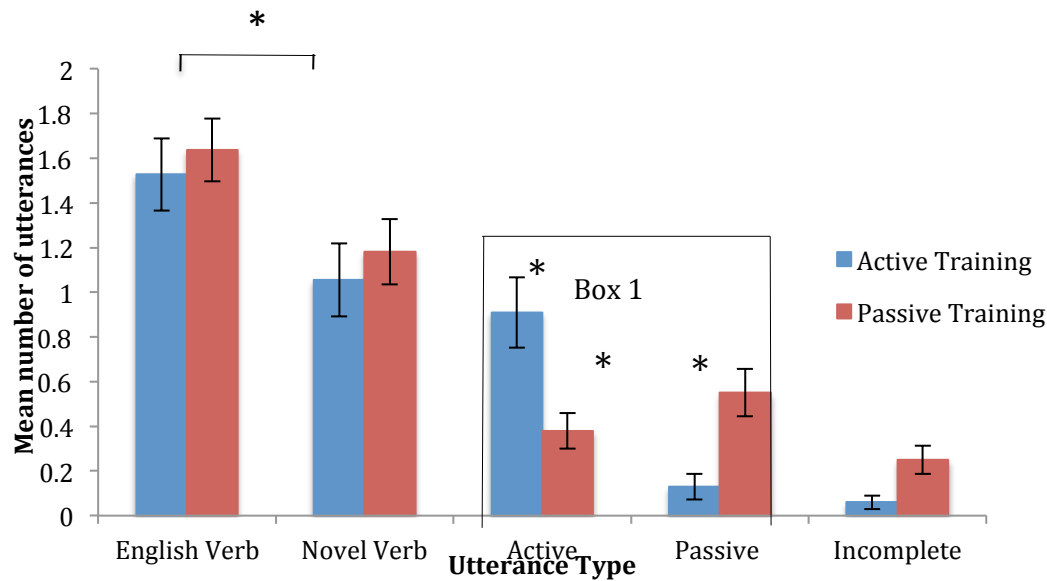


Figure 6.4. Overview table showing the number of English and novel verb utterances in the elicited phase and the transitive structure of those novel verbs produced. Error bars represent standard error. \* Represents a significant difference between bars  $p < .05$ .

Means were calculated from the total number of utterances children were prompted to produce and hence could be between 0-6. The mean number of English verbs produced=3.14, where the novel verb mean =2.1607. This difference was statistically significant with more English than novel verb utterances,  $t(55)=2.39$ ,  $p=.020$ . Novel verb here refers to the amalgamated scores of all utterances produced using the novel verb regardless of the linguistic construction in which they were used.

Box 1 highlights the constructions of interest, active and passive, that were produced using the novel verbs. There were significantly more active than passive novel utterances,  $t(55)=-2.64$ ,  $p=.011$  overall. There was a significant main effect of training condition on the number of active utterances, significantly more were produced using verbs actively trained,  $F(1,109)=7.7$ ,  $p=.007$ . Similarly there were significantly more passive utterances using passively trained verbs than actively

trained,  $F(1,109)=-12.96$ ,  $p<.001$ .

Additionally there was a main effect of training condition on the number of incomplete novel utterances produced overall,  $F(1,108)=6.65$ ,  $p=.011$ , with more incomplete utterances following passive training.

Table 6.2. *Mean number of active and passive utterances using novel verb following active or passive training.*

|                  | Number  | Mean | Standard Dev. |
|------------------|---------|------|---------------|
| Active Training  | Active  | .87  | 1.17          |
|                  | Passive | .17  | .43           |
|                  | Total   | .63  | .96           |
| Passive Training | Active  | .38  | .59           |
|                  | Passive | .56  | .79           |
|                  | Total   | .35  | .67           |

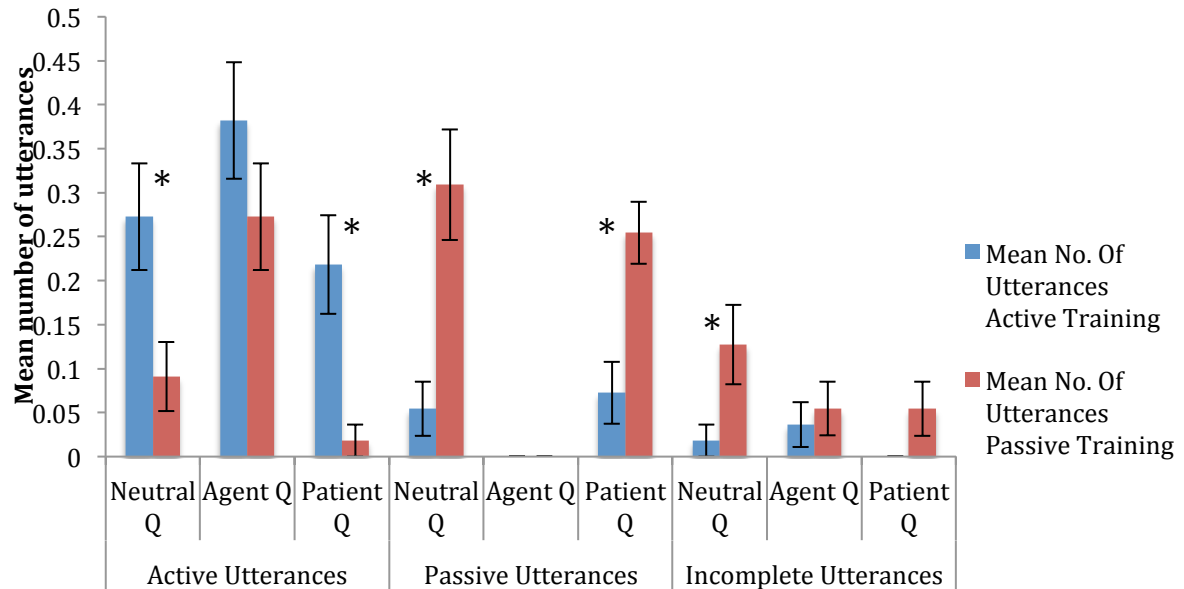


Figure 6.5.. Overview column chart showing the mean number of utterance types produced in response to the three elicited production question types and following both active and passive verb training. Error bars represent standard error. \* Represents a significant difference between bars  $p < .05$ .

**Active Utterances.** For active utterances there was a significant main effect of training condition on; the number of actives produced in response to the neutral question,  $F(1, 109)=6.35$ ,  $p=.013$  and the patient focused question,  $F(1,109)=11.46$ ,  $p=.001$ . There was no difference between conditions for the agent-focused questions.

There was a significant main effect of question type on the number of active utterances  $F(1,218)=14.69$ ,  $p < .001$ . Within the active utterances there were significantly more active utterances in response to agent focused than neutral focused questions,  $t(109)=-3.14$ ,  $p=.002$ , significantly more in the neutral than the patient focused questions,  $t(109)=-2.15$ ,  $p=.034$ , and significantly more in the agent than the patient focused,  $t(109)=5.1$ ,  $p < .001$ .

For the actively trained cases there was a significant difference between the agent and patient focussed questions in the form of fewer active utterances for the latter,  $t(54)=3.25$ ,  $p=.002$ .

For the passively trained condition there were significantly more active utterances in response to agent focussed than neutrally focussed questions,  $t(54)=-2.63$ ,  $p=.011$ , there were significantly more actives also in response to neutral focused questions than patient,  $t(54)=-2.058$ ,  $p=.044$ . Finally there were significantly more active utterances in response to agent than patient focused questions,  $t(54)=3.93$ ,  $p<.001$ .

**Passive Utterances.** For passive utterances there was a significant main effect of training condition on; the number of passives produced in response to neutral questions,  $F(1,108)=13.2$ ,  $p<.001$  and, the number of passives produced in response to patient focused questions,  $F(1,108)=6.94$ ,  $p=.01$ .

There was a significant main effect of question type on passive production,  $F(1,217)=21.33$ ,  $p<.001$ . Within the passive utterances there were significantly more passive utterances in response to patient focussed questions than agent focussed,  $t(109)=-4.62$ ,  $p<.001$ . There were also significantly more passives produced in response to neutral than agent focused questions,  $t(109)=4.92$ ,  $p<.001$ .

For the actively trained cases there was significantly more passives in response to patient than agent focused questions only,  $t(54)=-2.06$ ,  $p=.044$ . For the passively trained condition there were significantly more passive utterances in response to patient focussed than agent focussed questions,  $t(54)=-4.29$ ,  $p<.001$  and, significantly more passives in response to neutral than agent focused questions,  $t(54)=4.92$ ,  $p<.001$ .

**Incomplete Utterances.** For passive utterances there was a significant main effect of training condition on the number of incomplete utterances produced in response to the neutral focused question,  $F(1,108)=4.99$ ,  $p=.028$ .

Within the incomplete utterances there were no significant difference between the number of incomplete English utterances produced in response to the three questions types. Within both the actively and passively trained conditions there were no between question differences reaching significance.

**Question Type.** As reported earlier there is a significant main effect of question type on both the number of actives and passive produced overall. There is no main effect of question type when you look within the active training group only. The passively trained verb shows more variation with questions type. If you run the analysis with only the agent and patient questions and not the neutral, however, there are strong main effects on passive utterances, a significant interaction between question type and training condition training  $F(3,219)=6.94$ ,  $p=.009$ . There is also a significant main effect for actives  $F(2,164)=7.4$ ,  $p=.001$  and passives  $F(2,164)=12.4$ ,  $p<.001$  (see Figure 6.6.).

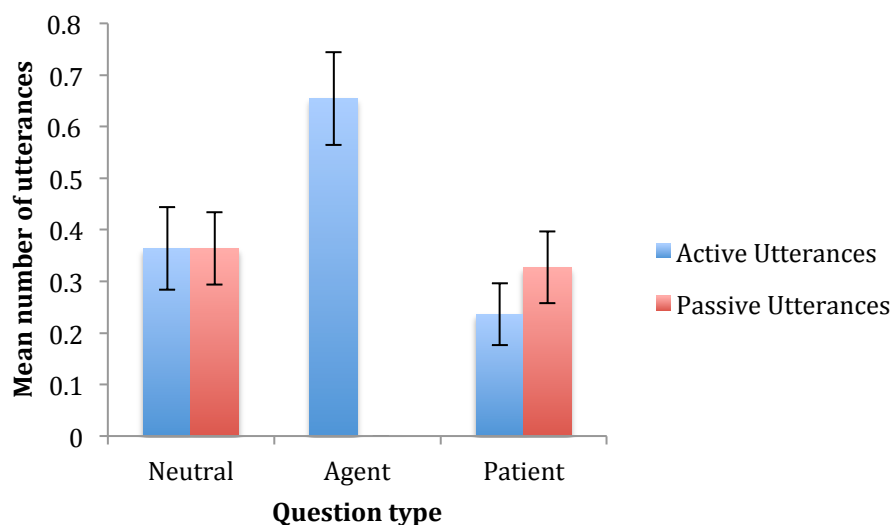


Figure 6.6. Mean number of passive and active utterances produced in response to the three question types. Error bars represent standard error.

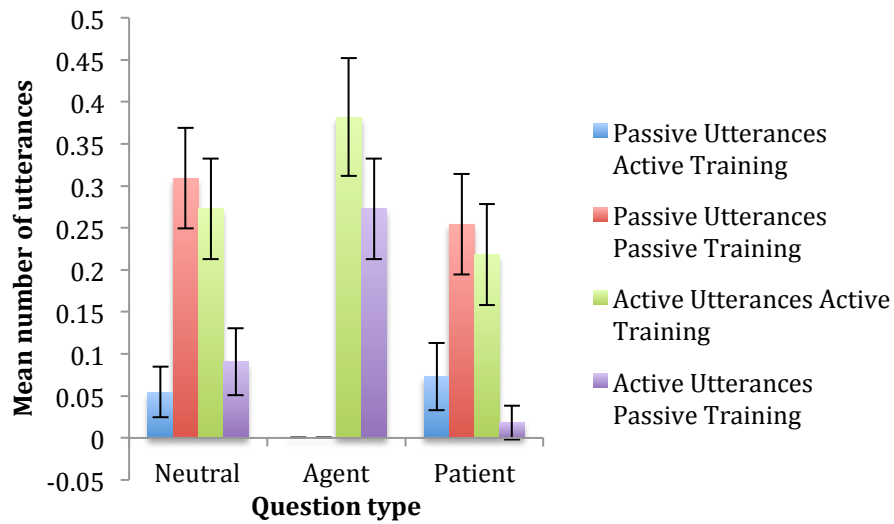


Figure 6.7. Column chart showing the number of English utterances, either active or passive, that were produced in response to the three question types. Error bars represent standard error.

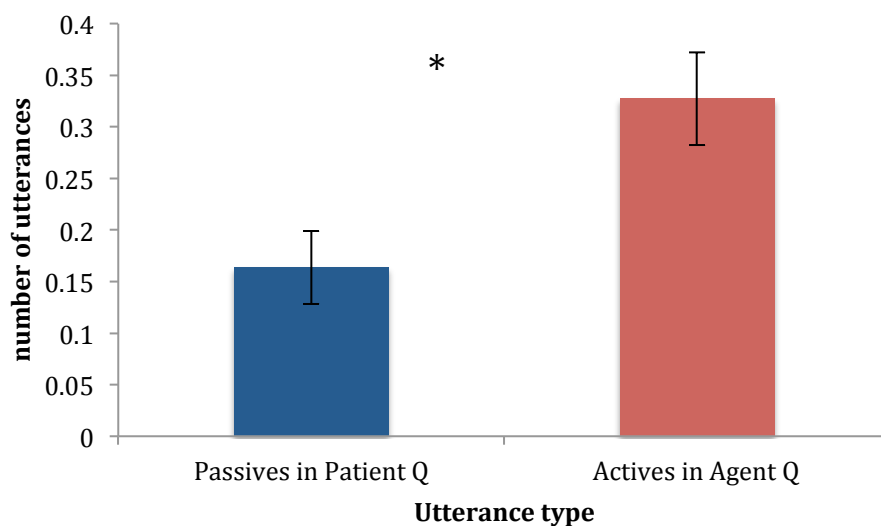


Figure 6.8. Column chart showing the relative number of novel passive utterances in response to patient focused questions compared to novel active utterances in response

to agent focused questions. Error bars represent standard error. \* Represents a significant difference between bars  $p < .05$ .

There were significantly fewer passives produced in response to patient focused questions than active utterances produced in response to agent focused questions,  $t(109) = -3.58$ ,  $p = .001$  (Figure 6.8.).

#### 6.3.2.4.2. English Utterances

Given that children preferentially chose to use an English verb rather than the modeled novel verb we also wanted to examine the transitive structure of their subsequent utterances. Our task was not only interested in novel verb production but also syntactic structure, hence in the following section we focus on the transitive structure of those utterances that contained familiar English verbs.

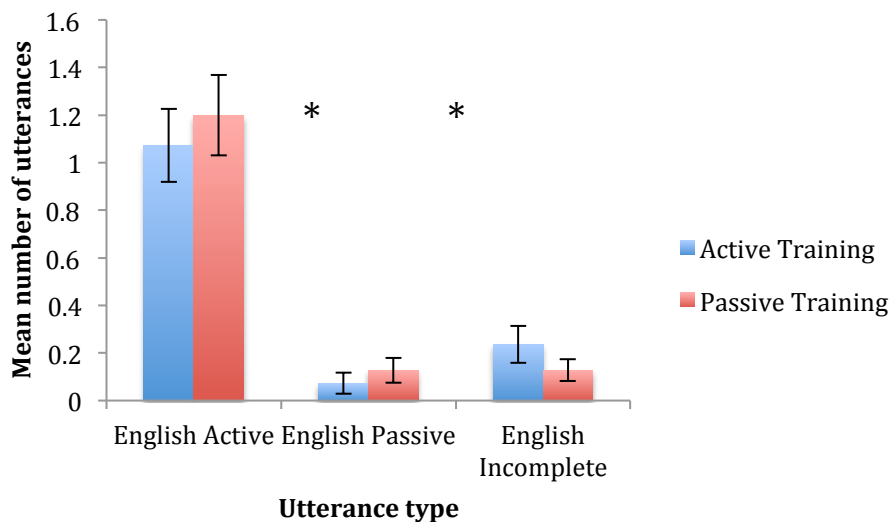


Figure 6.9. Column chart showing the number of active, passive and incomplete utterances using an English verb following both active and passive training. Error bars represent standard error. \* Represents a significant difference between bars  $p < .05$ .

There were significantly fewer passive than active utterances,  $t(54)=-7.52$ ,  $p<.001$  (Figure 6.9.). There were also significantly fewer incomplete utterances than active one,  $t(54)=7.21$ ,  $p<.001$ . There was no significant difference between the number of passive and incomplete utterances and no between training condition differences.

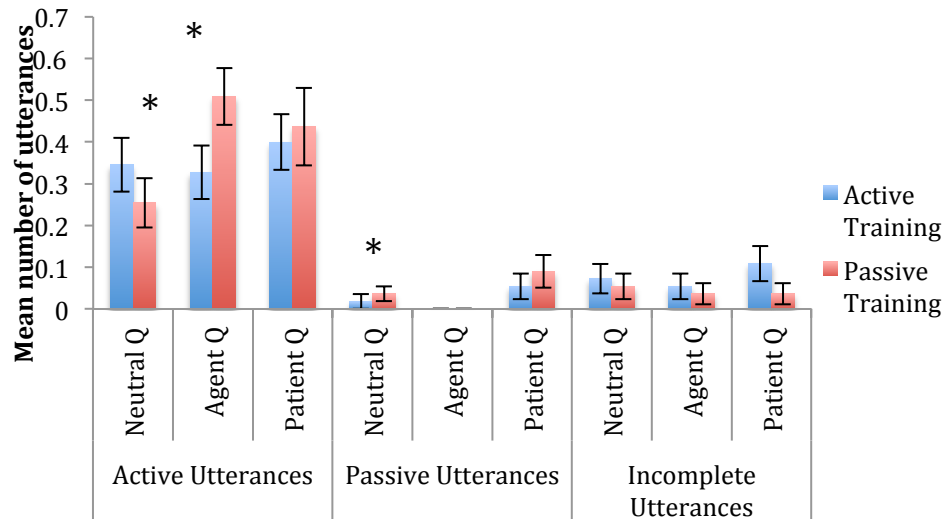


Figure 6.10. Overview column chart showing the mean number of English verb utterance types produced in response to the three elicited production question types and following both active and passive verb training. Error bars represent standard error. \* Represents a significant difference between bars  $p<.05$ .

**Active utterances.** For active utterances there was no significant main effect of training condition on the number of English actives produced overall,  $F(1,109)=.31$ ,  $p=.58$ , or in response to the different questions types.

Within the active utterances there were significantly more active utterances in response to agent focused than neutral focused questions,  $t(109)=-2.24$ ,  $p=.027$ , and significantly more in the patient than the neutral focused questions,  $t(109)=-2.17$ ,  $p=.032$ .



For the actively trained cases there were no significant differences in the number of active utterances produced for the different question types. For the passively trained condition there were significantly more active utterances in response to agent focussed than neutrally focussed questions,  $t(54)=-3.42$ ,  $p=.001$ . There were significantly more actives also in response to patient focused questions than neutral,  $t(54)=-2.21$ ,  $p=.032$ .

**Passive utterances.** For passive utterances there was no significant main effect of training condition on the number of English passives produced overall,  $F(1,108)=.64$ ,  $p=.34$ , or in response to the different questions types.

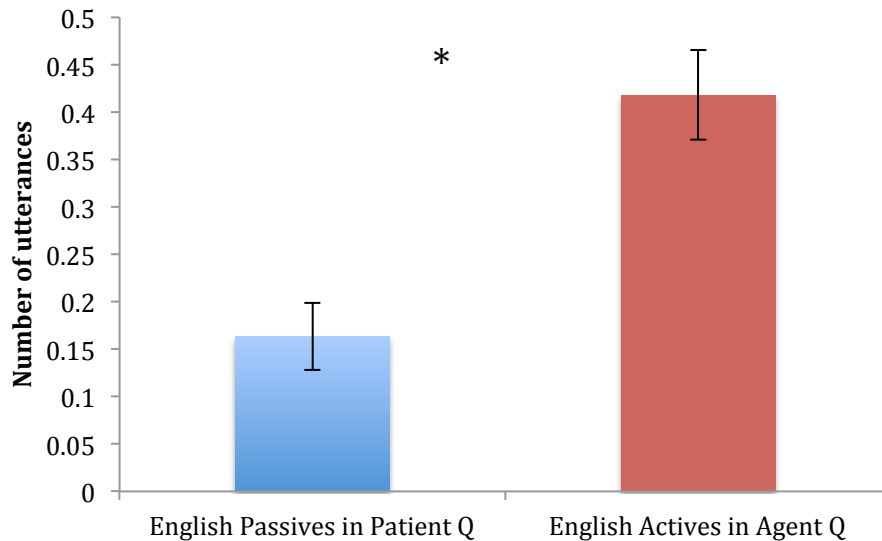
Within the passive utterances there were only significantly more passive utterances in response to patient focussed questions than agent focussed,  $t(109)=-2.9$ ,  $p=.004$ .

For the actively trained cases there were no significant differences in the number of passive utterances produced for the different question types. For the passively trained condition there were significantly more passive utterances in response to patient focussed than agent focussed questions,  $t(54)=-2.32$ ,  $p=.024$ .

**Incomplete utterances.** For incomplete utterances there was no significant main effect of training condition on the number of English incomplete utterances produced overall,  $F(1,108)=1.47$ ,  $p=.23$ , or in response to the different questions types.

Within the incomplete utterances there were no significant difference between the number of incomplete English utterances produced in response to the three questions types. Similarly within each training condition there were no significant differences across the question types.

There were significantly fewer passives produced in response to patient focused questions than active utterances produced in response to agent focused questions,  $t(109)=-6.6$ ,  $p<.001$  (Figure 6.11.).



*Figure 6.11.* Relative number of English passive utterances in response to patient focused questions compared to English active utterances in response to agent focused questions. Error bars represent standard error. \* Represents a significant difference between bars  $p<.05$ .

There were significantly fewer passives produced in response to patient focused questions than active utterances produced in response to agent focused questions across both English and novel verb utterances,  $t(109)=9.36$ ,  $p<.001$  (Figure 6.12.).

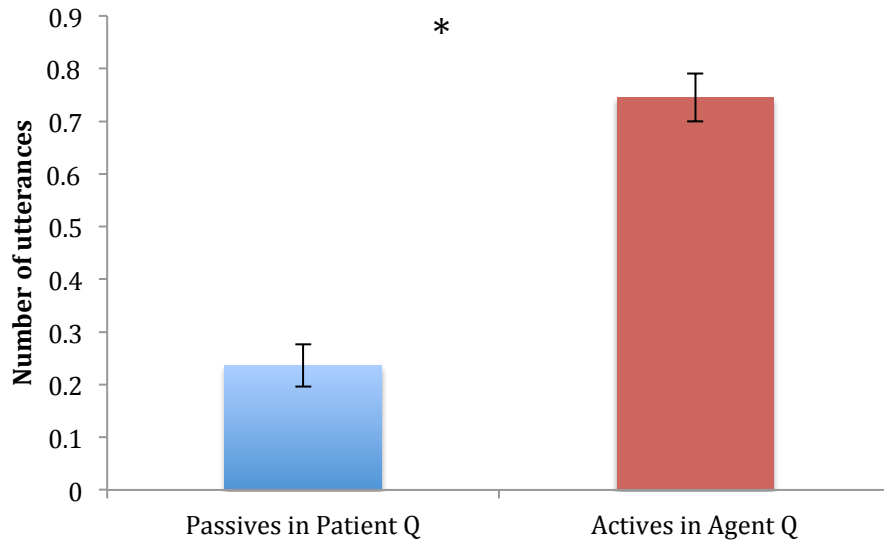


Figure. 6.12. Column chart showing the relative number of passive utterances in response to patient focused questions compared to active utterances in response to agent focused questions, using either English or a novel verbs. Error bars represent standard error. \* Represents a significant difference between bars  $p < .05$ .

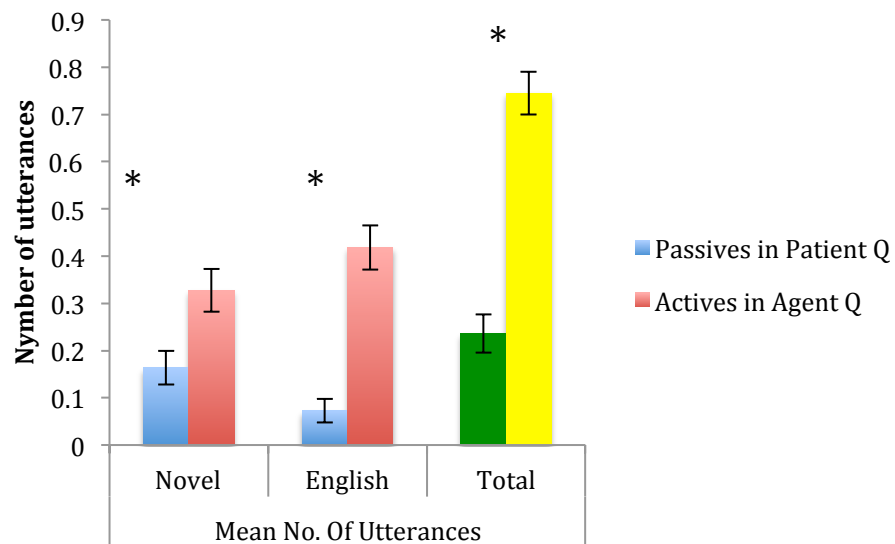
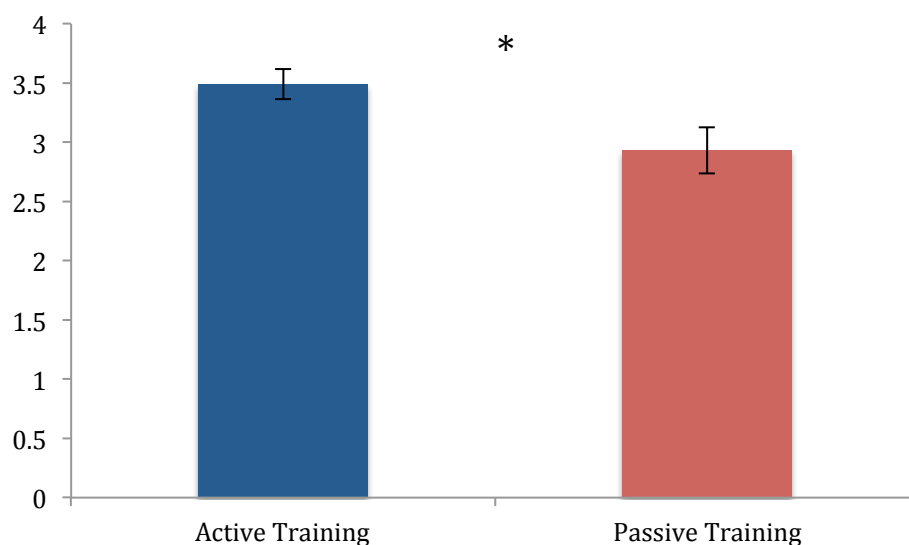


Figure 6.13. Composite of Figures; 8.5, 8.8 & 8.9.

### 6.3.2.4.3. Passive Verb Comprehension

The main effect for training type on comprehension scores was marginally significant,  $F(1,101)=3.79$ ,  $p=.054$ . When a pairwise comparison was carried out the mean comprehension score for actively trained verbs,  $M=3.5$  was significantly higher than for passively trained verbs,  $M=2.9$ ,  $t(50)=-2.71$ ,  $p=.009$  (Figure 6.14.).



*Figure 6.14.* Column chart showing the mean comprehension performance scores following active and passive training. Error bars represent standard error. \* Represents a significant difference between bars  $p<.05$ .

Comprehension was significantly better than production overall,  $t(51)=10.03$ ,  $p<.001$  (Figure 6.15.).

Children's passive, but not active comprehension scores related significantly and positively with their discourse sensitivity scores,  $r=.3$ ,  $p<.03$  versus  $r=.23$ ,  $p>.05$ . Children who were more responsive to the demands of the question were also more successful in comprehension of the more complex syntactic form.

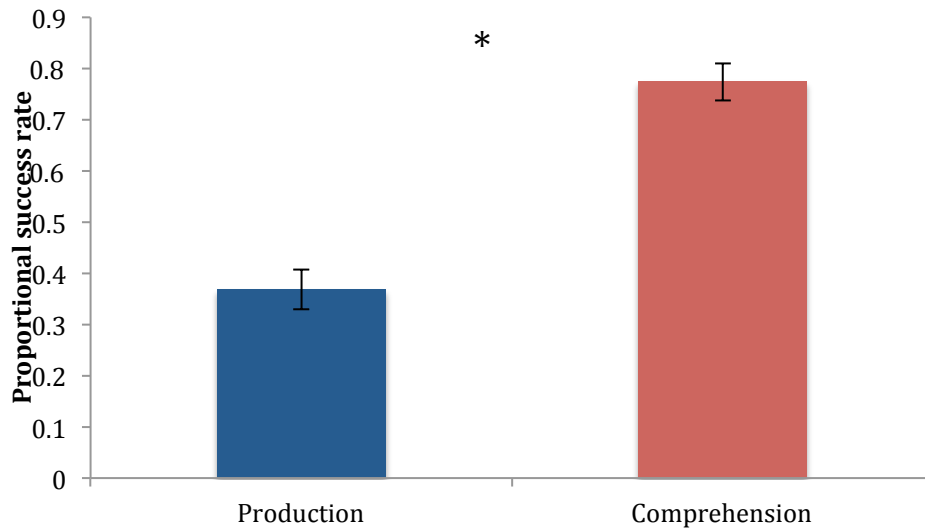


Figure 6.15. Proportional total scores for total verb comprehension and production.

Error bars represent standard error. \* Represents a significant difference between bars  $p < .05$ .

### 6.3.2.5. Experiment 2: Discussion

The passive verb construction was chosen for this experimental paradigm as children younger than age four to five years are not typically seen to produce it spontaneously (Harris & Flora, 1982; de Villiers & de Villiers, 1978). When given the opportunity to narrate transitive actions children consistently demonstrate a preference for the active construction (e.g., Kidd, 2012; Savage et al., 2006; Shimpi et al., 2007; Whitehurst et al., 1975). Comprehension of this construction is also much reduced when compared to the more common active construction (Lovell & Dixon, 1967; Turner & Rometveit, 1967a). There is disagreement in the literature as to how children develop their ability to produce and understand the passive construction. The use of syntactic priming studies in which children are given passive models prior to a free description opportunity has led some researchers to hypothesise that children hold abstract representations of the construction (e.g., Bencini & Valian, 2008;

Gentner et al., 2000; Shimpi et al., 2007; Thothathiri & Snedeker, 2008). The contrasting set of theories, item-based accounts instead deem children conservative verb users, instead preferring to use items only in the construction in which they have been previously heard (Tomasello, 2000). Using novel verbs, to ensure young children brought no prior experience to the experimental setting, Brooks and Tomasello (1999) reported that children were strikingly conservative with their verb use. They suggest that children build a verb-island and only with time and experience do verb-islands begin to connect, supported by semantic relationships, and syntactic rules be used flexibly with novel items.

We too chose to use novel items in our experimental paradigm. Although the majority of responses to our elicited production questions were answered in a semantically appropriate manner, overall children were more willing to substitute in familiar English verbs than produce utterances with the novel verbs. At the group level then children were unwilling to imitate a newly introduced novel verb, regardless of construction, into their own utterances. There were more English than Novel verbs generated in the passive training condition, however, these differences did not reach significance.

As reported in previous literature, there were significantly more active vs., passive utterances produced. In addition we observed the expected significant main effect of training condition on our sample's productive constructions using the novel verbs: there were significantly more novel active utterances produced following active training and; significantly more passive utterances produced following passive training. Additionally there were significantly more incomplete novel utterances, not adhering to a transitive structure following passive training. It is clear from these findings that at age four children are beginning to imitate the linguistic structures of a

social partner. In addition at first glance such results would seem to support a verb island hypothesis of early linguistic development as there is evidence of more actives following active models and vice versa (Tomasello, 1992). However, when looking back to Table. 7.1. or Figure. 7.1., children by no means exclusively retain the novel verb in the construction that they have heard it modelled: some children are being flexible in their production of newly learned verbs. Particularly in the case of the passively modelled verb a high proportion did revert to using the more familiar active construction with the newly presented verb, they were not construction specific. For children to have been flexible in this way they must have recognised these new lexical items as belonging to an abstract category pertaining to something like the adult classification of a “verb.” More over they would have had to import knowledge relating to the rules of this category and apply them to the novel item in order to switch between constructions in this way. Children have been able to take newly acquired knowledge from their experience of reality, integrate it with previous causal knowledge relating to the rules surrounding verb forms, and move into a new linguistic reality of their own making having generated a truly original utterance.

When looking at the effect of training condition and agent vs., patient elicited production question type there was a significant interaction. For active utterances there were significantly more produced in response to agent focused questions than patient focussed. Conversely then for passive utterances there were significantly more produced in response to patient than agent focussed questions. Our sample then showed evidence of discourse sensitivity with different question types eliciting different proportions of verb constructions in line with the attentional focus prompted. Particularly interesting was the observation that no child produced a passive utterance in response to a agent focussed question.

Also interesting, however, was the finding that following passive training, patient focussed questions yielded significantly fewer passive utterances than neutral questions. One explanation for this could be the unconstrained nature of the neutral question. No discourse pressure was being applied, congruent or otherwise, therefore no additional demands were being made on the participants to generate a particular linguistic state. In the case of the active construction, it being a more familiar form, both to comprehend and produce, may have worked to eliminate this effect.

Finally incomplete utterances did not vary as a function of question type, they were of a statistically comparable level for neutral, agent and patient focussed questions. Between training conditions, however, there were more incomplete utterances following passive training in both the neutral and patient focussed questions. There were significantly more novel active utterances overall than passive or incomplete utterances. Passive and incomplete utterances were at a statistically comparable level. In light of this it seems clear that passively trained novel verbs were difficult for children to produce transitively and more difficult to manipulate in response to discourse pressures. Despite having models of both active and passive constructions children were just as likely to produce passive utterances as produce intransitive, incomplete utterances. There was much more success with actively trained novel verb, highlighting that the less familiar construction, and not the novel verbs per se, was responsible for the productive difficulties. There were significantly fewer passive utterances produced in response to patient focussed questions than active utterances produced in response to agent focussed.

As seen in Figure. 6.4. there were a great deal more English active utterances than passive or incomplete. When children rejected the use of a novel lexical item they more often than not chose the most familiar type of transitive construction to fall



back on: the English active. There were no between condition differences in the numbers of English utterances of each type produced.

However, when looking only at the active English utterances, the most common type of English utterance, there were significantly more produced in response to the agent focussed question following passive training. This makes intuitive sense as children struggled to flexibly apply the necessary syntactic rules to the newly learnt passive verb in the face of discourse pressure. Instead then children reverted to a familiar English verb in the active form, thereby successfully adhering to discourse pressure and communicating successfully.

There were significantly more active English utterances in response to neutral questions following active training. Possible suggestions as to why this was the case were not so forthcoming. Perhaps some children were interpreting the neutral question as an opportunity to redescribe the scene, hence using different words.

There were very few passive English utterances, however, there were significantly more in response to patient focussed questions when compared to agent focused. This suggests that even when children abandon the use of a novel verb they are still attempting to conform to discourse pressure. Incomplete utterances were statistically comparable across training conditions and question types. Finally children were significantly poorer at comprehension questions relating the verbs they had encountered only in the passive construction when compared to actively trained verbs.

To summarise then (i) across training condition and question type children were more likely to revert to an English verb than incorporate a newly learnt lexical item. (ii) Children were more likely to revert to English verbs following passive training, however, highlighting a fundamental difficulty with this construction for young children. (iii) Children were more likely to use novel active, than novel passive

constructions in their own utterances. (iv) There was evidence of children being sensitive to discourse pressure with more passives produced following patient focussed questions and more actives produced following agent focussed questions. Finally, (v) children were significantly poorer at comprehending passively trained verbs than actively trained verbs.

**Imitation versus mimicry.** In our sample there was very little evidence to support a view that children were simply mimicking the language they heard modeled. Mimicry could be identified in situations where production scores exceeded comprehension scores. Firstly comprehension was significantly higher than production with the novel verbs: children were more able to comprehend than produce novel items in any form. Secondly the comprehension scores for those children who did not produce a novel verb were significantly lower than for those children who did. Looking only at children low on the passive verb comprehension measure though they showed significantly fewer novel transitive utterances and significantly more English utterances: they were less imitative overall. Children in this context were often fairly innovative in their linguistic responses, however, the children who were more lexically innovative (by substituting an English verb) may have been doing so to mask the fact they did not fully understand that the novel item related to the novel action. They had not been able to build a robust enough representation of the verb to support either comprehension or production.

In terms of construction faithfulness only ten participants on our sample stuck to the modeled construction only. Furthermore there were no differences in the comprehension scores of these groups. It does not seem then that children are producing novel transitive utterances in a mimicking way, without comprehension. Imitative production patterns alone are similar regardless of comprehension scores.

Children's imitation, particularly the flexibility of imitation does seem then to be related to comprehension. Language comprehension supports language production rather than the converse. Willingness to use a novel verb then seems to be indicative of advanced development, a verb seems more likely to be used when its semantic, causal properties are understood. Understanding of syntax facilitates the production of novel items.

To explore this further we inspected children's discourse sensitivity, our measure of how flexible participants were in response to questions exerting different agent foci. This further index of linguistic competence was also positively related to passive comprehension. This would suggest that flexibility within imitation is indeed related to comprehension. Critically then discourse sensitivity was very highly correlated with novel verb usage,  $r=.53$ ,  $p<.001$ , but negatively correlated with the number of English utterances,  $r=-.33$ ,  $p<.02$ . This lends more support to our theory that causal understanding about a verb's semantic properties supports, and the causal affordances of syntax more generally support language production.

It is clear that children who better understand language are more willing to imitate it.<sup>1</sup> It is not the case that children are simply imitating linguistic items in a construction specific way. This chimes with the verbal imitation task where grammatical sentences were more likely to be imitated, and with the action imitation task where imitation was greater when actions were understood as necessary.

Imitation is initially organized instrumentally, with goals being hierarchically defined

---

<sup>1</sup> In the following chapter we generate a language competence score combining aspects of children's productive and receptive language skill in addition to their flexible use of language assessed by the discourse sensitivity score mentioned. Verbatim imitation in both the grammatical and ungrammatical conditions of the first task introduced in this chapter correlated positively with language competence scores. This lends further support to our assertion that language imitation is not based on simple mimicry but that it is grounded in and supported by linguistic knowledge.

(Bekkering et al., 2000). Children often imitate the goal of an action before they imitate the means through which that goal was achieved. The primary goal of language is communication and the vast majority of participants responded in some form in the vast majority of cases, fulfilling this instrumental goal through whichever means they could. The attentional focus of the question can subtly modulate this communicative goal, if realized, and hence make one transitive form more or less appropriate than the other. In addition to this instrumental, communicative goal then the means of communication can be entered in the hierarchy. Then the exact surface features of the utterance may be imitated including the lexical items or syntactic constructions utilized.

Surface imitativity was clearly evidenced through: the majority of children using a novel verb transitively at least once, 75% and; children using a novel or English transitive utterance over an intransitive or incomplete one,  $t(54)=10.6$ ,  $p<.001$ . So having established that the majority of children replicated the communicative goal of using a transitive utterance to describe an action sequence we further showed that the majority of children imitated the use of a novel lexical item at least once. The next question to ask is there any contiguity in children's imitativeness across different linguistic and action tasks?

### **6.3.3.1. Analysis 3: Introduction**

In the first two sections of this chapter we presented data concerning children's imitation and innovation when given the opportunity to imitate the language of others around them. Children were avid imitators of both grammatical and ungrammatical sentences, they reverted to English verbs more often than they adopted novel verb forms and they showed evidence of imitating the syntactic structure of a model. The next step for us then was to look for patterns of imitation

across the linguistic tasks and finally relate children's imitativity in the language domain with their imitativity in the action domain as assessed in Chapter 5.

#### **6.3.3.1.1. Imitation of Actions and Speech**

As discussed in earlier sections (see 5.1 and 8.1.5) children's imitation is influenced by numerous structural and social factors and are often organized around a hierarchy of goals. Goals are extracted from the individual's representation of some event in the world and can be identified whether or not the model actually achieved the goal in some situations (Meltzoff, 1995). Furthermore in very young children Sakkalou et al. (2013) reported that there was stability in goal directed imitation with participants having correlated scores across imitative tasks where intentionality was indicated via diverse vocal and object means. Based on these two seminal findings we expected there to be a strong individual relationships in imitation within the verbal tasks and also across the action and language domains. We expected children who imitated more faithfully in one task to also imitate more faithfully in the other.

Indeed within and across many of the action and verbal imitation tasks reported in Chapters 5 and 6 children showed a striking tendency towards faithful imitation. Additionally there was evidence of selective imitation more often in one condition than the other across domains. In this final analytical section of Chapter 6 we will investigate the cross-domain relationships between action and verbal imitation.

#### **6.3.3.1.2. Aims**

In Analysis 3 we shall combine the experimental verbal imitation data collected in Experiments 1 and 2 of this chapter with the action imitation data reported in Chapter 5. We expected that imitation patterns in all three would be vary together given that imitation is reported to be driven by both the intentional behaviors

of the model and the necessity of these behaviors as was reported in Chapter 5 (eg., Bekkering et al., 2000; McGuigan et al., 2007). Furthermore we expected there to be strong cross-domain relationships with children's imitative proclivity being consistent in both action and verbal domains. Although this cross domain relationship from a productive stand-point has not, to our knowledge been explored as fully before a contiguity in object versus vocally cued imitative behavior has been reported (Sakkalou et al., 2013).

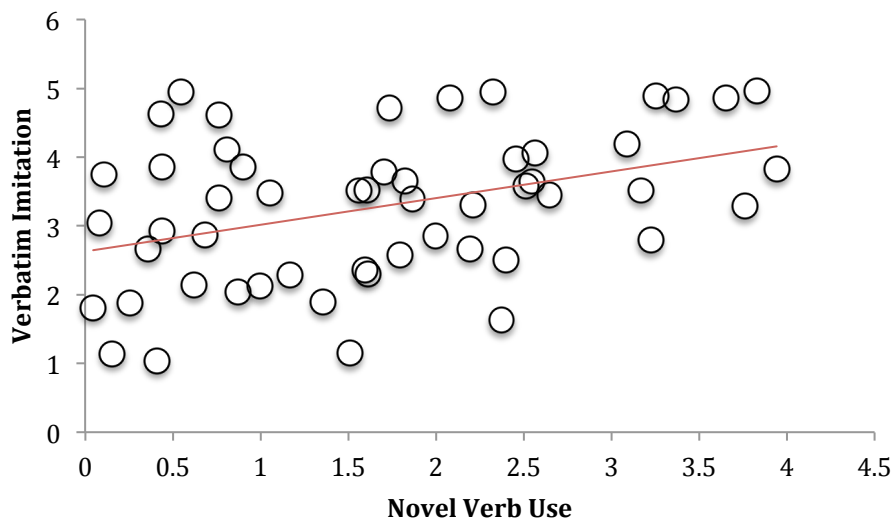
Our first aim for this final section was to investigate whether there are correlations between the imitative verbal behaviors related to our three features of language themselves: repetitions; novel verbs and; syntax. This would help us answer the question of whether verbal imitation patterns are consistent across different features of productive language before we set about extending the analysis to include object related behaviors. Our second aim then is to relate these language imitation indices to the imitation of actions reported in Chapter 5.

### **6.3.3.2. Results**

We first looked at imitation within the verbal domain following which we moved to include action imitation also. Results are separated into sections for clarity. We were looking specifically for contiguity between imitation patterns, between verbal tasks and across the verbal and actions tasks.

#### **6.3.3.2.1. Language Imitation: Between Task Relationships**

**Repetitions and novel verbs.** When the total number of novel verb utterances were compared to the total number of sentences repeated verbatim the positive correlation found was marginally significant  $r=.24$ ,  $p=.08$ . When only the target sentences containing a repetition were included this correlation increased in strength but remained marginal  $r=.25$ ,  $p=.069$  (see Figure 6.16.).



*Figure 6.16.* Scatterplot showing novel verb production scores following passive training on the x-axis and the total number of verbatim repetitions in the target conditions on the y-axis.

Following active training, novel use was not related to children's imitation of sentences containing repetitions  $r=.04$ ,  $p=.75$ , be they grammatical or ungrammatical. Following passive training, however, children's use of a novel verb was positively related to their verbatim imitation across the grammatical and ungrammatical sentences  $r=.35$ ,  $p<.01$ . This was driven mainly by the ungrammatical condition with the correlation here reaching  $r=.36$ ,  $p<.01$  (see Figure 6.17). This same pattern was marginally significant in the grammatical condition,  $r=.24$ ,  $p=.08$ .

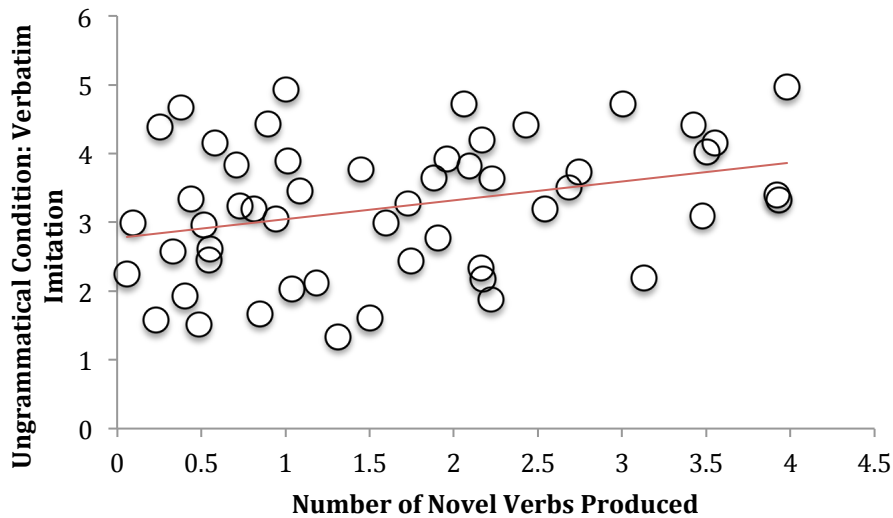


Figure 6.17. Scatterplot (including jitter) showing the number of novel verbs produced following passive training on the x-axis and the number of verbatim repetitions in the ungrammatical sentence condition on the y-axis.

**Repetitions and syntax.** Overall children who recounted sentences containing a repetition verbatim were more likely to produce a transitive utterance  $r=.33$ ,  $p<.05$ . This was driven by the grammatical condition where verbatim imitation correlated positively and significantly with the number of transitive utterances produced  $r=.31$ ,  $p<.05$  (see Figure 6.18.). Conversely the number of repetitions removed in the grammatical condition related negatively with the number of transitive utterances produced  $r=-.32$ ,  $p<.02$ . The number of novel passives but not actives or intransitives showed a marginally significant positive correlation with verbatim responses including a repetition  $r=.27$ ,  $p=.05$ . This relationship was driven by more novel passives being produced by children also retaining the repetitions in the ungrammatical condition  $r=.27$ ,  $p=.05$ . There was no such relationship seen with active or intransitive utterance.



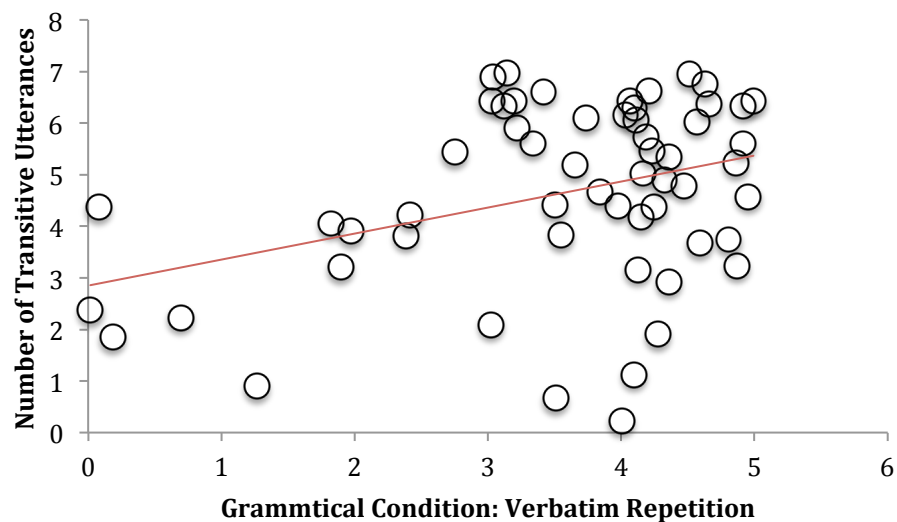


Figure 6.18. Scatterplot (including jitter) showing the number of transitive utterances produced on the y-axis and the number of verbatim repetitions in the across condition on the x-axis.

When looking at construction faithfulness, that is the number of passives after passives and actives after actives total construction faithfulness did not correlate with repetition imitation. However, when looking at passive construction faithfulness individually and repetition imitation in the ungrammatical condition a significant correlation was present  $r=.28$ ,  $p<.05$  (see Figure 6.19) but not in the grammatical condition  $r=.056$ ,  $p=.7$ . Children who, regardless of verb, chose a passive construction after hearing it modeled were also more likely to retain ungrammatical repetitions. Structure not semantics remained key.

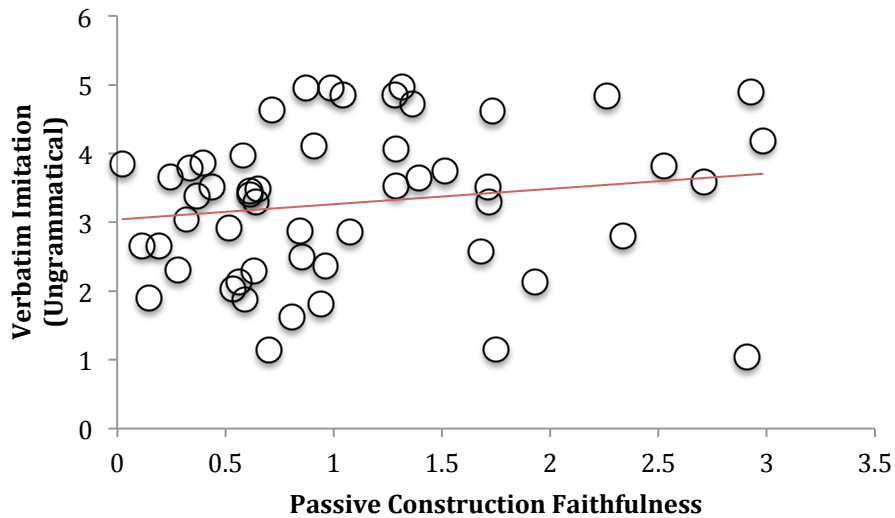


Figure 6.19. Scatterplot showing passive construction faithfulness on the x-axis and verbatim repetition in the ungrammatical condition on the y-axis.

### 6.3.3.2.2. Imitation of Language and Actions: Cross Domain Relationships

**Repetitions and actions.** Initially each imitation apparatus was inspected separately. Neither two-action, or Just B responses were correlated with performance in the repetition task. However, when looking at the temporal order of the responses the number of simultaneous responses in the necessary condition did correlate significantly negatively with verbatim imitation in the grammatical condition  $r = -.28$ ,  $p < .05$  (see Figure 6.20). The opposite pattern was found in the unnecessary condition but did not reach significance  $r = .2$ ,  $p = .09$ .

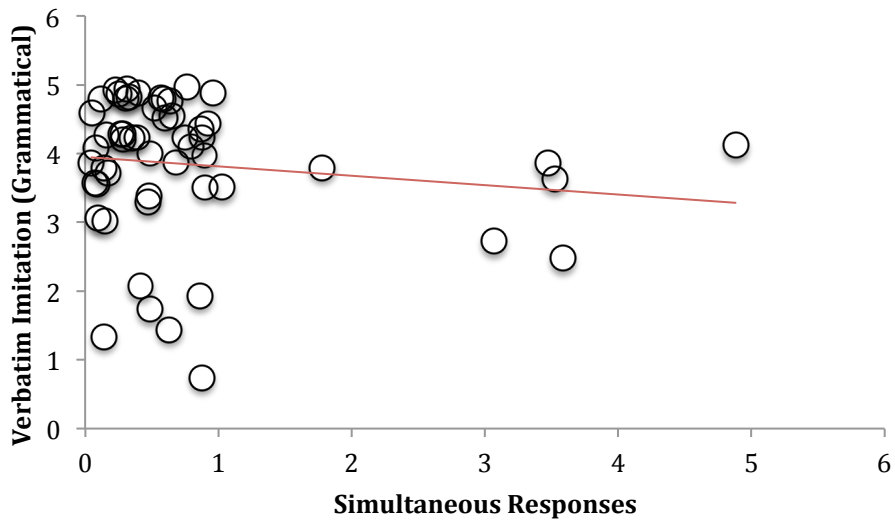


Figure 6.20. Scatterplot showing simultaneous responses in the necessary condition on the x-axis and verbatim repetition in the grammatical condition on the y-axis.

However, when looking at the specific temporal responses there were some significant associations. For the Tower, unnecessary condition there was a significant negative correlation between the number of B then A responses and verbatim imitation in the ungrammatical condition  $r=-.46, p<.02$ . Children who ignored the temporal sequence in the action imitation paradigm were less likely to imitate an ungrammatical repetition faithfully in a verbal imitation paradigm. Furthermore there was significant positive correlation between repetition removal in the grammatical condition and B then A responses in the unnecessary Tower condition  $r=.46, p<.03$ .

For the Rake apparatus then it was B then A responses in the necessary condition that correlated positively with additional repetitions in the grammatical condition,  $r=.46, p<.03$ . Children were ignoring the sequence of intentional actions in across both domains.

Across apparatus, however, the number of faithfully imitated actions carried out in the critical unnecessary condition did not correlate with the total number of sentences containing repetitions recounted verbatim  $r=.01$ ,  $p=.97$ .

**Imitation of novel verbs and actions.** Again each imitation apparatus was inspected separately. For the Tower there was a significant negative correlation between novel verb use and single action, Just B responses in the unnecessary condition,  $r=-.4$ ,  $p<.05$  (see Figure 6.21). Novel verb use correlated negatively with selective imitation. Furthermore the number of novel passives correlated negatively with the number of simultaneous responses carried out on the tower in the unnecessary condition,  $r=-.52$ ,  $p<.01$ . For the Rake there were no such significant correlations observed.

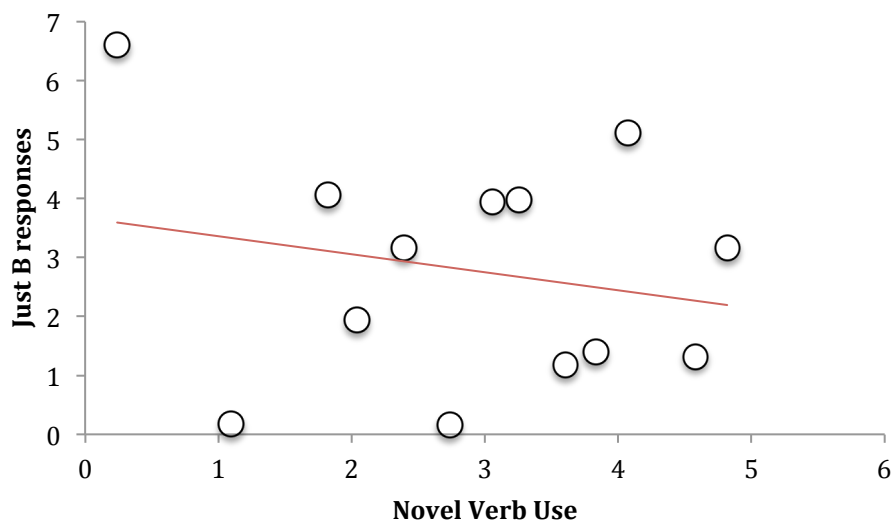


Figure 6.21. Scatterplot showing number of Just B responses in the Tower Unnecessary condition on the y-axis and total novel verb use on the x-axis.

**Imitation of syntax and actions.** With the Tower we inspected the number of novel and familiar transitive and intransitive utterances correlated with our action imitation measures. The number of Just B responses in the unnecessary condition and

the number of novel transitive utterances correlated negatively,  $r=-.42$ ,  $p<.04$  (see Figure 6.22).

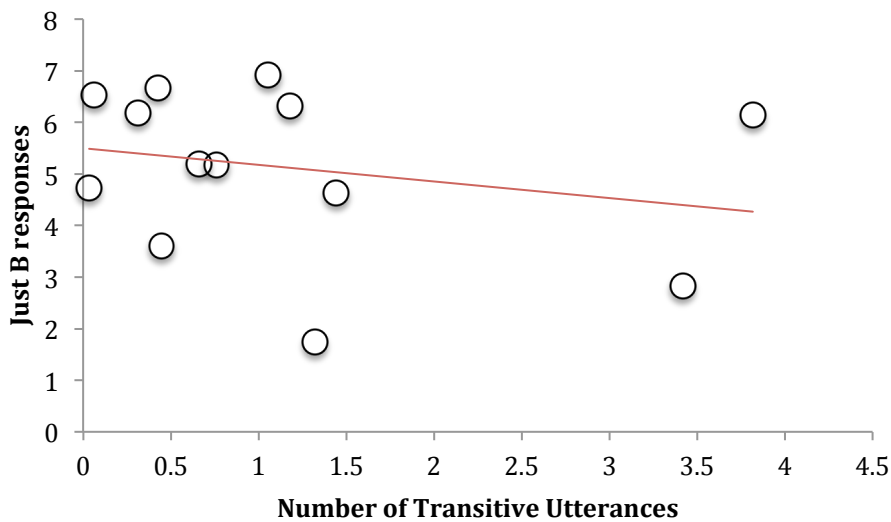


Figure 6.22. Scatterplot showing number of Just B responses in the Tower Unnecessary condition on the y-axis and total transitive production on the x-axis.

With the Rake there was a significant positive correlation between the number of faithful responses in the necessary condition and the number of transitive utterances,  $r=.41$ ,  $p<.05$  (see Figure 6.23). A then B utterances specifically were positively correlated in this way,  $r=.42$ ,  $p<.04$ . This was mirrored by a significant negative correlation between Just B responses and number of transitive utterances,  $r=-.41$ ,  $p<.05$ . There were no relationships seen between our action imitation indices in the unnecessary condition and novel transitive production.

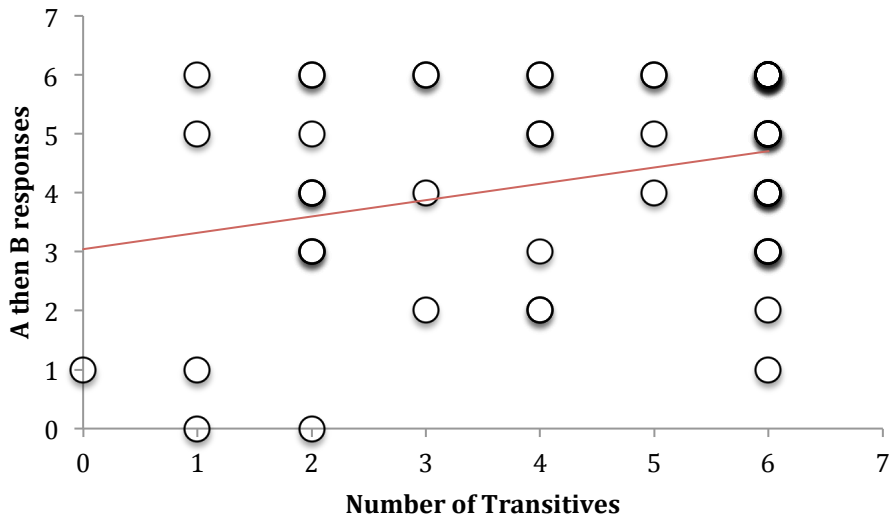


Figure 6.23. Scatterplot showing number of A then B responses to the Rake on the y-axis and total transitive production on the x-axis.

Across apparatus the total number of transitive utterances correlated positively with the total number of temporally faithful A then B responses, however, this relationship was marginal,  $r=.23$ ,  $p=.08$ . A then B responses in the necessary condition only correlated significantly and positively with the total number of transitive utterances produced  $r=.31$ ,  $p<.05$ . Total construction faithfulness and A then B responses in the necessary condition marginally correlated in the same direction  $r=.25$ ,  $p=.06$ .

The total number of transitive utterances, however, did correlate significantly negatively with the total number of Just B responses  $r=-.28$ ,  $p<.05$  (see Figure 6.24). Just B responses further correlated negatively with the number of active utterances after active models,  $r=-.35$ ,  $p<.01$ . Particularly in the unnecessary condition Just B responses correlated negatively with active construction faithfulness  $r=-.27$ ,  $p<.05$ , with the necessary condition only being marginally significant  $r=-.23$ ,  $p=.09$ . Total construction faithfulness correlated marginally negatively with these Just B responses

$r=-.26$ ,  $p=.051$ . B then A and Simultaneous responses showed no relationship with syntax imitation.

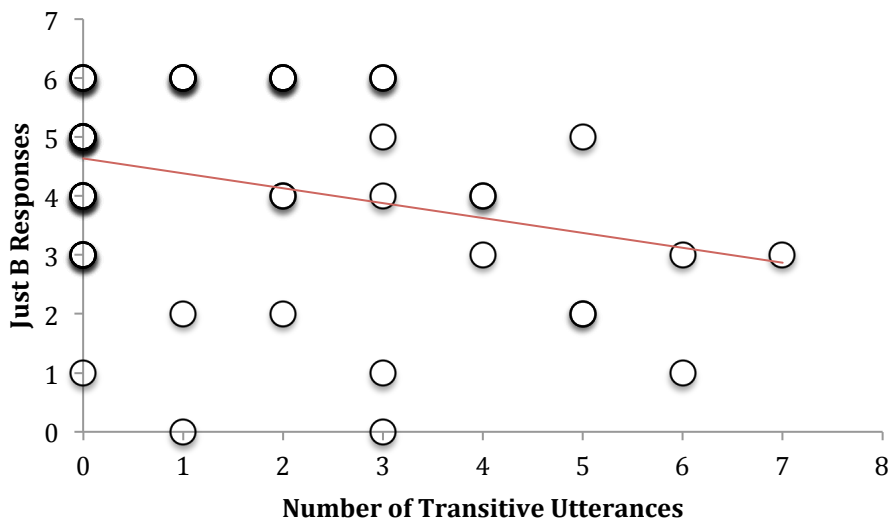


Figure 6.24. Scatterplot showing number of Just B responses on the y-axis and total transitive production on the x-axis.

### 6.3.3.3. Discussion

Children can be highly imitative at age four as evidenced in this data set and many more. In some cases they go above and beyond what is necessary to imitate. However, not all imitate-able behaviors are equal: children more often than not imitate a simple two-action sequence on an object; they more often than not imitate repetitions in short sentences; they more often than not do not imitate a novel verb; they imitate transitive versus intransitive structure at chance levels; they more often than not imitate active syntactic structure; they more often than not do not imitate passive syntactic structure. Some things children at this age readily imitate, some things they do not.

The first aim for our third analytical section was to identify whether there were correlations between children's imitative behavior during two differing tasks, with differing structures and features of language eligible for imitation. We had

expected there to be contiguity across tasks given that children's imitation is organized around the intentions and goals of others. We had hypothesized that this shared underlying structure of imitation would result in cross task behavioral similarities. It was clear, however, that children much more readily imitated repetitions, be they grammatical or ungrammatical, than they imitated novel verbs or unfamiliar syntax. However, there were some patterns of significance evident between our two language tasks.

Children who retained a transitive structure, producing either the active or passive construction in their own utterances, were more likely to repeat grammatical sentences verbatim. Furthermore there was a negative correlation between repetition removal in the grammatical condition, and the number of transitive utterances produced. There was a strong relationship then between children choosing to produce transitive sentences and choosing to recount grammatical sentences verbatim. Furthermore there was a positive correlation between novel passive production and verbatim imitation in the ungrammatical condition. Those children who produced the most complex lexical and syntactic combination in our paradigm were also more likely to have faithfully imitated an ungrammatical model. At first this might seem odd as one is grammatically correct and the other not so, but critically in the passive verb's case no grammatically incorrect forms were produced. The interesting point then is that children who are productively grammatically competent in one sphere seem to show a paradoxical propensity to replicate an incorrect grammatical form in another task. This would suggest that imitating a modeled form is a beneficial strategy for children as it supports their syntactic competence. Children are imitative regardless of correctness in this case.



Particularly following passive training, there was a positive relationship between the number of novel verbs produced and the number of sentences imitated verbatim in the ungrammatical condition. Specifically the production of passives following passive training correlated with the verbatim imitation of ungrammatical sentences. The ungrammatical and passive training conditions were the more conceptually interesting of the two conditions in each task, and the place where imitation would have been the most challenging due to the unconventional, less familiar or simply ungrammatical content. Imitation in these two conditions would not have been bolstered by their intuitive representations as a result therefore imitation was more likely to be socially driven, based upon the intentions of the model. There were no instrumental or semantic reasons to retain the ungrammatical or unfamiliar structural elements of the utterances therefore we must conclude that a desire or strategy to imitate was responsible for this cross task relationship.

Our second aim for analysis three was to investigate the cross-domain relationships between the imitation of actions and the imitation of the linguistic features reported. In our verbal imitation paradigm there were no clearly significant relationships between faithful action and verbatim verbal imitation when looking at both our experimental action apparatus together. There was, however, a negative correlation between the number of simultaneous responses in the necessary condition and the level of verbatim imitation in the grammatical condition. These children were not adhering to the models exact behavior in either domain.

It seems that in this case the instrumental and semantic goals of the model were adhered to by some children over and above the conventional, social goals. Rather than imitating the exact sequences executed by the model children retained information in a more cognitively efficient way. Those children who were less likely

to imitate the conventional sequence of actions were also less likely to imitate the verbatim presentation of sentences. In both then the instrumental and semantic goals were retained, i.e., those goals that were most salient, at the expense of the intentional, social goals.

In addition to these results from the repetition imitation paradigm there were supporting results with novel verb imitation. For the Tower apparatus, in the critical unnecessary condition the number of Just B, selective imitation responses, correlated negatively with the number of novel verbs used. Children who were more often choosing to selectively imitate in the action domain were less likely to imitate the precise verb modeled to them during experimental trials. Furthermore those children who faithfully imitated but disregarded temporal sequence by carrying out both actions simultaneously produced fewer novel passive utterances. Again children here who rejected carrying out actions in a given sequence also rejected the use of a novel verb, in favor of a more familiar and easily accessible option. In both cases children selected the path that most quickly and efficiently led to their instrumental end goal. Carrying out two actions simultaneously took a shorter time than doing so consecutively. Using a familiar verb took less cognitive effort than using a novel one. This additional evidence strengthens the idea that being highly imitative is a domain general feature of children's cognition rather than being restricted to the type of task being modeled. Children again were imitating instrumental goals by generating the modeled end result or conveying correct semantic information but they were not doing so in an entirely faithful way.

Finally then we reported how action and syntax imitation were related. Here the most striking evidence for cross-domain imitative contiguity was found. When looking at children's production of transitive utterances there was a positive

relationship with both two action sequence globally and A then B (temporally faithful) responses specifically. Conversely children who produced fewer transitive utterances, and less active utterances after active training, were more likely to give selective Just B responses in the unnecessary condition. There is some form of structural chiming here in that children in both tasks here were either carrying out the most appropriate form of behavior across tasks or showing less structural and intentional awareness, therefore carrying out fewer conventionally imitative behaviors across tasks.

For the Tower apparatus the number of Just B responses in the unnecessary condition correlated negatively with the number of novel transitive utterances. This pattern was replicated in Rake data. Furthermore two-action, specifically A then B sequence actions, using the Rake were correlated positively with the number of transitive utterances produced.

It seems that overall children's imitative patterns are relatively contiguous across action and verbal domains. Children who faithfully imitate those elements only conventionally necessary in action sequences are also more likely to copy sentences faithfully, use novel lexical items and respond in modeled syntactic forms.

The relationships found were more complicated than we had originally envisaged. We had expected clear relationships between the imitation of language and the imitation of actions at the group level. Indeed as seen in earlier sections children at the age of four are sensitive to the necessity of actions, the grammaticality of utterances, the novelty of verbs and the familiarity and complexity of syntax. Children more readily imitate causally necessary, grammatically correct and familiar behaviors. However, perhaps due to the high levels of faithful and verbatim imitation we saw no bold and direct link between faithful two action responses and verbatim imitation as a

whole. It was only when inspecting more closely did we uncover the wealth of between task relationships contained within our two action imitation apparatus and our experimental language tasks.

#### **6.4. General Discussion**

Many researchers have presented theories that seek to identify the fundamental principles that underlie imitative behaviors. Meltzoff's (1995) behavioral re-enactment paradigm beautifully illustrated the importance of goals to the patterns of imitation seen in young children. Furthermore authors such as Bekkering et al. (2000), and Carpenter et al. (2005) highlighted that such goals are organized hierarchically and are identified differentially given certain task demands. In addition the intentions of a model are crucial in influencing what out of a sequence is replicated as illustrated so well by Gergely et al. (2002) using the head-touch paradigm. The identification of goals and the intentions of a model are just two of the aspects involved in determining what gets imitated in an action sequence.

Other researchers such as Nielsen (2006) have emphasized the social component of imitation. Imitation by its' very nature is a highly imitative behavior and has been show to be influenced by social mediators. As discussed in Chapter 5 numerous types of social mediators influence imitation in children and these manipulations fall broadly into three categories (1) the model's social behaviors (e.g., Bandura, 1971; Brugger, Lariviere, Mumme & Bushnell, 2007; Hartup & Coates, 1970; Nielsen, 2006; Nielsen, Simcock & Jenkins, 2008); (2) the model's characteristics (e.g., Buttelmann, Zmyj, Daum & Carpenter, 2013; Jaswal & Neely, 2006; Kinzler, Corriveau & Harris, 2011; Ryalls, Gul & Ryalls, 2000; Seehagen & Herbert, 2011; Zmyj et al., 2012) and finally; 3) increasing goal salience through social means (e.g., Brubacher, Roberts & Obhi, 2013; Chen & Waxman, 2013; Elsner

& Pfiefer, 2012; Hermann, Legare, Harris & Whitehouse, 2013; Simpson & Riggs, 2011).

Given these readily identifiable modifiers of imitative behavior (goals, intentions and social factors) and the conceptual assumption that such features could refer to either the action or verbal worlds we had expected there to be some parsimony between imitative decisions both within and across behavioral domains. If the underlying pillars of imitation are the same, goal, intention and socially driven, then it follows that imitative behaviors across a variety of tasks should be contiguous.

At age four children: preferentially imitated repetitions faithfully, be they grammatical or ungrammatical; imitated grammatical repetitions more faithfully than ungrammatical repetitions; preferentially rejected the use of novel verbs following eight transitive models; preferentially used the transitive following transitive modeling; vastly more often imitated active than passive models; were sensitive to discourse pressure and; were poorer at comprehending passively introduced verbs when compared to those actively introduced.

Children readily imitated the instrumental goals of an action across the board, be it the achievement of some physical goal or the communication of some set of events. Across the action and language tasks there was evidence that children do imitate elements above and beyond those that are necessary, they imitate intentional but not causally necessary features of another's behavior. Children need not imitate ungrammatical repetitions, they need not adopt novel words and they need not utilize unfamiliar syntactic constructions, yet they do. Of course some behavioral and linguistic features are imitated more readily, however, each element was imitated in differential ways dependent on condition. This provides evidence for children's ability to flexibly incorporate elements of a social partner's past behavior into their

own behaviors moving forward. In our sample children were not entirely imitative, and not entirely innovative: they were capable of different types of imitative responses across different tasks, and crucially these imitative responses were related across tasks.

There were some interesting relationships between task performances. Children who rejected the modeled sequence in one task did so more often in another. Some children were rejecting the implied, conventional structure modeled in favor of their own interpretation of or response to the event sequence. Conversely children who faithfully imitated in one task often imitated certain features more often in another. Certainly it was not the case that all children imitated all things, or indeed the opposite. However, there were demonstrable links between action imitation patterns, repetition imitation, the imitation of novel lexical items and the imitation of syntax. In every aspect of language production measured here there was some relationship with action imitation. These tantalizing aspects of contiguity across the tasks would suggest that imitation its' self is a strategy that can be deployed in a similar manner in response to a variety of different situations, both physical and verbal. This could be based upon children's identification of goals and intentions being crucial for both.

In the general introduction we set up two reasons as to why language is an invaluable habitat in which to study children's navigation of actual versus possible worlds. Firstly language is a tool that can be manipulated and used to change the world. Secondly language is paradoxical as it relies on the contrasting behaviors of imitation and innovation. In our sample children proved capable at using language in both tasks, repeating target sentences and producing many transitive utterances. In addition they showed both imitativeness and innovativeness through their use of novel verbs and discourse sensitivity. It was not the case that children only used these newly

learned verbs in the constructions that they had been modeled, although they did follow that general trend. The verb island hypothesis (Tomasello, 1992) predicts that children around this age only have access to item-based constructions therefore should not typically apply the necessary verb general rules to produce novel verbs in a different construction. Our results then clearly challenge this hypothesis. Although children more often used the verb in the modeled construction, this was not an exclusive strategy.

In our introduction to Experiment 1 we set out some features of imitation evident in the action domain that we may expect to see in the verbal domain if indeed a similar process is at work. In particular the first three can be spoken to by our data.

(i) *Children would more readily selectively imitate causally relevant features of the language* i.e. those that achieve some communicative goal. Indeed the children in Experiment 1 were more willing to imitate grammatically correct than incorrect utterances. Furthermore they were less likely to imitate novel verbs and syntactic constructions that were not necessarily related to communicative goals. Bannard et al. (2013) reported that three year-old children were more likely to use novel words in situations that demanded them, unlike ours where English verbs could plausibly be substituted.

(ii) *Children would more readily imitate intentionally produced features of the language.* Although we did not have conditions in which unintentional features of language were modeled it was clear that many children imitated intentionally produced, grammatically incorrect and novel items.

(iii) *Children would faithfully imitate more incidental or non-critical features of language with age.* Again it was clear that children often imitated repetitions and syntax, two elements that were non-critical to the communicative outcome. However,

they were more likely to imitate repetitions than complex syntax at this age. Older children and adults have been shown to be more readily imitative with regards syntax in previous studies (e.g., Bock & Griffin, 2000).

In conclusion then the imitation of certain features of language follows some similar trends to those evidenced in the action domain. Furthermore there were some direct patterns of correlation between performance both within the different language tasks and between the action and language data. Children are able to flexibly use the behaviors they see others model in their own repertoire as a means of manipulating the world through either their interaction with an object or their production of language. Children can be imitative and innovative in the same breath and their willingness to utilize novel lexical items and syntactic constructions in related to their understanding of the underlying causal structure of the utterances they have seen modeled. This replication is based on the understanding, and hierarchical organization of goals and intentions.

#### **6.4.1. Limitations**

The first experimental paradigm detailed in this chapter saw children playing a game where there were asked to say what the experimenter had previously said. This task was adapted from a paradigm used with younger children, aged three years (Over & Gattis, 2010). As a result it may have been more appropriate for us to have generated and tested new target sentences. Levels of verbatim imitation were very high and it may have been the case that the sentences were simply too easy for our age group to elicit the levels of variance we had hoped to see. An example sentence is “Sam was a big big cat.” We could have manipulated sentence complexity in several ways whilst retaining the same premise of having grammatical versus ungrammatical repetitions. Sentences could have been made longer, the lexical items chosen could



have been less familiar or difference syntactic structures could have been utilized. With more complex sentences some children may have relinquished a faithful strategy as the communicative goal became increasingly tricky to fulfill alongside the social goal.

Bannard, Klinger & Tomasello (2013) generated a novel experimental situation in which imitation could be interpreted as instrumentally necessary (a contrastive adjective) in some situations and not so in others. Children were playing a game in which they requested an object from a confederate. Regardless of whether they modeled faithfully or otherwise the child received the requested object. They showed that the instrumental necessity of the word impacted upon its usage in their three year-old participants. Utilizing a similar paradigm but one in which utterance production did determine whether the instrumental goal was achieved would have been a fascinating addition to our repertoire of tasks. For example you could set up a necessary condition where the use of an additional contrastive adjective was necessary to distinguish between two similar objects i.e., *Please pass me the blue duck* when the options were a blue duck or a red duck. In the unnecessary condition then the model could use the same contrastive adjective but the options were a blue duck or a blue frog. It would have allowed verbal imitation to be more closely compared and contrasted with action imitation with one condition having two necessary actions modeled and the other having one necessary and one unnecessary for goal completion.

In many ways the passive verb production task yielded the opposite problem in that performance rates for the target passive construction in particular were low. In the original paper using this methodology (Brooks & Tomasello, 1999) children were given eighty-eight models, this is compared to our sixteen. Production rates of the

passive have been shown to increase with increased production (e.g., Messenger, 2009; Vasilyeva et al., 2006). Perhaps increasing the number of models would have supported children's representations of the unfamiliar syntactic structure just enough to increase production rates. Furthermore explicitly tasking the children with repeating the models would have increased the salience and availability of the structure for subsequent manipulation in their own utterances (Shimpi et al., 2007).

#### **6.4.2. Conclusions**

In Chapter 5 we reported that children were prolific imitators in an action imitation paradigm. In the current chapter we shifted our focus to language production. Competent language production embodies a necessarily paradoxical combination of imitation and innovation. We reported results that characterize children as prolific imitators of some aspects of language and not others. Despite being sensitive to grammaticality children preferentially imitated both correct and incorrect repetitions. Conversely children more frequently described a novel action using an English verb than the novel verb modeled during a demonstration. Further children more often than not rejected a complex syntactic form in favor of a more familiar one. Despite these group level findings there was also evidence that children were more likely to use the more difficult syntactic form following a model that used the same structure. This provides support for a theory that some children at least are beginning to imitate complex surface features of language. Additionally our results suggest that some children demonstrate discourse sensitivity, that is, they modulate the syntactic structure of their responses to the attentional focus of the questioner.

Finally we showed that although being highly imitative in one task does not dictate high levels of imitation elsewhere there are striking links between behavior both within the language tasks and across the language and action domains. How

children represent and replicate the intentional, goal directed actions of others, be they actions or utterances, show relative contiguity.

**Chapter 7: The World as it Is, Was, Could Have Been and Could Be: The  
Relationship Between Reasoning About, and Manipulating the World Through  
Language**

**7.1. Abstract**

In Chapter 6 we focused on the imitation of language. We characterized how readily the children in our sample would imitate certain features of intentionally produced language. However, as discussed in Chapter 6 language is paradoxical in that it is a behavior encompassing both imitative and innovative aspects. To become a truly competent language user children must be able to use the conventional tokens and structure of their native language in productive and flexible ways. The communicative flexibility that language affords us is one of, if not *the* single most powerful tool humans have to manipulate the world. Furthermore language is arguably the essential element required for the true navigation of the actual and the possible as it is through language that we as humans can uniquely represent and reflect upon how the world is, how the world was and crucially how the world could be. In this chapter we sought to characterize our sample's use of language as a tool. We derived a language competence score for each participant from discrete aspects of the experimental language data we collected throughout the two testing sessions.

We then sought to investigate whether the indices of reasoning and imitation reported earlier were too related to children's use of language as a tool. If both causal and mental state understanding are involved in imitation, then it may be expected that they too are related to language learning more generally based upon language's highly imitative origins. This chapter speaks to the question of whether reasoning about and acting based upon ones knowledge of the physical and social world as it was, could have been and could be, is predictive of linguistic competence more generally.

## 7.2. Introduction

In the previous chapter the imitation and innovation of children's utterances was our main focus. However, as outlined briefly at the start of Chapter 6 language is interesting for the purpose of our study not only because of this paradoxical relationship between imitation and innovation, but also because of language's integral function as a tool. Language, in addition to action, is another instrument for change children can add to their arsenal. Productive language learners may suddenly be able to gain a social partner's attention, request objects or information from others and provide information in return. Receptive language learners become privy to more complex ideas and can learn about events temporally or physically removed from their current reality. With increasingly complex lexical and syntactic skills children are better able to interact within the world. They are better able to appraise situations based on linguistic input, and better placed to alter their own worlds through verbal means. Language then is a tool, a facilitator of expressing physical information, a vehicle for establishing conditional relationships and a means for communicating mental state information. In addition, according to some evolutionary theories of language, it is a pre-requisite for representing and reasoning about complex models of the world. Only with a competent grasp of language can we engage in counterfactual or mental state reasoning.

In Chapter 5 (5.2.7.) we discussed in detail the types of cognitive demand that imitation may place upon children. Indeed many studies have probed some of the specific links mentioned in the action domain. However, imitation in the language domain could be hypothesized to require these very same cognitive and socio-cognitive skills.

Executive functions are integral to imitation as children are required to represent a situation in linguistic terms, switch between their own and others descriptions of a linguistic and select their own language from a potentially infinite number of combinatory possibilities. As with the action domain if children were not engaging EFs then they would replicate failed and successful behaviors equally through not being able to select or inhibit one representation in favor of another (Behne, Carpenter, Call & Tomasello, 2005; Meltzoff, 2005; Olineck & Poulin-Dubois, 2009). We see in both our action and verbal imitation paradigms that children do treat causally necessary or grammatically correct models differently. Children do not copy the coughs, stutters or other incidental vocal actions of others: they copy selectively at some level.

In the action domain children build representations based around goals they have identified (Bekkering et al., 2002). Additionally, information about the model's social behavior and their social characteristics are further incorporated into this representation. Children require the key executive functions of inhibitory control, cognitive flexibility and working memory to successfully imitate based on these representations in a way that fulfills both their instrumental and social goals.

Identifying these linguistic goals, be they instrumental, i.e., communicative, or social, then is another key ability needed for imitation. In early childhood, social attention (Olineck & Poulin-Dubois, 2009), point following (Carpenter et al., 1998) and, pointing, reaching, showing and checking back (Bretherton, McNew & Beeghly-Smith, 1987), are all suggested to facilitate imitative behavior. Measures related to sociability more generally also have been related to enhanced imitative behaviors (Chartrand & Bargh, 1999; DiYanni, Nini & Rheel, 2011; Hilbrink et al., 2013; Uzgiris, 1981). These socio-cognitive skills highlight the goals of an action for the

observer and may too then be implicated in the imitation of language. Indeed early socio-cognitive indexes have been linked to language development specifically. Tamis-LeMonda and Bornstein (1989) reported that five month-olds' who more rapidly habituated to visual stimuli have larger vocabularies at thirteen months. This work was replicated and extended by Dixon and Smith (2008) who found that in addition to decrement of attention, attentional focus also predicted vocabulary size over a year later. Much like imitation, temperamental easiness relates to greater, and temperamental difficulty relates to lesser vocabularies in early childhood (Salley & Dixon, 2007; Todd & Dixon, 2010).

Such features contribute to joint attention: a state of shared, triadic focus between two social partners and some object or event in their shared environment (Mundy & Gomes, 1998). In turn joint attention is widely regarded as a key facilitator in language development (e.g., Akhtar, Dunham & Dunham, 1991; Brooks & Meltzoff, 2005). In imitation joint attention, although not always strictly necessary as children can imitate in a third party context, can be seen to play a huge part. A “referential triangle” constituted of children: following the gaze of others; using others as social “reference points” and subsequently; imitating their actions was suggested by Carpenter, Nagell and Tomasello (1998). It seems possible then that language imitation may similarly benefit from such joint attentional skills. In the language domain, however, there is often no present physical object or situation being discussed. Children may be reliant wholly on generating their own representations based on the language they are hearing. In our two experiments the demands of generating ones own representation entirely independently was present in the verbal imitation task, but not the passive verb task where the scene being described was plainly visible.

In addition to being able to attend to certain linguistic features, being able to understand the communicative or linguistic intentions of a model would also be expected to facilitate imitation as seen in the action domain (e.g., Lyons, Young & Keil, 2007; Over & Gattis, 2010; Gergely et al., 2002; Sakkalou et al., 2012) highlighted the need to represent behavior in terms of mental states. This too could be extended to include not only physical manipulations of objects, but linguistic behaviors also.

In Section 5.2.9 we discussed at length the potential theoretical links between reasoning skill and imitative behaviors in the action domain. These included: the crucial need for mental state understanding when reasoning about goals and intentions; the importance of being able to reason about causality and counterfactuals for imitative behavior and; the necessity of executive functions during imitation. These same elements would be expected to contribute during the flexible use of language. As far as the authors know language competence has not been investigated in the context of MSU, CFR, EF and imitation in the same sample and at the same time point. Language competence is more regularly associated with reasoning skill. Guajardo et al. for example (2009) reported significant positive correlations between children's language skills (as measured by the TACL-3: Carrow-Woodfolk, 1999) and their false belief scores ( $r=.72$ ), generative counterfactual scores ( $r=.69$ ) and several measures of EF. In fact Guajardo et al. suggest that the relationship between counterfactual reasoning (as measured using a generative paradigm) and MSU (as measured by false belief) is driven by shared variance in language and EF. Language has further been associated with CFR and MSU separately (e.g. Milligan, Astington & Dack, 2007; Guajardo & Turley-Ames, 2004; Perner et al. 2004). It is unsurprising that language is so closely related to reasoning. Comprehending language increases



the amount of information about situations and mental states that children are exposed to. Language provides access to mental states that may otherwise go unrecognized due to their internal nature. Conversely, however, reasoning about the words people use, and the way they chose to use them can allow children to infer meanings. In a sense then language and reasoning become synergistic. Language is a means to represent causal structure in a way that may be impossible without it (Chomsky, 2002). Language can highlight mental states that may otherwise remain unknown and undetectable. Language too then may be able to influence, change or inform these previously incomprehensible situations in a positive way.

Our aim for this chapter was to go above and beyond relating children's language and reasoning. We wanted to assess whether their use of language as a tool was related to: their imitation of both actions and linguistic features and; their reasoning about and performance on tasks integrating information about different possible worlds. Linguistic competence reflects how skillful children are at understanding, producing and, accurately and appropriately using language as a tool to change their world. Being readily able to imitate certain linguistic features in a restricted context alone will not aid children in the real world to the same extent as their being able to understand and produce language flexibly, crucially integrating both imitative and innovative strategies. Causal knowledge affords an individual the ability to make wide ranging predictions about the world (Gopnik, Sobel, Schulz & Glymour, 2004). This world may be physical or social and may be altered through actions and utterances: these are children's tools for manipulating the world based upon their knowledge of how their decisions will impact upon a desired future world and state of affairs. Competence of this sort is a more practical assessment of

children's language sophistication than reliance on an assumption that the imitation of language is a suitable proxy.

### **7.3. Method**

In order to assess children's linguistic competence in this broader, productive way we utilized elements of the two experimental language tasks we had administered during testing. We chose not to use a parent report method or extensive vocabulary style measure but instead integrate a number of diverse features of children's productive and receptive language skills inherent in our verbal imitation and novel verb tasks.

We first choose a comprehensive list of the features of language we could acutely observe through our experimental methods but that were distinct from the experiments primary goals. From the verbal imitation task then we included children's imitation of the filler sentences. These five sentences were grammatical utterances of a matched length to the target experimental sentences that contained repetitions. This was sensitive to children's ability to recall and reproduce simple multiword utterances.

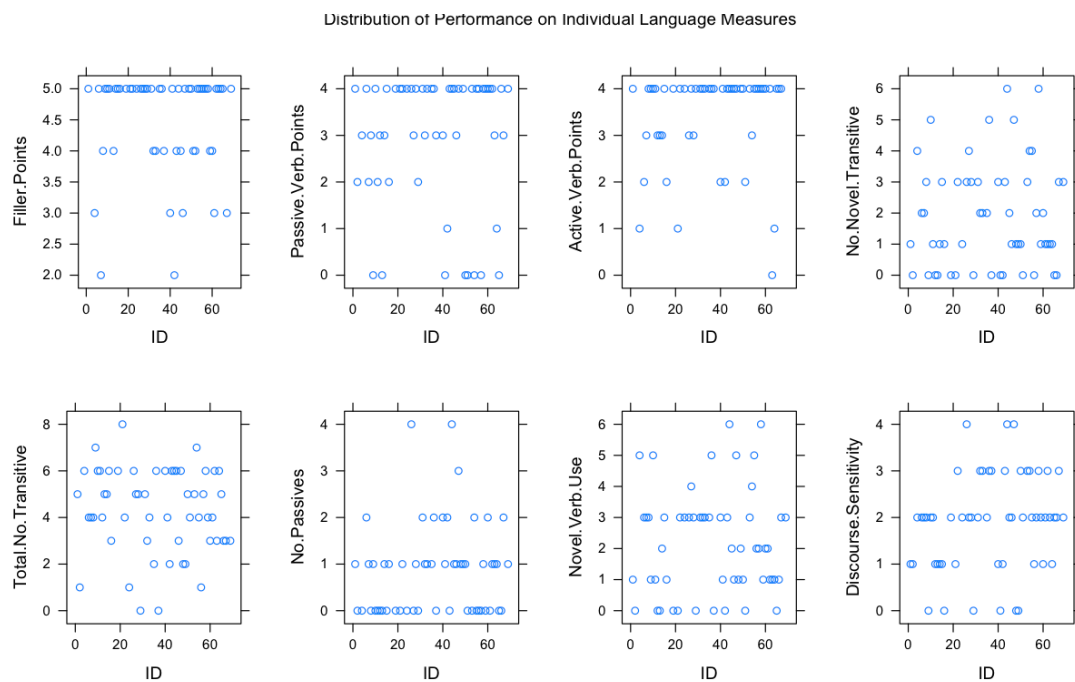
From the transitive verb task we extracted seven more potential competence variables. The first two of these were comprehension based. Each participant was asked eight comprehension questions during this task, four following passive construction training and four following active construction training. These questions assessed whether children had understood the sentence structure fully enough to act out a requested scene using novel agents and patients. Two separate competence scores were yielded then, passive and active, each with a maximum of four.

Four features of language production were then included for inspection. These were extracted from the utterances children produced when asked to describe a

modeled scene in either the agent, patient or neutral focused questions. We chose; how many transitive utterances, of either the active or passive construction were produced by the participants using the novel verb and overall; how many passive utterances were produced overall and; how many novel verbs were produced in any construction. Each of these variables could range from zero to six.

Finally we coded each utterance made in response to an agent or patient focused question as being discourse sensitive or otherwise. Discourse sensitivity was based upon the syntactic construction of the utterance being appropriate in the context of the question (i.e., an active construction following and agent focused question) and could have included a novel or English verb. Scores on this final variable could range from zero to four.

We first plotted each of these variables to visually inspect their distributions (see Figure 7.1.). Filler sentence imitation and verb comprehension had comparatively high scores in contrast to passive construction production that had lower scores.



*Figure 7.1.* Scatterplots showing the distribution of scores on our eight potential language competence indices. Scores are plotted against the y-axis and participants against the x-axis.

We further carried out pairwise correlations between each of these variables to identify whether any were superfluous through large positive or negative associations (see Table 7.1. below). Some variables were unsurprisingly correlated due their procedural overlap with items being counted in two or more categories. These included novel and total transitive production scores, passive production scores and the number of novel verbs produced.

Table 7.1.

*Non-parametric correlation table showing size of linear correlations and significance values for our eight potential linguistic competence items.*

|                    | Filler | Passive<br>Comprehension | Active<br>comprehension | Novel<br>Transitive<br>Utterances | Total<br>Transitive<br>Utterances | Passives<br>Produced | Novel Verbs<br>Produced | Discourse<br>Sensitivity |
|--------------------|--------|--------------------------|-------------------------|-----------------------------------|-----------------------------------|----------------------|-------------------------|--------------------------|
| Filler             |        | .18                      | .16                     | .08                               | .24                               | .02                  | .1                      | -.03                     |
| Passive Comp       |        |                          | .32*                    | .4**                              | .06                               | .09                  | .45**                   | .3*                      |
| Active Comp        |        |                          |                         | .14                               | -.11                              | -.15                 | .18                     | .27*                     |
| Novel Trans        |        |                          |                         |                                   | .39**                             | .37**                | .94***                  | .56***                   |
| Total Trans        |        |                          |                         |                                   |                                   | .12                  | .29*                    | .26                      |
| Passives Pro       |        |                          |                         |                                   |                                   |                      | .33*                    | .21                      |
| Novel Verbs<br>Pro |        |                          |                         |                                   |                                   |                      |                         | .52***                   |
|                    |        | p<.05*                   |                         | p<.01**                           |                                   |                      | p<.001***               |                          |

Only looking at relationships between independent indices then there were numerous significant positive correlations evident. Only active comprehension scores showed any negative associations and these were non-significant. Filler comprehension scores was the only item not to correlate significantly with another measure, however, a positive association with the total number of transitive utterances produced was approaching significance,  $r=.24$ ,  $p=.09$ .

Passive comprehension scores correlated significantly positively with the number of novel transitive utterances,  $r=.4$ ,  $p<.01$  but not the total number of transitive utterance,  $r=.06$ ,  $p=.67$ . Passive comprehension further positively correlated with the number of novel verbs produced,  $r=.45$ ,  $p<.001$  and, discourse sensitivity,  $r=.3$ ,  $p<.03$ . Active comprehension only correlated significantly with discourse sensitivity,  $r=.27$ ,  $p<.05$ .

The number of novel transitive utterances correlated significantly positively with discourse sensitivity,  $r=.56$ ,  $p<.001$ , whilst the correlation between the total number of transitive utterances produced was much smaller and marginally significant,  $r=.25$ ,  $p=.052$ . Finally discourse sensitivity also positively correlated with the number of novel verbs produced,  $r=.52$ ,  $p<.001$ .

We selected five of these variables to comprise our linguistic competence score. We wanted to included variables that showed variation between individuals. As such we rejected active verb comprehension as there was only a limited amount of variance due to the generally high performance.

Furthermore we wanted variables that were not highly correlated in order to avoid excess shared-variance being incorporated. Significant strong positive correlations were seen between the number of novel transitive utterances, and the use of novel verbs with many of the other variables so these two variables were

eliminated. Passive and active-verb comprehension were also highly correlated, strengthening the decision to eliminate active comprehension from our composite score.

Finally we could not include sets of variables that were theoretically overlapping. Both novel and total transitive utterances scores for example would count many of the same responses therefore only total transitive utterances were used. The same shared items would exist between novel transitive and total novel verb utterances again giving reason for us to eliminate them from our final composite score.

Based on these statistical and theoretical requirements we selected: filler sentence verbatim imitation score; passive verb comprehension score; the total number of transitive utterances produced; the number of passive constructions produced and; discourse sensitivity. The distribution of these five final variables can be seen below in Figure 7.2.

Only one significant positive correlation was present between these variables, that was passive comprehension score and discourse sensitivity with a correlation of  $r=.3$ ,  $p<.05$ . This was deemed a low enough correlation to have avoided excessive parsimony. In addition one variable derived from participants' language production and the other their language comprehension.

To ensure that each variable contributed an equal amount of variance to the final composite variable we computed standardized scores between 0-1 for every participant on each of the five selected variables. These were then combined to generate their final language comprehension score. The resultant composite language competence score ranged from zero to five,  $M=2.9$ ,  $Std. Dev. = .73$ . (see Figure 7.3.)

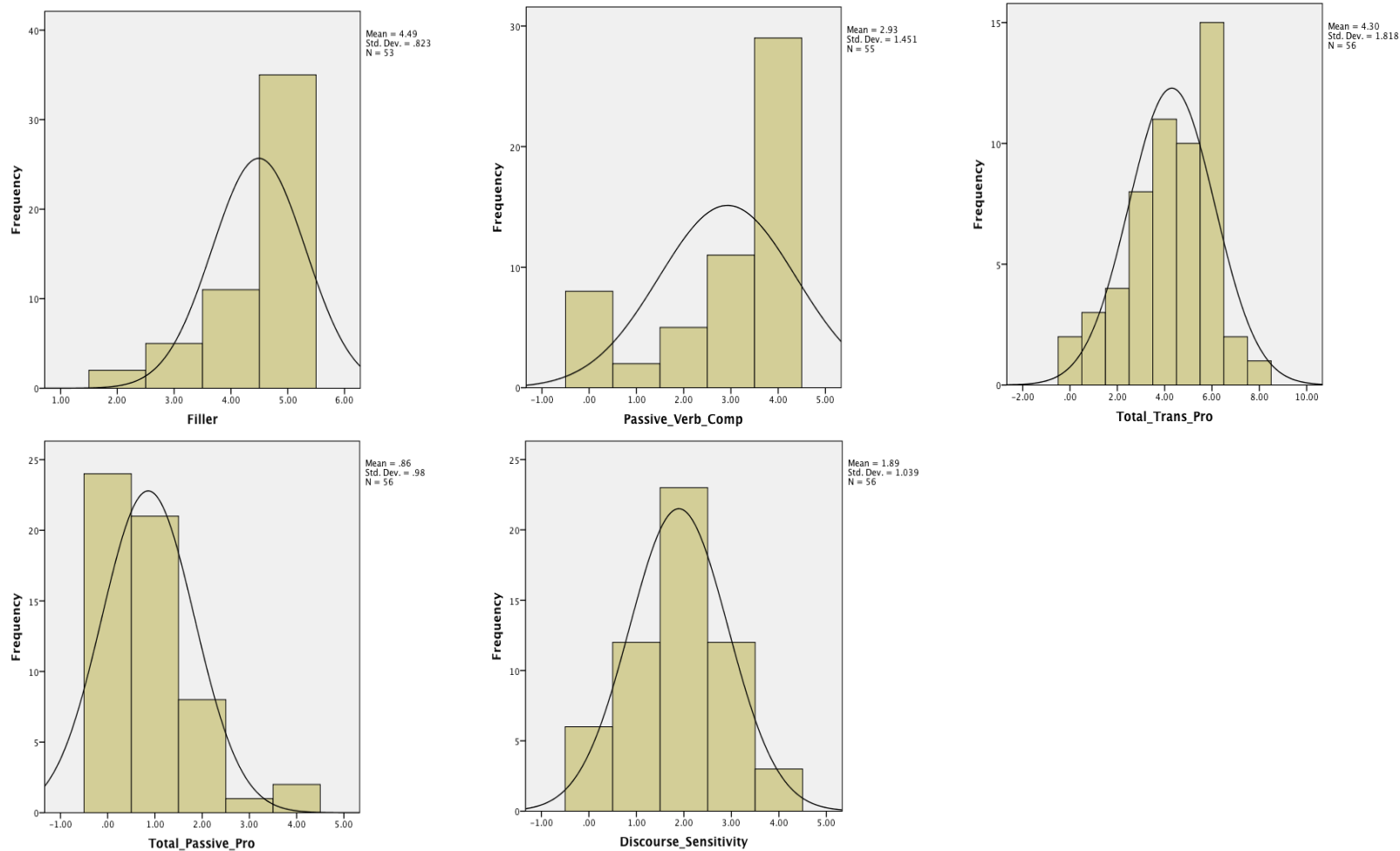


Figure 7.2. Histograms showing the distribution of scores on filler sentence imitation, passive verb comprehension, total number of transitive utterances produced, total number of passives produced and discourse sensitivity.



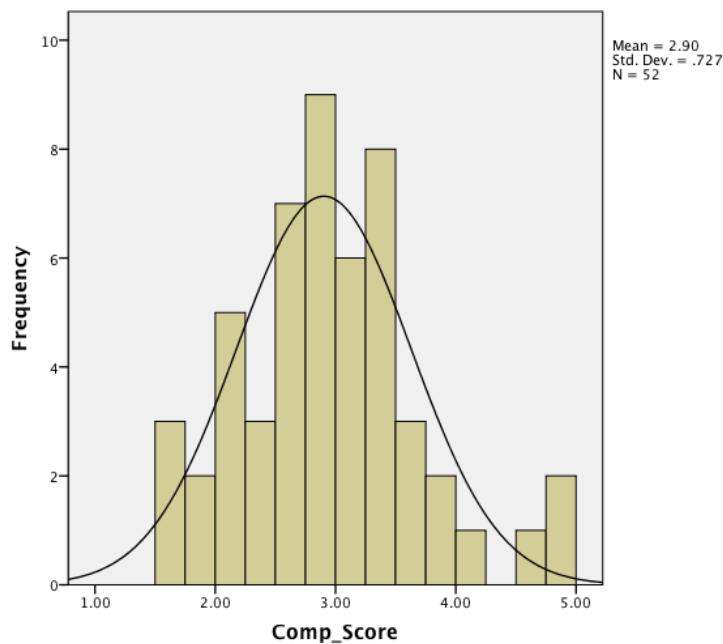


Figure 7.3. Histogram showing the distribution of language competence scores.

## 7.4. Results

This section shall inspect the relationships between our language competence score and data collected during all our earlier experimental paradigms; counterfactual conditional reasoning, executive functioning, mental state understanding, object imitation and, verbal imitation.

### 7.4.1. Counterfactual Conditional Reasoning

Parametric correlations were carried out between language competence and counterfactual reasoning performance scores (see Table 7.2.). There was a positive correlation between standard slide performance and language competence,  $r=.35$ ,  $p<.001$ . Furthermore there was a strong positive correlation between story performance and language competence,  $r=.33$ ,  $p<.05$ . This relationship was driven by the atypical condition with a correlation of  $r=.23$ ,  $p<.05$ , while no such significant relationship was seen in the typical condition,  $r=.15$ ,  $p>.05$ . The composite CFR score, combining the three tasks, was also correlated with language competence,  $r=.37$

$p < .01$  (see Figure 7.4.). A linear regression was carried out to assess the variance in language competence that could be attributed to CFR reasoning. 9% of the variance in language competence can be explained by CFR performance  $R^2 = .09$ ,  $F(1,53) = 4.73$ ,  $p < .05$ . It was found that CFR performance significantly predicted language competence  $\beta = .1$ ,  $p < .05$ .

Table 7.2.  
*Non-Parametric Correlations Between Language Competence and Scores on the Three Counterfactual Conditional Reasoning Tasks.*

|                     | Standard Slide | Typical Stories | Atypical Stories | Composite CFR    |
|---------------------|----------------|-----------------|------------------|------------------|
| Language Competence | .35***         | .15             | .23*             | .35***           |
|                     | $p < .05^*$    |                 | $p < .01^{**}$   | $p < .001^{***}$ |

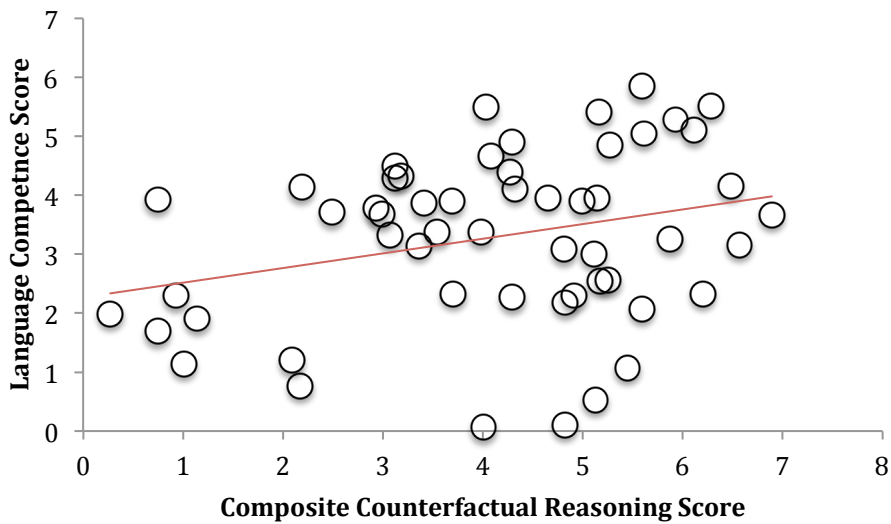


Figure 7.4. Scatterplot (including jitter) showing composite counterfactual reasoning scores on the x-axis and language competence scores on the y-axis.

### 7.4.2. Executive Functioning

Given the distribution of our EF scores (discussed in Chapter 2) we performed one-way ANOVAs using first Flanker performance and then DCCS performance as independent variables and language competence as the dependent variable (see Table 7.3.). Language competence scores were significantly higher for those in the high Flanker group than those in the low group,  $F(1,52)=5.1, p<.05$  (see Figure 7.5.).

Although this same pattern was evident when DCCS scores were used as the independent variable the results did not reach significance  $F(1,52)=2.5, p>.05$ .

Table 7.3.

*F values Obtained Using One-Way ANOVA with Executive Functioning Scores Used as Predictors for Language Competence.*

|                     | Flanker   | DCCS           |
|---------------------|-----------|----------------|
| Language Competence | 5.1*      | 2.5            |
|                     | $p<.05^*$ | $p<.01^{**}$   |
|                     |           | $p<.001^{***}$ |

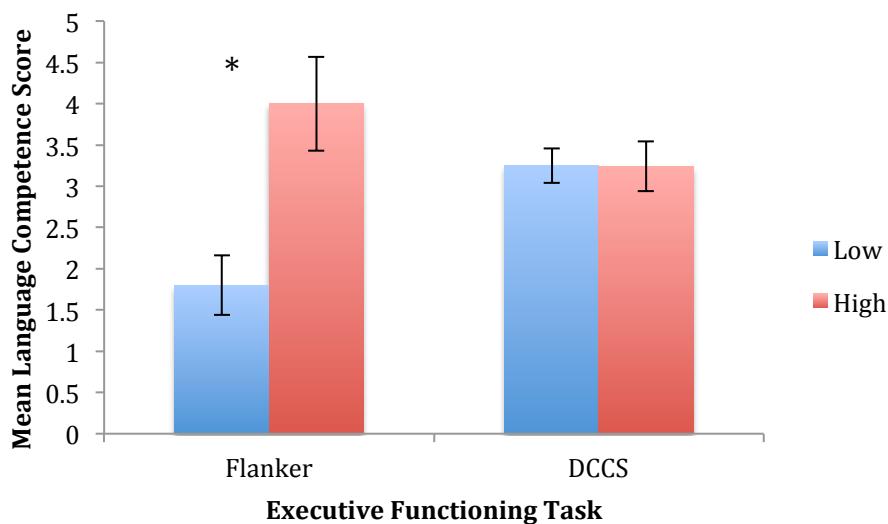


Figure 7.5. Bar chart showing mean language competence scores on the y-axis for executive functioning task separated into high and low performance groups. \*denotes a significant difference between group  $p<.05$ .

### 7.4.3. Mental State Understanding

Non-parametric correlations were carried out using the total MSU battery score. The correlation was positive but very weak and far from reaching statistical significance  $r=.14$ ,  $p>.05$  (see Figure 7.6.).

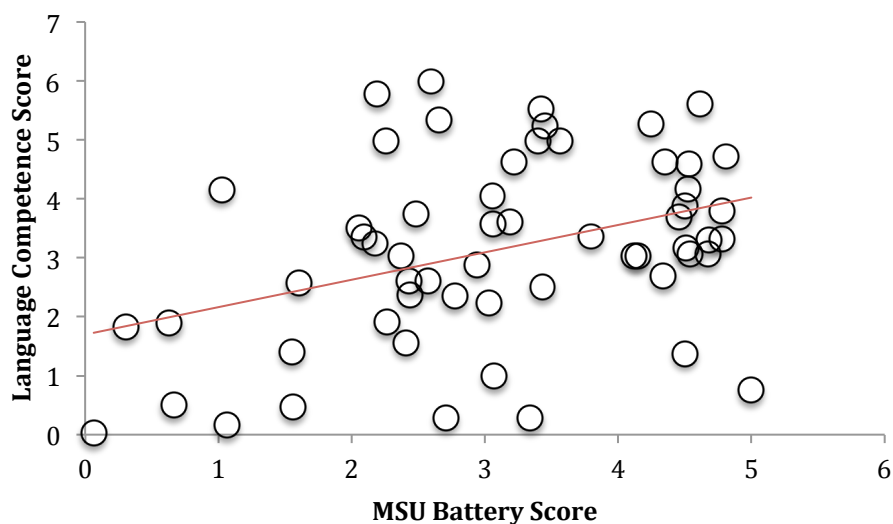


Figure 7.6. Scatterplot (including jitter) showing MSU battery scores on the x-axis and language competence scores on the y-axis.

We then chose to inspect each battery item individually to ascertain whether performance on any of the items would be related to language competence scores. A series of one-way ANOVAs were performed using each MSU battery item in turn as the independent variable while language competence was the independent variable (see Table 7.4). One item, Q3 showed the expected pattern. Participants who answered this question correctly had significantly higher language competence scores than those participants who answered it incorrectly,  $F(1,52)=4.0$ ,  $p<.05$ .

Table 7.4.

---

*F-values Obtained from One-Way ANOVAs Using Each of the Five Mental State Understanding Items as Predictors for Language Competence.*

|                     | Q1     | Q2  | Q3      | Q4  | Q5        |
|---------------------|--------|-----|---------|-----|-----------|
| Language Competence | .4     | .01 | 4.0*    | .14 | 1.2       |
|                     | p<.05* |     | p<.01** |     | p<.001*** |

#### **7.4.4. Object Imitation**

The first step taken was to perform non-parametric correlations between the number of two action responses in each object imitation condition and language competence scores (see Table 7.5.). Correlation coefficients for both the necessary (see Figure 7.7.) and unnecessary condition (see Figure 7.8.) were very small and did not reach significance. Results were also separated into apparatus but this pattern remained. Just B, that is single action responses were also correlated against language competence scores with the same outcome, very small coefficients and results not nearing significance.

Table 7.5.

*Non-Parametric Correlations Between Language Competence and Object Imitation Performance, Two Action and Single Action Responses, in Each Condition and Separated by Apparatus.*

|                     | <u>Necessary Condition</u> |       | <u>Unnecessary Condition</u> |           |
|---------------------|----------------------------|-------|------------------------------|-----------|
|                     | <u>Two Action</u>          |       | <u>Two Action</u>            |           |
| Language Competence | -.03                       |       | -.02                         |           |
|                     | Rake                       | Tower | Rake                         | Tower     |
| Language Competence | -.29                       | .08   | -.16                         | .1        |
|                     | <u>Just B</u>              |       | <u>Just B</u>                |           |
|                     | Rake                       | Tower | Rake                         | Tower     |
| Language Competence | .29                        | .01   | -.04                         | -.29      |
|                     | p<.05*                     |       | p<.01**                      | p<.001*** |

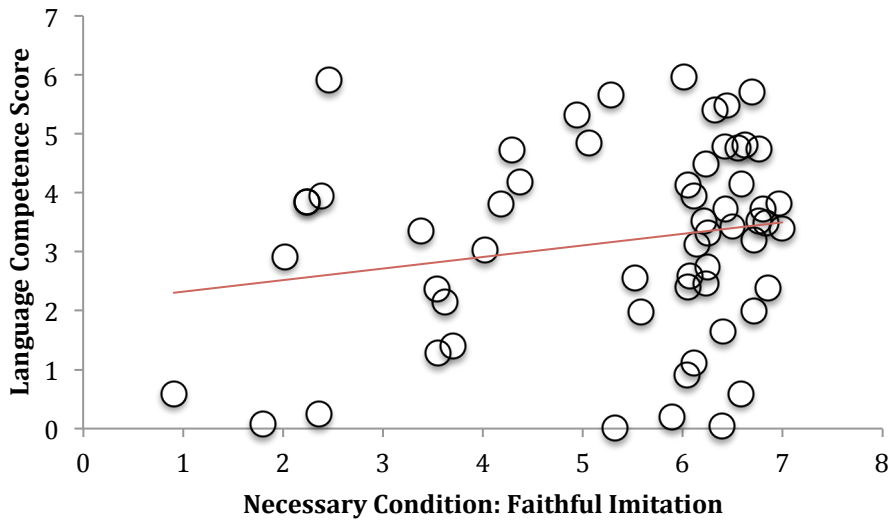


Figure 7.7. Scatterplot (including jitter) showing faithful, two-action responses in the necessary condition on the x-axis and language competence scores on the y-axis.

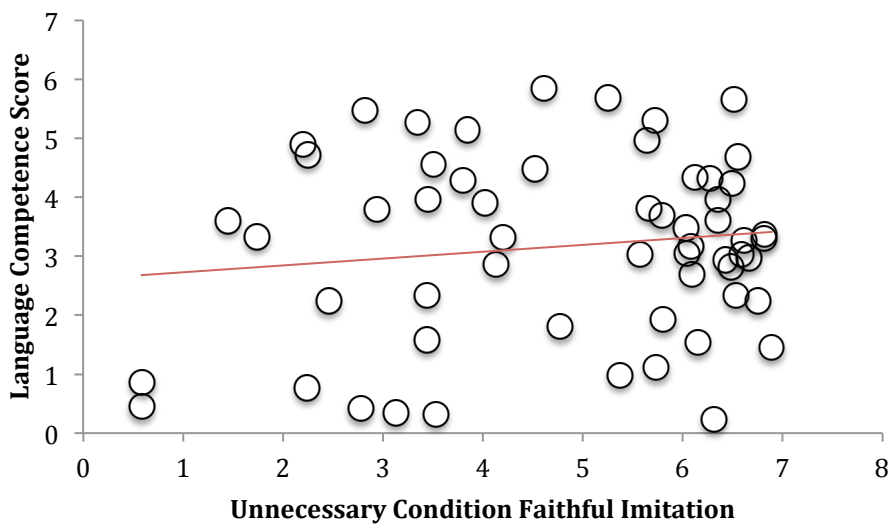


Figure 7.8. Scatterplot (including jitter) showing faithful, two-action responses in the unnecessary condition on the x-axis and language competence scores on the y-axis.

### 7.5. Discussion

Our aim for Chapter 7 was to extend our investigation of children’s imitation of language to include their proficiency in using language as a tool. We were interested in examining the relationship between children’s reasoning about and

acting based upon their knowledge of the physical and social world to see if these skills were predictive of linguistic competence more generally.

Rather than use a parent report or checklist type measure of language competence we created a composite variable derived from the language opportunities given to our participants during their experimental visits. The resultant variable encapsulated elements of grammatical production and comprehension in addition to information about how responsive children were to the questions directed toward them. Language competence scores were then correlated with performance scores on the other main experimental tasks carried out over the two testing sessions.

### **7.5.1. Counterfactual Conditional Reasoning and Language Competence**

Counterfactual reasoning was investigated in Chapter 2. Participants were asked questions of varying difficulty aimed at elucidating their counterfactual reasoning skill. The slide task (Beck et al., 2006; Beck & Guthrie, 2011) was a simple state change counterfactual with a readily observable causal structure, no intentional agent involvement and language cues inherent in the questioning. The story task was much more challenging with causal structure not readily observable but supplied verbally by the experimenter, an intentional agent (the protagonist) was present and the question structure was less transparent. Despite these very different task demands performance on both the slide and story tasks correlated strongly with language competence. Language and counterfactual reasoning's positive relationship has been reported before in Guajardo and Turley-Ames (2004). Furthermore when looking separately at the typical and atypical conditions of the story task only the atypical condition retained this strong positive correlation. Remember, the typical condition was so called as the correct counterfactual reasoning conclusion was also the conclusion supported by typicality and therefore could have been answered through

basic conditional, as opposed to counterfactual conditional reasoning processes. The atypical story condition's correct answer could only have been derived through true counterfactual conditional reasoning through adhering to the nearest possible world hypothesis and the counterfactual world differing from the actual only in those elements causally necessary given the counterfactual assertion. That the atypical, and not the typical story performance correlated significantly with language competence is a striking result. It seems that children's ability to reason about the causal structure of the world in this hypothetical, and truly counterfactual way is distinctly related to language competence. Such a relationship, upon inspection is not entirely surprising. Children being able to reason not only about how the world is, but how the world was and could be, allows for them to cognitively explore and reflect upon more possibilities and situations than they have physically experienced or perceived: it increases the size of the child's cognitive world immeasurably. Just like physical cause effect relations, linguistic cause effect relations can too be reasoned about and experimented with through counterfactual conditional reasoning processes. What seems clear in our data is that children's reasoning about physical counterfactual conditional scenarios is consistently positively correlated with their linguistic competence at age four.

### **7.5.2. Executive Functioning and Language Competence**

Next we focused on the relationship between executive functioning and language competence in our sample. Executive functioning has previously been shown to predict linguistic ability in young children (e.g., Carlson & Moses, 2001; Hughes, 1998). As expected those children who performed better on our two measures of executive functioning also had higher language competence scores. Those high in inhibitory control, as measured by the Flanker task, had significantly



higher language competence scores. Those in the high DCCS group also had higher scores but this difference did not reach significance. Executive functions are the hallmark of effortful, goal directed and flexible actions and as such we expect them to be involved in a diverse variety of actions. The positive relationship between EF and language in our sample support this assertion. That inhibitory control, a process so vital in avoiding realist answers during counterfactual reasoning, relates to language alongside counterfactual reasoning as seen earlier is very interesting. Children who are more reflective and base their actions or responses on causal structure rather than heuristic, pre-potent responses also have concurrently better language skills.

### **7.5.3. Mental State Understanding and Language Competence**

In Chapter 3 we collected data on our sample's mental state understanding using a five-item battery (Wellman & Liu, 2004). Mental state understanding has many times been related to language sophistication (e.g., Bretherton & Beeghly, 1982; Brown et al., 1996; Chiarella et al., 2013; deVilliers & Pyers, 2002; Milligan et al., 2007; Olineck & Poulin-Dubois, 2005). This is despite many different ways of investigating mental state understanding and of characterizing language performance. Our language competence score, however, did not correlate significantly with battery score performance. This result was puzzling due to the large amount of literature reporting strong significant correlations between these two abilities. When we looked at each item individually only one, question three, showed the expected pattern. Those who answered this item correctly had significantly higher language competence scores than their peers who answered it incorrectly. Item three is a knowledge access question. An opaque box is presented to the participant and they are asked to guess what is inside. The contents are then revealed: a plastic ball. The box is closed again and Polly the doll is introduced to the scene. Polly has never ever seen inside this box.

The participants are asked, “Does Polly know what is in the box?” It seems then that relationships between language and mental state understanding are more specific than definable through our battery and language competence scores alone. What is clear, however, is that mental state understanding and language are broadly related and task demands mediate this relationship greatly.

#### **7.5.4. Imitation and Language Competence**

We then moved on to looking at how object imitation patterns related to general language competence. We expected a strong positive relationship due to language’s inherently imitative nature. However, there were no significant correlations between either two action or single action responses and language competence scores. Even when looking at apparatus and condition separately correlation coefficients were small and did not reach significance.

The newly assembled language competence score proved a worthy addition to the language data already collected as it allowed us to see more general relationships between language and cognitive development. There were striking relations between language competence and counterfactual conditional reasoning, inhibitory control, certain elements of mental state understanding and verbal imitation. What was surprising, however, was the lack of relationship between object imitation and language competence. Such a positive relationship was expected due to the inherently imitative properties of language and language development and the relationships shown in the previous chapter between action and language imitation.

#### **7.5.5. Limitations**

Our limitations for this Chapter reflect those set out earlier in the individual chapters. We had hoped for a score of between 0-8 for the final counterfactual conditional reasoning measure. However, the unconstrained nature of the open slide

task meant that in the end we yielded a score of 0-6. Furthermore having used two conditions in the stories task we removed chances for children to demonstrate their truly counterfactual reasoning abilities. Rafetseder et al. (2010) used a different paradigm in which counterfactual reasoning only could be used to correctly identify the correct state of affairs. Had we used this paradigm and perhaps generated complimentary scenarios to increase the number of test questions we could have generated a more comprehensive score of reasoning skill.

For the executive functioning measure it would have been preferential to also include a measure of working memory to give a fuller picture of how executive functions work alongside reasoning, action and language at this age. Further having measures that yielded ratio scores as opposed to pass/fail distinctions would have been highly beneficial and increased the scope of any analyses carried out.

For our mental state understanding measure the seeming independence of many of the measures, evidenced through the distinct patterns of relationship identified with other skills was problematic. The distinctive nature of the questions made the use of the battery as a scale difficult. Battery scores themselves were not as informative as the pass/fail information for each item. Further we did not include a classic false belief task that included deception such as the Sally/Ann task. The inclusion of such a prolifically reported task would have made our results more directly comparable to others.

It would have been more powerful for our subsequent analyses had the apparatus used in Chapter 5, the Tower and Rake been more directly comparable. The affordances of the objects were diverse enough for distinctly different behavioral types to be seen for both. Particularly in the necessary condition children were less likely to be temporally faithful imitators using the tower than the rake given that the

temporal sequence of action was critical for modeled goal success in one and not the other. Having had two apparatus that had identical causal sequence outcomes, whilst maintaining independent forms would have allowed for the data from both to be combined across conditions increasing the size and power of the data set.

The imitation of repetitions too could have been improved as discussed in the previous chapter. Levels of verbatim imitation were almost at ceiling level and as such more complex sentences and sentence structures could have been used to challenge the children and produce a greater spread of data points to work with. Further it would have been an interesting addition to look at how the causal implications of the language we use impacted upon children's imitation in a paradigm influenced by Bannard et al. (2013). Finally supporting the production of higher rates of novel verb use and passive verb constructions would have allowed for differences between participants to be identified more readily. In contrast to the imitation of repetitions the imitation of passive verb constructions were very low. We concluded that the decreased number of models we chose to use in comparison to the original study (Brooks & Tomasello, 1999) had induced this suppression of production. Had we seen more variance in verb production and syntax choices we may have been able to identify a richer story of how reasoning, imitation and language support one another.

Finally we chose to use a measure of linguistic competence derived from the elements of our own experimental procedures. This decision was based upon our desire to assess language in an interactive and relevant manner. It may have been beneficial to have further included a standardized linguistic measure such as the Peabody Picture Vocabulary Test-III (PPVT-III; Dunn & Dunn, 1997). Doing so would have added an additional layer of information about children's linguistic

competence and would have been free from any task specific demands placed upon the sample. Further it would have broadened the scope for comparison with other work.

#### **7.5.6. Conclusions**

Language competence as measured using our composite variable was positively predicted by counterfactual conditional reasoning. For executive functions too, those high on the Flanker task had significantly better scores than their peers in the low group. Finally the mental state understanding battery as whole did correlate with linguistic competence. Only participants who passed Q3 showed significantly higher language competence scores than their peers in the low group who failed. We then moved to look at how imitative behaviors in the action domain related to language scores. None of our action imitation behaviors showed any significant relationships with language competence.

We have provided evidence that children's ability to successfully utilize language as a tool is related to their ability to reason about the causal structure of the world more generally. Counterfactual conditional reasoning, inhibitory control and to a lesser extent mental state understanding were all positively associated with language competence. Acting about the world physically, in contrast to linguistically, however, showed no strong correlation. Children's actions on novel objects with a clearly implied causal structure did not relate to their language competence as formulated here.

When speaking about linguistic competence more generally, that is not only language repetition but language productivity, comprehension and discourse sensitivity there are strong links with reasoning about the world abstractly, but not so with acting upon physical objects in this context. Action imitation as measured

through our necessary and unnecessary conditions was not related to language competence more generally as derived from performance on our experimental tasks.

## **8. Navigating the Boundary Between the Actual and the Possible at Age Four:**

### **Conclusions**

#### **8.1. Summary of Results**

The explorative theme for this thesis was centered around investigating and understanding how children navigate the boundary between the actual world and other possible worlds at age four. These possible worlds may be different by virtue of being semantically incongruent with the actual world, being temporally distinct from the actual world or, being represented differently in the mind of another. We wanted to identify whether children's proficiency, or lack thereof, in reasoning about such actual versus possible worlds related to their use of actions and utterances as a tool for manipulating the world. Along the way we have reported specific details about children's counterfactual reasoning, executive functioning, mental state understanding, imitation of action and language and, productive and receptive language skills.

Firstly in Chapter 2 we looked at counterfactual reasoning ability at age four as measured by two structurally and conceptually contrasting tasks. Children were very good at reasoning about a counterfactual alternative when; the number of counterfactual possibilities are constrained, when causal structure is simple and, the question being asked is unambiguous. As these parameters were extended, however, children's reasoning became increasingly poor. When additional counterfactual options were introduced into a counterfactual world where the causal structure was not concretely observable, children's reasoning performance was significantly reduced. Despite children having greater difficulty predicting the counterfactual

location of a protagonist, when compared to identifying the correct counterfactual location of a ball, performance on these two tasks was positively statistically related.

Performance on two Executive Functioning tasks, the Flanker and the DCCS, were reported in Chapter 3. Children were better able to inhibit incongruent stimuli than they were at moving flexibly between differing object sorting rules. We chose to include EF for two distinct reasons. Firstly EFs are domain general cognitive processes required for all effortful, purposeful and goal directed actions (e.g. Carlson, 2005; Gioia, Isquith & Guy, 2001). As such we wanted to eliminate any relationships between tasks being solely a result of shared EF involvement. Secondly EF tasks themselves challenge children in much the same way as CFR and MSU. EF tasks ask children to apply some previously imposed rule to a current situation in order to select the ‘correct’ possible response. Children must respond to the actual world and select the correct possible world to be successful.

In Chapter 4 then we presented our final cognitive reasoning measure, mental state understanding. A five-item battery was used to generate a MSU score for each participant (Wellman & Lui, 2004). Mental state reasoning, as with CFR requires children to reasoning about a world different to their own. MS worlds are different by virtue of psychological perspective whilst CF worlds are different due to them once, but no longer, being possible. Children performed as expected on the battery with items becoming increasingly difficult along the scale.

Performance between our three cognitive measures was then investigated. CFR, EF and MSU are all forms of conditional reasoning where past information must be integrated with ones pre-existing knowledge in order to respond appropriately. As such we expected strong relationships between all three variables. There was a trend for children with higher EF scores to perform better on both the



CFR and MSU tasks, however, there was not the initial strikingly relationship we had anticipated. As such it was concluded that EF was necessary for, but not sufficient to support CFR and MSU. In addition we did not find any observable linear relationship between CFR and MSU as we measured them. This was surprising as both tasks challenge children to select a response less salient and personal than their own. Furthermore CFR and MSU questions both ask children to reason not in line with reality as they actual experience it, but in line with the causal structure of another possible reality. Instead we saw that generally MSU questions were answered more successfully than conditional reasoning questions, particularly those that challenged CFR processes uniquely. It was concluded then that perhaps the difference in content knowledge, physical causal understanding versus mental state understanding can alone account for such divergence in performance. Whilst both rely on conditional reasoning CFR also requires a disengagement from temporal reality whilst MSU requires a disengagement from egocentric reality, children appeared to find this egotistical disengagement significantly easier than the temporal alternative. Reasons as to why this may be the case were discussed further in Chapter 4.

In Chapter 5 we moved away from reporting how children answer questions about the world and towards how they choose to impact upon it. Children were highly imitative of an adult model performing two actions on a novel object. Over imitation was the preferred strategy regardless of causal necessity or goal salience. Much like as was the case with CFR and MSU, EF showed some positive relationships with this over-imitation phenomenon but was not predictive of it at the group level. The domain general processes we reported alone do not account for faithful imitation. CFR too was related to imitative performance with positive correlations being found between faithful imitation performance using both of our experimental apparatus. In

addition negative correlations were found between single action responses on the Tower and CFR. The temporal order of the two action responses across apparatus was further related to CFR performance as children who replicated but reversed actions in the unnecessary condition had lower CF story scores. Altogether then this is striking evidence that faithful imitation is driven by causal knowledge with children who reject socially implied causal structure also being poorer at reasoning about causal structure more generally in a counterfactual way.

Chapter 5 also investigated any link between MSU and imitation. Only one of the battery items, and not the battery as a whole related to imitative performance. Item 5 target-question performance alone correlated with faithful imitation in the critical unnecessary condition. No children answered the control memory questions of this item correctly, however, making this link additionally surprising. Reasons as to why this item alone may be significantly related were explored in detail in Chapter 5. Item 5 was unique in contrasting two different faces of the same mental state in a way much more cognitively akin to the contrasting realities of an actual and counterfactual world. This may have driven the relationships between these two items. MSU, CFR and Imitation then, despite their strikingly similar structural demands had relatively independent performance patterns for the most part. Relationships although evident are diverse enough to conclude that task specific demands play a significant role in performance.

In Chapter 6 we extended our exploration into imitation through the inclusion of the language as a to be imitated behavior. Three distinct features of language imitation were investigated; grammatical vs., ungrammatical repetitions, novel verb usage and, transitive structures. Like with action imitation children preferentially imitated repetitions faithfully, be they grammatical or ungrammatical. Furthermore as

with action imitation, verbal imitation was lower in the ungrammatical condition. This mirrored the necessary/unnecessary action imitation condition differences. Children, however, were reluctant to imitate newly introduced verbs and less familiar transitive verb structures. Children often reverted to English verbs, rather than novel verbs. Further they often reverted to the more familiar active transitive construction despite having heard numerous models of the passive construction. Their comprehension of passive sentences was also much lower than their comprehension of active sentences. There was evidence of discourse sensitivity, however, with children attempting to answer questions with a clear attentional focus in the congruent transitive structure.

Within our language imitation tasks then there were numerous relationships between imitation patterns. Children who retained a transitive structure in their own utterances (regardless of active or passive training) were more likely to repeat grammatical repetitions verbatim. Furthermore novel passive production, although relatively infrequent, was related to ungrammatical repetition imitation. Novel verb use following passive training was also uniquely related to repetition imitation in the ungrammatical condition.

When the imitation of actions and language were looked at together there were clear similarities between children's imitative proclivity. Patterns of imitation, regardless of the medium being observed, were not consistent enough to conclude that if children imitate one item, they shall imitate another. Instead a more complex inter-task picture was observed. This was in part due to the differing causal configurations of our two imitation apparatus. When each was looked at separately, however, there were several interesting findings. The temporal order of actions following a necessary condition tower model related to the removal of repetitions in our first language task. Those children who altered modeled actions also altered modeled speech. Using the

rake apparatus children who reversed responses in the necessary condition were more likely to add repetitions in the ungrammatical condition. Although this pattern is strikingly different it further suggests that instrumental goals are imitated faithfully while social goals may be more often altered.

When looking at the imitation of novel verbs further evidence of cross-domain relationships were unearthed. Novel verb use correlated negatively with single action, selective imitation responses using the tower following an unnecessary model. Children who rejected the modeled causal structure of tower and imitated only the modeler's instrumental goal were less likely to utilize the newly introduced verb. Finally we looked at children's use of transitive structures during the experimental task and found that temporally faithful response patterns in the action domain were positively correlated with the use of a transitive structure. Children who adhered to the modeled action structure also retained a transitive construction when answering questions about a verb action sequence.

This combined evidence provides support for a theory of imitation in which there are some cross task consistencies, despite obviously contrasting task demands. These task demands come on the form of both the medium of input and output, however, there are strikingly similarities across our four conditions.

Finally in Chapter 7 we compiled a language competence score to combine elements of comprehension, production and flexibility in performance. This competence score was designed to generate a proxy for children's skill at using language as a tool rather than simply a to-be imitated feature of behavior. Competence scores were positively associated with counterfactual conditional reasoning, inhibitory control and certain distinct elements of MSU. Language competence, above and beyond the ability to repeat what is heard, is related to how

well four year-olds reason about the world abstractly. Language competence is related to their ability to apply causal rules, to reason about events removed in time and events as observed and experienced by others. How children interact with the world through the use of language as a tool is directly related to their conditional reasoning about not only the actual world, but also other possible divergent worlds. How well children navigate the actual and possible then is directly related to how proficiently they can use language as a tool.

## **8.2. General Discussion**

### **8.2.1. Reasoning at Age Four**

As humans navigate the world we build a representation of reality and causality based on our inferences about the events and behaviors we are exposed to. As we act, both physically and verbally we alter this reality and often times these selected actions and utterances are grounded in past experiences (either personal or vicarious) that help us predict that they will achieve the desired outcome from a host of possibilities. It is this limitless and flexible number of possible world decisions that make human cognition remarkable and unique.

When we engage in counterfactual reasoning the perception of this reality can be changed through the addition of positive or negative affect resultant from said musings. We may also engage in counterfactual reasoning to plan the behaviors required to reach a positive future reality, often prompted by goal failure on the first attempt. Our sample provided ample evidence that children's conditional and counterfactual reasoning still has a long way to come. When comparing performance on two different tasks designed to assess children's counterfactual reasoning we showed that questions that can be answered using basic conditional processes are significantly easier than those requiring counterfactual reasoning exclusively. As the

demands on the reasoner increased, with the addition of multiple possibilities and a less concrete causal structure, performance decreased. At age four children are struggling to correctly integrate all the necessary causal information of a presented situation, and therefore fail to consistently select the logical counterfactual outcome from a set of possibilities.

Executive functioning tasks measure how well children process and deal with information free from context and the influence of prior knowledge. They, on a very basic level, challenge children to select a possible world response based on some feature of the actual world. They can give us an idea of how well children choose a possible response that contrasts with the actual in some way. In our sample there was a great diversity in the performance of children on tasks designed to assess their inhibitory control and cognitive flexibility. Children are not yet entirely proficient in rejecting one possible response in favor of another.

Mental state understanding further allows us to gather information about the world to aid our own decisions. Observing and making inferences based upon how another person's mental state changes in response to some stimuli or situation embellishes our causal knowledge for future use. At age four children were highly competent in MSU questions related to diverse desires and beliefs but with the addition of more complex items related to knowledge access, expectations and real versus apparent emotions children in this age group began to struggle. Children's conditional and mental state reasoning is clearly still developing and we were interested in seeing whether these developments in reasoning about the world were impacting upon concurrent developments in impacting upon it.

### **8.2.2. Reasoning Skill versus Action on the World**

Our data set is the first to combine executive functions, counterfactual conditional reasoning and mental state understanding with imitation. It was clear that children who were better at reasoning conditionally and counterfactually were generally and consistently performing to a higher level in other tasks. Conditional and counterfactual reasoning success was reflected in increased faithful imitation of actions and in particular the replication of the temporal order of modeled actions on objects. We showed that the children in our sample who reasoned in line with the causal structures presented during CFR tasks were not only more likely to replicate all the actions modeled by an experimenter, but they did so in the same temporal order. Children who reversed the temporal order of actions had lower CFR scores than their peers, as did children who selected a single action response as opposed to imitating faithfully.

Similar links were found in the verbal domain as the same pattern was seen between CFR and language competence more generally. These results suggest that being able to successfully incorporate information about the actual world, into a representation of the possible world that supports both inferences and behaviors is a pervasive ability spanning both counterfactual reasoning tasks and linguistic proficiency. Having good counterfactual reasoning skills support the comprehension and incorporation of novel lexical items into representations that support said items' replication in a semantically appropriate way. Finally children's sensitivity to discourse pressure was uniquely related to our simplest conditional reasoning task, the Slide task. We have shown that reasoning counter to fact is related to language comprehension and production, in addition to being related to complex social

language skills such as discourse sensitivity. How well children reason counter to fact is positively related to their use of language as a tool for change in the world.

In our sample we found that executive functions were related to action imitation and the imitation of language. We reported that children with lower inhibitory control were less likely to faithfully imitate a model's necessary and intentional action. Furthermore, we showed that participants who performed well on EF tasks recorded significantly better comprehension scores for both active and passive verb constructions. We therefore showed EFs to be influential in numerous facets of imitation and language.

Finally, to a lesser extent MSU also showed similar relations to CFR and EF. We reported evidence that MSU is involved in imitation in that children who answered the target questions only correctly for item 5 of the MSU battery (the real-apparent emotion question) had significantly higher rates of faithful imitation in the unnecessary condition. Here it is clear that how children interpret the 'possible world' that is the world seen by others, impacts upon how they choose to move forward. Furthermore children who answered item 3 of the battery correctly had significantly higher language competence scores. Although the scale as a whole did not relate to our measures of tool use as expected it was clear that certain aspects of mental state understanding do indeed mediate how children interpret the actual world and inform their choices when acting upon it.

### **8.2.3. Imitation Across Actions and Utterances**

Children copy behaviors to induce physical outcomes and they copy behaviors to induce psychological outcomes for both themselves and others. Children begin copying actions and vocalizations in selective and concrete ways but with time they begin to copy many behavioral features, both intentional and incidental. This move



may initially emerge due to social reasons (Hilbrink et al., 2013) or alongside developments in their navigation of the actual versus the possible, but is a prevalent and highly robust strategy used by four-year-olds in our sample and many others (e.g., McGuigan et al., 2007; 2011; Lyons et al., 2007; Nielsen, 2006). Often proclivity of imitation is related to the necessity of the behavior. Concurrently the complexity or familiarity of behavior impacts upon imitation. In our action imitation chapter it was evident that faithful imitation is a highly pervasive phenomenon. Causal necessity, although influential, did not remove this propensity.

Further when looking at verbal imitation some features of language were more readily imitated than others, regardless of their causal or grammatical properties. Children readily imitated ungrammatical repetitions, although less so than grammatical ones. Sometimes they even added repetitions to ungrammatical utterances in a ‘hyper-extension of imitation’. Children less frequently copied simple novel words, but the majority did so at least once. Children even less readily imitated an unfamiliar syntactic structure with the majority not doing so at all. With age children and adults do begin to robustly imitate novel words and less frequent syntactic forms as their ability to manipulate and integrate linguistic rules and tokens increases. They also more readily imitate more complex items if task demands are reduced or exposure is increased.

When relationships between these two distinct imitative domains were examined there was definitive evidence for some contiguity. In both action and language imitation tasks children readily imitated intentional features of the model that went above and beyond those required for an instrumental or communicative goal. There were correlations between the rejection of a modeled sequence in physical and verbal tasks. Although certainly not the case that children’s imitativity was stable

across all tasks there were tantalizing glimpses of contiguity that suggest imitation is a strategy not confined to either/or behavioral domain, but that can be flexibly deployed across domain and is influenced by causal knowledge, executive functions and intention understanding. There were some shared features of action and verbal imitation identified including children more readily imitating features of a modeled sequence that were instrumentally or communicatively conventional than otherwise. Further children in both domains clearly imitated intentional actions that were not necessary or correct for instrumental goal completion. Finally in both genres children showed a willing-ness to copy actions that were incidental or non-critical to the primary goal.

The message we present here is clear, children's reasoning skill related to navigating various forms of actual versus possible world boundaries are instrumental in their subsequent actions upon the world moving them into the realms of the possible. In conclusion, children's skill at reasoning about a reality temporally distinct from their own, or psychologically distinct from their own is directly reflected in their manipulations of the world as it is in pursuit of some representation of a future possible goal state.

#### **8.2.4. Limitations**

In Chapter 2 it was hoped that the inclusion of two forms of the slide task would add to an embellished our determination of their basic conditional reasoning skill related to a concrete, readily observable causal structure. However, the open slide task was not constrained enough in that children could theoretically respond with a location anywhere other than the realist response. As such the breadth of conditional reasoning tasks was diminished and the subsequent variance in conditional reasoning scores was decreased. On reflection the inclusion of the typical

and atypical stories were superfluous to our most important aim: characterizing the counterfactual reasoning ability of our sample. It would have been preferable to utilize the Sweet story example given in Chapter 2 (Rafetseder, Cristi-Vargas & Perner, 2010). The sweets paradigm focused solely on teasing apart CFR vs., BCR and did not include any between congruency manipulations. Using this structure of paradigm would have afforded us a more comprehensive picture of true CFR reasoning ability and given the results more clarity.

In Chapter 3 what we had really aimed for was a continuous measure of executive functions. This would have given us much clearer results and afforded us a more utilizable variable when identifying relationships with other variables. Furthermore, given the mix of results pertaining to EF and CFR in particular we would ideally have given children a working memory task in order to have data on the three main pillars of EF.

In Chapter 4 there were two main concerns about the battery used. Firstly Question 5 seemed far removed from the other, both conceptually and given the low number of correct memory check responses. Additionally, to have given our study greater comparability to other contemporary researchers a traditional false belief task could have been added. Although Item 4 the unexpected contents task did in some ways tap children's reasoning about a false belief the addition of a deceptive element could have proved crucial in unifying our results. The battery items often related to different elements of the other variables and as such did not act like the true continuous measure of mental state understanding that we had striven for. Perhaps we could have used a more interactive paradigm. One in which the child played an integral role in the experimental procedure and as such truly had to contrast their perceptions of an event to another's, i.e., the experimenter or a confederate.

When looking at Chapter 5 we saw that in a similar way to the typical and atypical story conditions contrasting experimentally so too did the necessary and unnecessary conditions of the action imitation apparatus. Rather than having two, similar but fundamentally different models it would have been better to have two diverse, but structurally comparable ones to increase the amount of comparable data collected between them and bolster the assertions drawn from the results.

Another consideration was that the tasks, or apparatus themselves were too easy. Had they been left alone children would have presumably carried out the same actions given a few opportunities to explore the object independently. As such there was very little variation in performance with children completing two actions in the majority of cases. Have a more complex, causally challenging structure inherent in the testing apparatus could have yielded more diverse results to feed forward. This formulation may prove more useful when looking for a link with language competence. Language competence reflects children's skill at both being imitative with language but also innovative in response to situational and social pressures. The inclusion of a more complex action imitation task, or a task where some action innovation is required alongside imitation may be more fruitful in helping to understanding the link between imitation as a strategy more generally and language competence in an ecologically valid way.

In Chapter 6 the verbal imitation task again yielded very high verbatim imitation rates. The sentences were short, and despite containing repetitions very simple. Longer sentences, sentences with a more challenging structure or even sentences encapsulating less familiar or more physically challenging lexical items could have been used to generate more diverse results. The passive verb paradigm had the opposite problem in that production rates for the target utterance structure were

low. In the original paper using this methodology (Brooks & Tomasello, 1999) children were given 88 models, this is compared to our 16. Perhaps having even doubled the number of models would have supported children enough to increase production rates.

### **8.2.5. Conclusions**

We have presented the first in depth investigation into how children aged four years navigate the boundary between the actual and the possible through a variety of experimental tasks. To our knowledge no other projects have attempted to unify the underlying global similarities between navigating the actual and the possible through diverse reasoning tasks and realizing the move from the actual to the possible through physical or linguistic manipulations on the world. We uniquely within one sample collected performance measures from four cognitive tasks we had identified as challenging children to reasoning about possible based on the actual world: counterfactual reasoning, the Flanker task, the DCCS task and a mental state understanding battery. Further more we concurrently investigated performance on these tasks alongside performances in imitative tasks where children could carry out the process of navigating the actual and the possible through their own actions: an action imitation paradigm, a grammatical versus ungrammatical repetition paradigm and a novel verb learning task. In addition we related children's navigation of the actual and the possible, both through reasoning and behavior to their competence in using language as a tool for change in the world.

Our aim was to first identify any correlations between cognitive reasoning tasks that challenge children to select a response and make inferences based on a possible world that contrasted from the actual world in some way. For counterfactual reasoning tasks the possible world adhered to all the same conditional rules as the

actual world but was different due to some counterfactual antecedent: here contrast between the actual and the possible included a temporal distinction. For the executive functioning tasks the target questions challenged children to select an incongruent response that contrasted with their pre-potent response in some way. Finally mental state understanding tasks the possible world differed by virtue of the perspective of another intentional agent: their thoughts, beliefs, desires and access to knowledge contrasted with the reasoners'. There were indeed correlations between CFR and EF, and EF and MSU but not a bold link between CFR and MSU despite the striking structural similarities between the task demands.

We have shown that all three cognitive reasoning measures were in some way implicated in the imitation of actions and utterances in our sample. How well children understand the actual versus possible boundaries in the world indeed impacts upon their attempts to manipulate it. We have reported evidence that performance on measures of counterfactual conditional reasoning, executive functioning and mental state understanding all uniquely relate to the way children use imitation as a strategy to impact upon the world. Better performance on these three tasks cognitive tasks relates to higher levels of faithful imitation, more adherence to the temporal qualities of imitation and even in some circumstances a hyper-extension of imitation. How children think about the world, and the contrasts between actual and possible worlds, impacts upon their own actions moving them from the actual to the possible.

Furthermore we reported that features of our samples' performance on counterfactual reasoning, executive functioning, mental state understanding and language imitation tasks were related to language competence more generally. How well children use language as a tool for change in the world is related to how well they can reason about actual versus possible distinctions within it.

We chose to take a global view of numerous cognitive development indices together rather than selecting several of these tasks in isolation. We did so because we wanted to understand whether the conceptual similarities between the tasks chosen, all requiring children's navigation of the actual and the possible, was indeed a powerful underlying feature unifying performance on them at age four. Studies have in the past related performance on some of the tasks we administered, particularly in the case of executive functions and language. None, however, have taken such a broad view in an attempt to identify an underlying supporting ability in tasks which on the surface look quite different: the ability to flexibly navigate the boundary of what actually occurred, what could have occurred and crucially what could yet occur in the world.

#### **8.2.6. Future Directions**

We choose to look at one age snap-shot of how well children navigate between the actual and the possible. For us, the next logical step would be to assess these same abilities longitudinal in a single sample of typically developing children. For all of the abilities were characterized in this thesis there are comparable paradigms and methods for both younger and older samples. Being able to observe the emergence and development of conditional and counterfactual reasoning, executive functions and mental state understanding and relating it to concurrent tool use, both physically and linguistically would provide an invaluable insight into how these relationships may coalesce. Indeed within the sample presented throughout this thesis were children who had been involved in data collection from birth. A range of data was collected every month from birth to 18-months, and again at 24-months relating to children's imitative proclivity, cognitive development, motor development, temperament and other socio-cognitive/attentional abilities. In addition maternal-

infant free-play interaction data and parent diary data were recorded (see Ellis-Davies, Fowler, Hilbrink, Sakkalou & Gattis, 2012 for more information on this CUE diary method). As such in future we would hope to look retrospectively within this portion of our data set to understand if tasks relating to this actual versus possible navigation in the pre-school years predicted later reasoning and tool use performance. Given that reasoning flexibly and in a way that integrates what is actual and what is possible is central to adult human cognition, being able to observe the emergence and development of this skill will, we believe, be an exciting and invaluable endeavor.



## References

- Abbot-Smith, K., Lieven, E., & Tomasello, M. (2004). Training 2; 6-year-olds to produce the transitive construction: the role of frequency, semantic similarity and shared syntactic distribution. *Developmental Science*, 7(1), 48-55.
- Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development*, 67, 635–645.
- Akhtar, N., & Tomasello, M. (1996). Two-year-olds learn words for absent objects and actions. *British Journal of Developmental Psychology*, 14, 79 – 93.
- Alexander, M. P., & Stuss, D. T. (2000). Disorders of frontal lobe functioning. *Seminars in Neurology*, 20(4), 427-438.
- Ali, N., Schlottmann, A., Shaw, A., Chater, N., & Oaksford, M. (2010). Causal discounting and conditional reasoning in children. *Cognition and conditionals: Probability and logic in human thought*, 117-134.
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child neuropsychology*, 8(2), 71-82.
- Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Mikiewicz, O. (2002). Relationships between cognitive and behavioral measures of executive function in children with brain disease. *Child Neuropsychology*, 8(4), 231-240.
- Aschersleben, G., Hofer, T., & Jovanovic, B. (2008). The link between infant attention to goal-directed action and later theory of mind abilities. *Developmental Science*, 11(6), 862-868.
- 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.

- Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, 20(2), 395 – 418.
- Baldwin, D. A. (2000). Interpersonal understanding fuels knowledge acquisition. *Psychological Science*, 9, 40–45.
- Baldwin, D. A., & Baird, J. A. (1999). Action analysis: A gateway to intentional inference. In Rochat, P. (Ed). *Early social cognition: Understanding others in the first months of life*. Mahwah: Lawrence Erlbaum Associates.
- Bandura, A. (1971). *Social learning theory*. Morristown.
- Bannard, C., Klinger, J., & Tomasello, M. (2013). How selective are 3-year-olds in imitating novel linguistic material? *Developmental psychology*, 49(12), 2344-2356.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, 21(1), 37-46.
- Baron-Cohen, S., Ring, H., Moriarty, J., Schmitz, B., Costa, D., & Ell, P. (1994). Recognition of mental state terms. Clinical findings in children with autism and a functional neuroimaging study of normal adults. *The British Journal of Psychiatry*, 165(5), 640-649.
- Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar: individual differences and dissociable mechanisms*. Cambridge, UK: Cambridge University Press.
- Beaumont, J. L., Havlik, R., Cook, K. F., Hays, R. D., Wallner-Allen, K., Korper, S. P., ... & Gershon, R. (2013). Norming plans for the NIH Toolbox. *Neurology*, 80(11:3), S87-S92.

- Beck, S. R., & Crilly, M. (2009). Is understanding regret dependent on developments in counterfactual thinking? *British Journal of Developmental Psychology*, 27, 505–510.
- Beck & Guthrie, (2011). Almost Thinking Counterfactually: Children's Understanding of Close Counterfactuals. *Child Development*, 82(4), 1189-1198.
- Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's thinking about counterfactuals and hypotheticals as possibilities. *Child Development*, 77, 413-426.
- Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: infants' understanding of intentional action. *Developmental psychology*, 41(2), 328-337.
- Behrend, D. A. (1989). Default values in verb frames: cognitive biases for learning verb meanings. *Proceedings of the 11th-annual meeting of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Behrend, D. A. (1990). The development of verb concepts: children's use of verbs to label familiar and novel events. *Child Development*, 61, 681 – 696.
- Behrend, D. A., Harris, L. L., & Cartwright, K. B. (1995). Morphological cues to verb meaning: verb inflections and the initial mapping of verb meanings. *Journal of Child Language*, 22, 89-106.
- Behrend, D. A., & Scofield, J. (2006). Verbs, actions, and intentions. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. New York: Oxford University Press.

- Bekkering, H., Wohlschlagel, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *The Quarterly Journal of Experimental Psychology: Section A*, 53(1), 153-164.
- Bencini, G. M., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, 59(1), 97-113.
- Benson, J. E., Sabbagh, M. A., Carlson, S. M., & Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Developmental Psychology*, 49(9), 1615-1627.
- Bever, T. G. (1970). The cognitive basis for linguistic structures. In Hayes, R. (Ed). *Cognition and Language Development*. New York: Wiley & Sons.
- Bertenthal, B., & Von Hofsten, C. (1998). Eye, head and trunk control: the foundation for manual development. *Neuroscience & Biobehavioral Reviews*, 22(4), 515-520.
- Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, 18(5), 382-386.
- Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, 78(2), 647-663.
- Bloom, L. (1978). The semantics of verbs in child language. Paper presented at the meeting of the Eastern Psychological Association, New York.
- Bloom, L. (1993). *The transition from infancy to language: Acquiring the power of expression*. New York: Cambridge University Press.

- Bloom, L., Tinker, E., & Margulis, C. (1993). The words children learn: evidence against a noun bias in early vocabularies. *Cognitive Development*, 8, 431 – 450.
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology*, 129(2), 177.
- Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S. Y., Pascual, L., ... & Vyt, A. (2004). Cross-Linguistic Analysis of Vocabulary in Young Children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, 75(4), 1115-1139.
- Boyland, J.T., & Anderson, J.R. (1998). Evidence that syntactic priming is long-lasting. In M. A. Gernsbacher & S. J. Derry. (Eds). *Proceedings of the 20th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum Associates.
- Bowerman, M. (1974). Learning the structure of causative verbs: A study in the relationship of cognitive, semantic, and syntactic development. *Papers and Reports on Child Language Development*, 8, 142 – 178.
- Brandone, A. C., Pence, K. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Action speaks louder than words: Young children differentially weight perceptual, social, and linguistic cues to learn verbs. *Child Development*, 78(4), 1322-1342.
- Branigan, H. P., Pickering, M. J., McLean, J. F., & Cleland, A. A. (2007). Syntactic alignment and participant role in dialogue. *Cognition*, 104(2), 163-197.

- Bretherton, I., McNew, S., & Beeghley-Smith, M. (1987). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a “theory of mind”? *Cognitive development in infancy*, 219-246.
- Brooks, P. J., & Braine, M. D. (1996). What do children know about the universal quantifiers all and each? *Cognition*, 60(3), 235-268.
- Brooks, P. J., & Tomasello, M. (1999). Young children learn to produce passives with nonce verbs. *Developmental Psychology*, 35(1), 29.
- Brooks, P. J., Tomasello, M., Dodson, K., & Lewis, L. B. (1999). Young children's overgeneralizations with fixed transitivity verbs. *Child Development*, 70(6), 1325-1337.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. *Child Development*, 67(3), 836-849.
- Brubacher, S. P., Roberts, K. P., & Obhi, S. S. (2013). Gaze, goals and growing up: Effects on imitative grasping. *British Journal of Developmental Psychology*, 31(3), 318-333.
- Brugger, A., Lariviere, L. A., Mumme, D. L., & Bushnell, E. W. (2007). Doing the right thing: Infants' selection of actions to imitate from observed event sequences. *Child Development*, 78(3), 806-824.
- Buchsbaum, D., Bridgers, S., Weisberg, D. S., & Gopnik, A. (2012). The power of possibility: causal learning, counterfactual reasoning and pretend play. *Philosophical Transactions of the Royal Society B*, 367, 2202-2212.

- Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition*, 120(3), 331-340.
- Burns, P., Riggs, K. J., & Beck, S. R. (2012). Executive control and the experience of regret. *Journal of Experimental Child Psychology*, 11, 501-515.
- Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2013). Chimpanzees, < i> Pan troglodytes</i>, recognize successful actions, but fail to imitate them. *Animal Behaviour*, 86(4), 755-761.
- 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.
- Byrne, R. M. J. (2005). *The rational imagination: How people create alternatives to reality*. Cambridge, MA: MIT Press.
- Call, J., & Tomasello, M. (1999). A nonverbal false belief task: The performance of children and great apes. *Child Development*, 70(2), 381-395.
- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, 28(2), 595-616.
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72(4), 1032-1053.
- Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, 69(3), 672-691.
- Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, 21, 315 – 330.

- Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science*, 8(1), 13-20.
- Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the society for research in child development*, i-174.
- Carrow-Woolfolk, E. (1999). *Test for Auditory Comprehension of Language: TACL-3*. Pro-Ed.
- Cazden, C. B. (1968). The acquisition of noun and verb inflections. *Child Development*, 433-448.
- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234-272.
- 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.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of personality and social psychology*, 76(6), 893.
- Chelune, G. J., & Baer, R. A. (1986). Developmental norms for the Wisconsin Card Sorting test. *Journal of clinical and experimental neuropsychology*, 8(3), 219-228.
- Chen, M. L., & Waxman, S. R. (2013). "Shall we blick?": Novel words highlight actors' underlying intentions for 14-month-old infants. *Developmental Psychology*, 49(3), 426-431.



- Chiarella, S. S., Kristen, S., Poulin-Dubois, D., & Sodian, B. (2013). Concurrent Relations Between Perspective-Taking Skills, Desire Understanding, and Internal-State Vocabulary. *Journal of Cognition and Development*, 14(3), 480-498.
- Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, 38, 967-978.
- Childers, J. B., & Tomasello, M. (2006). Are nouns easier to learn than verbs? Three experimental studies. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*, New York: Oxford University Press.
- Chomsky, N. (2002). *On nature and language*. Cambridge University Press.
- Clark, R. (1977). What's the use of imitation? *Journal of Child Language*, 4(03), 341-358.
- Clark, E. (1978). Discovering what words can do. In Farkas, D., Jacobsen, W., & Todrys, K. (Ed.), *Papers from the parasession on the lexicon*. Chicago: Chicago Linguistic Society.
- Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9(4), 377-395.
- Colonesi, C., Rieffe, C., Koops, W., & Perucchini, P. (2008). Precursors of a theory of mind: A longitudinal study. *British Journal of Developmental Psychology*, 26(4), 561-577.
- DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in Cognitive Science*, 8, 66-70.

- De Villiers, J. G., & Pyers, J. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief understanding. *Cognitive Development*, 17, 1037–1060.
- Diamond, A. (2013). Executive functions. *Annual review of psychology*, 64, 135-168.
- Diamond, A. (2005). Attention-deficit disorder (attention-deficit/hyperactivity disorder without hyperactivity): A neurobiologically and behaviorally distinct disorder from attention-deficit/hyperactivity disorder (with hyperactivity). *Development and psychopathology*, 17(03), 807-825.
- Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to “Do as I say, not as I do”. *Developmental psychobiology*, 29(4), 315-334.
- Dias, M. G., & Harris, P. L. (1988). The effect of make-believe play on deductive reasoning. *British Journal of Developmental Psychology*, 6, 207–221.
- DiYanni, C., Nini, D., & Rheel, W. (2011). Looking good versus doing good: Which factors take precedence when children learn about new tools? *Journal of experimental child psychology*, 110(4), 575-591.
- Echols, C. H., & Marti, C. N. (2004). The identification of words and their meanings: From perceptual biases to language-specific cues. In D.G. Hall & S.R. Waxman. (Eds.), *Weaving a lexicon*. Cambridge, MA: MIT Press.
- Eckerman, C. O., Davis, C. C., & Didow, S. M. (1989). Toddlers' emerging ways of achieving social coordinations with a peer. *Child Development*, 440-453.
- Ellis-Davies, K., Sakkalou, E., Fowler, N. C., Hilbrink, E. E., & Gattis, M. (2012). CUE: The continuous unified electronic diary method. *Behavior research methods*, 44(4), 1063-1078.

- Elsner, B., & Pfeifer, C. (2012). Movement or goal: Goal salience and verbal cues affect preschoolers' imitation of action components. *Journal of experimental child psychology*, 112(3), 283-295.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, 16(1), 143-149.
- Fein, G. G. (1981). Pretend play in childhood: An integrative review. *Child Development*, 1095-1118.
- Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Applied Psycholinguistics*, 21(01), 95-116.
- Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human perception and performance*, 1(3), 288-299.
- Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science*, 5, 55-64.
- Fisher, C., Hall, G. D., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. In L. Gleitman & B. Landau (Eds.), *The acquisition of the lexicon*. Cambridge, MA: MIT Press.
- Flavell, J. H. (2000). Development of children's knowledge about the mental world. *International journal of behavioral development*, 24(1), 15-23.
- Flavell, J. H., Beach, D. R., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development*, 283-299.

- Flobbe, L., Verbrugge, R., Hendriks, P., & Krämer, I. (2008). Children's application of theory of mind in reasoning and language. *Journal of Logic, Language and Information*, 17(4), 417-442.
- Forbes, J. N., & Farrar, M. J. (1993). Children's initial assumptions about the meaning of novel motion verbs: Biased and conservative? *Cognitive Development*, 8, 272-290.
- Forbes, J. N., & Farrar, M. J. (1995). Learning to represent word meaning: Initial training events affect children's developing action verb concepts. *Cognitive Development*, 10, 1- 20.
- Forbes, J. N., & Poulin-Dubois, D. (1997). Representational change in young children's understanding of familiar verb meaning. *Journal of Child Language*, 24, 389- 406.
- Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *Journal of verbal learning and verbal behavior*, 2(2), 121-135.
- Frosch, C. A., McCormack, T., Lagnado, D. A., & Burns, P. (2012). Are causal structure and intervention judgments inextricably linked? A developmental study. *Cognitive Science*, 36(2), 261-285.
- Frye, D., Zelazo, P. D., Brooks, P. J., & Samuels, M. C. (1996). Inference and action in early causal reasoning. *Developmental Psychology*, 32(1), 120-131.
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, 10(4), 483-527.
- Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: II. The development of executive function in childhood. *Current Directions in Psychological Science*, 121-126.

- Gelman, S. A., & Taylor, M. (1984). How two-year-old children interpret proper and common names for unfamiliar objects. *Child Development*, 55, 1535 – 1540.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Language, thought, and culture* (Vol. 2). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature*, 415(6873), 755-755.
- Gergely, G., Nadasdy, Z., & Csibra, G. (1995). Taking the intentional stance at 12 months of age. *Cognition*, 56, 165 – 193.
- Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: performance of children 3;2–7 years old on a stroop-like day-night test. *Cognition*, 53(2), 129-153.
- Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning Words and Rules Abstract Knowledge of Word Order in Early Sentence Comprehension. *Psychological Science*, 17(8), 684-691.
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.
- Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development*. New York: Oxford University Press.
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73, 135 – 176.
- Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, 1, 3 – 55.
- Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, 1, 23 – 64.

- Gleitman, L., & Gillette, J. (1995). The role of syntax in verb learning. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language*. London: Blackwell.
- Gleissner, B., Bekkering, H., & Meltzoff, A. N. (2000). Children's coding of human action: cognitive factors influencing imitation in 3-year-old. *Developmental Science*, 3(4), 405-414.
- Glymour, C. N. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. MIT press.
- Gioia, G. A., Isquith, P. K., & Guy, S. C. (2001). Assessment of executive functions in children with neurological impairment. *Psychological and developmental assessment: Children with disabilities and chronic conditions*, 317-356.
- Goldberg, A. E. (1995). *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M. & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14, 23 – 45.
- Golinkoff, R. M., Hirsh-Pasek, K., & Schweisguth, M. A. (2001). A reappraisal of young children's knowledge of grammatical morphemes. J. Weissenborn & B. Hoehle (Eds.), *Approaches to bootstrapping: Phonological, syntactic and neuro-physiological aspects of early language acquisition*. Amsterdam-Philadelphia: John Benjamin.
- Golinkoff, R. M., Jacquet, R. C., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. *Child Development*, 67, 3101 – 3119.

- Gopnik, A. & Meltzoff, A. N. (1986). Relations between semantic and cognitive development in the one-word stage: The specificity hypothesis. *Child Development, 57*, 1040 – 1053.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological review, 111*(1), 3.
- Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology, 37*(5), 620.
- Gordon, P., & Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition, 36*(3), 227-254.
- Gordon, A. C., & Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68*(1), 70-83.
- Grice, H. P. (1970). *Logic and conversation* (pp. 41-58). na.
- Guajardo, N. R., Parker, J., & Turley-Ames, K. (2009). Associations among false belief understanding, counterfactual reasoning, and executive function. *British Journal of Developmental Psychology, 27*(3), 681-702.
- Guajardo, N. R., & Turley-Ames, K. J. (2004). Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cognitive Development, 19*(1), 53-80.

- Guttentag, R., & Ferrell, J. (2004). Reality compared with its alternatives: Age differences in judgments of regret and relief. *Developmental Psychology*, 40, 764–775.
- Hagmayer, Y., Sloman, S. A., Lagnado, D. A., & Waldmann, M. R. (2007). Causal reasoning through intervention. *Causal learning: Psychology, philosophy, and computation*, 86-100.
- Hanna, E., & Meltzoff, A. N. (1993). Peer imitation by toddlers in laboratory, home, and day-care contexts: implications for social learning and memory. *Developmental Psychology*, 29(4), 701-710.
- Harris, M. (1976). The influence of reversibility and truncation on the interpretation of the passive voice by young children. *British Journal of Psychology*, 67(3), 419-427.
- Harris, F. N., & Flora, J. A. (1982). Children's use of Get passives. *Journal of Psycholinguistic Research*, 11(4), 297-311.
- Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, 61(3), 233-259.
- Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, 75(2), 27-39.
- Hartup, W. W., & Coates, B. (1970). The role of imitation in childhood socialization. Early experiences and the processes of socialization, 109-142.
- Harwood, F. W. (1959). Quantitative study of the speech of Australian children. *Language and Speech*, 2(4), 236-271.
- He, Z., Bolz, M., & Baillargeon, R. (2012). 2.5-year-olds succeed at a verbal anticipatory-looking false-belief task. *British Journal of Developmental Psychology*, 30(1), 14-29.



- Henning, A., Spinath, F. M., & Aschersleben, G. (2011). The link between preschoolers' executive function and theory of mind and the role of epistemic states. *Journal of experimental child psychology*, 108(3), 513-531.
- 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.
- Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience & Biobehavioral Reviews*, 34(4), 575-583.
- 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), 828-840.
- Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar*. Cambridge, MA: MIT Press.
- Hirsh-Pasek, K., & Golinkoff, R. M. (Eds.). (2006). *Action meets word: How children learn verbs*. New York: Oxford University Press.
- Hirsh-Pasek, K., Golinkoff, R. M., & Hollich, G. (2000). An emergentist coalition model for word learning. In R.M., Golinkoff et al. (Eds.), *Becoming a word learner: A debate on lexical acquisition*. New York: Oxford University Press.
- Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model of the origins of word learning. *Monographs of the Society for Research in Child Development*, 65 (3, Serial No. 262).
- Horgan, D. (1978). The development of the full passive. *Journal of Child Language*, 5(01), 65-80.

- Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, 16(2), 233-253.
- Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology*, 43(6), 1447-1459.
- Hume, D. (1978). A treatise of human nature [1739]. *British Moralists*, 1650-1800.
- Huttenlocher, J., Smiley, P., & Charney, R. (1983). Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review*, 90, 72 – 93.
- Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, 45(3), 337-374.
- Imai, M., Haryu, E., & Okada, H. (2002). Is verb learning easier than noun learning for Japanese children? 3-year- old Japanese children's knowledge about object names and action names. In B. Skarabela, S. Fish & A. H. A. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development*, Somerville, MA: Cascadilla Press.
- Imai, M., Haryu, E., Okada, H., Lianjing, L., & Shigematsu, J. (2006). Revisiting the Noun-Verb Debate: A Cross-Linguistic Comparison of Novel Noun and Verb Learning in English-, Japanese-, and Chinese-Speaking Children. In *Action meets word: How children learn verbs*. K. Hirsh-Pasek & R. M Golinkoff (Eds.). Oxford University Press.
- Imai, M., Haryu, Etsuko., Hirsh-Pasek, K., Li, Lianjing., Okada, Hiroyuki., Golinkoff, R.M., & Shigematsu, J. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese- speaking children. *Child Development*, 79(4), 979-1000.

- Jacques, S., Zelazo, P. D., Kirkham, N. Z., & Semcesen, T. K. (1999). Rule selection versus rule execution in preschoolers: an error-detection approach. *Developmental Psychology*, 35(3), 770.
- Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best preschoolers use past reliability over age when learning new words. *Psychological Science*, 17(9), 757-758.
- Jespersen, O. (1924). *The Philosophy of Grammar*. London: Holt
- Johnson-Laird, P. N., & Byrne, R. M. (2002). Conditionals: a theory of meaning, pragmatics, and inference. *Psychological Review*, 109(4), 646.
- Jusczyk, P. W., Cutler, A., & Redanz, N. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, 64, 675 – 687.
- Kaschak, M. P., & Glenberg, A. M. (2004). This construction needs learned. *Journal of Experimental Psychology: General*, 133(3), 450-467.
- Kersten, A.W., & Smith, L.B. (2002). Attention to novel objects during verb learning. *Child Development*, 73, 93–109.
- Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, 89(1), 25-41.
- Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, 48(1), 171.
- Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Developmental Science*, 14(1), 106-111.
- Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science*, 6(5), 449-467.

- Kray, L. J., Galinsky, A.D., & Wong, E.M. (2006). Thinking within the box: Relational processing styles elicited by counterfactual thinking. *Journal of Personality and Social Psychology*, 91(1), 33-48.
- Kristen, S., Thoermer, C., Hofer, T., Aschersleben, G., & Sodian, B. (2006). Skalierung von "Theory of Mind "-Aufgaben. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 38(4), 186-195.
- Kuczaj, S. A., & Maratsos, M. P. (1975). On the Acquisition of " Front, Back", and " Side". *Child Development*, 202-210.
- Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *The Journal of the Acoustical Society of America*, 100(4), 2425-2438.
- Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, 96, 1–33.
- Lampinen, J. M., & Smith, V. L. (1995). The incredible (and sometimes incredulous) child witness: Child eyewitnesses' sensitivity to source credibility cues. *Journal of Applied Psychology*, 80(5), 621-627.
- Lempert, H. (1990). Acquisition of passives: the role of patient animacy, salience, and lexical accessibility. *Journal of Child Language*, 17(03), 677-696.
- Lieven, E. V., Pine, J. M., & Barnes, H. D. (1992). Individual differences in early vocabulary development: Redefining the referential-expressive distinction. *Journal of Child Language*, 19, 287 – 310.
- Loeb, D. F., & Allen, G. D. (1993). Preschoolers' imitation of intonation contours. *Journal of Speech, Language, and Hearing Research*, 36(1), 4-13.

- Lovell, K., & Dixon, E. M. (1967). The growth of the control of grammar in imitation, comprehension, and production. *Journal of Child Psychology and Psychiatry*, 8(1), 31-39.
- Low, J. (2010). Preschoolers' Implicit and Explicit False-Belief Understanding: Relations With Complex Syntactical Mastery. *Child Development*, 81(2), 597-615.
- Luna, B. (2009). Developmental changes in cognitive control through adolescence. *Advances in child development and behavior*, 37, 233.
- Luria, A. R. (1976). *The working brain: An introduction to neuropsychology*. Basic Books.
- 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.
- Ma, W., Golinkoff, R., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2006). Imagine that! Imageability predicts verb learning in Chinese children. Unpublished manuscript, University of Delaware.
- Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science*, 11(4), 628-634.
- Maguire, M. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2006). A unified theory of word learning: Putting verb acquisition into context In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. New York: Oxford University Press.
- Maguire, M. J., Pruden, S. M., Hirsh-Pasek, K., Hansell, N., & Meyer, M. (2004). More isn't always better: Three-year-olds master intransitive verbs. *Poster*

*presented at the biennial meeting of the International Society on Infant Studies*, Chicago, IL.

- Maratsos, M., & Chalkley, M.A. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. Nelson (Ed.), *Children's language*. New York: Gardner Press.
- Maratsos, M., Kuczaj, S. A., Fox, D. E., & Chalkley, M. A. (1979). *Some empirical studies in the acquisition of transformational relations: Passives, negatives, and the past tense. Children's language and communication*. Hillsdale, NJ: Lawrence Erlbaum.
- Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of child language*, 21(02), 339-366.
- Marchman, V., Bates, E., Burkhardt, A., & Good, A. (1991). Functional constraints on the acquisition of the passive: Toward a model of the competence to perform. *First Language*, 11, 65-92.
- Markman, E. M. (1991). *Categorization and naming in children: Problems of induction*. MIT Press.
- Markman, K. D., Lindberg, M. J., Kray, L. J., & Galinsky, A. D. (2007). Implications of counterfactual structure for creative generation and analytical problem solving. *Personality and Social Psychology Bulletin*, 33(3), 312-324.
- Markovits, H. (1985). Incorrect conditional reasoning among adults: Competence or performance? *British Journal of Psychology*, 76(2), 241-247.
- Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review*, 22(1), 5-36.

- Martin, A., Onishi, K.H., & Vouloumanos, A. (2012). Understanding the abstract role of speech in communication at 12 months. *Cognition*, 123, 50-60.
- Mason, W. A., Jones, A., & Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, 137(3), 422.
- Matheson, H., Moore, C., & Akhtar, N. (2013). The development of social learning in interactive and observational contexts. *Journal of experimental child psychology*, 114(2), 161-172.
- McClure, K., Pine, J. M., & Lieven, E. V. (2006). Investigating the abstractness of children's early knowledge of argument structure. *Journal of Child Language*, 33(4), 693-720.
- McCrea, S. M. (2008). Self-handicapping, excuse making, and counterfactual thinking: consequences for self-esteem and future motivation. *Journal of Personality and Social Psychology*, 95(2), 274.
- McCune-Nicolich, L. (1981). Toward symbolic functioning: Structure of early pretend games and potential parallels with language. *Child Development*, 52, 785 – 595.
- McDonough, C., Lannon, B., Hirsh-Pasek, K., Golinkoff, R. M. (2006). The effect of imageability on early word learning. Unpublished manuscript, Temple University, Philadelphia, PA.
- McGuigan, N., Makinson, J., & Whiten, A. (2011). From overimitation to superimitation: Adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal of Psychology*, 102(1), 1-18.

- McGuigan, N., Whiten, A., Flynn, E., & Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3- and 5-year-old children. *Cognitive Development, 22*(3), 353-364.
- Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts of by 18-month-old children. *Developmental Psychology, 31*, 838 – 850.
- Meltzoff, A. N. (2007). ‘Like me’: a foundation for social cognition. *Developmental Science, 10*(1), 126-134.
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*(4312), 75-78.
- Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child development, 702-709*.
- Meltzoff, A. N., & Moore, M. K. (1989). Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Developmental psychology, 25*(6), 954.
- Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development & Parenting, 6*(3-4), 179.
- Messenger, K. (2009). *Syntactic priming and children’s production and representation of the passive*. PhD Thesis, University of Edinburgh.
- Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children’s passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language, 66*(4), 568-587.
- Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*(2), 622-646.



- Milner, B. (1964). Some effects of frontal lobectomy in man. *The frontal granular cortex and behavior*, 313-334.
- Mink, D., Henning, A., & Aschersleben, G. (2014). Infant shy temperament predicts preschoolers Theory of Mind. *Infant Behavior and Development*, 37(1), 66-75.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., ... & Caspi, A. (2011). *A gradient of childhood self-control predicts health, wealth, and public safety*. Proceedings of the National Academy of Sciences, 108(7), 2693-2698.
- Moore, C., & Frye, D. (1991). The acquisition and utility of theories of mind. *Children's theories of mind: Mental states and social understanding*, 1-14.
- Morton, J. B., & Munakata, Y. (2002). Are you listening? Exploring a developmental knowledge–action dissociation in a speech interpretation task. *Developmental Science*, 5(4), 435-440.
- Müller, U., Zelazo, P. D., & Imrisek, S. (2005). Executive function and children's understanding of false belief: How specific is the relation? *Cognitive Development*, 20(2), 173-189.
- Nagy, E., Pilling, K., Orvos, H., & Molnar, P. (2013). Imitation of tongue protrusion in human neonates: Specificity of the response in a large sample. *Developmental Psychology*, 49(9), 1628.
- Naigles, L. G., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child Development*, 64, 1665 – 1687.
- Naigles, L. G. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17, 357 – 374.
- Naigles, L. R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, 58, 221 – 251.

- Nathanson, A. I., Sharp, M. L., Aladé, F., Rasmussen, E. E., & Christy, K. (2013). The relation between television exposure and theory of mind among preschoolers. *Journal of Communication*, 63(6), 1088-1108.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 38 (149).
- Nelson, K., Hampson, J., & Shaw, L. (1993). Nouns in early lexicons: Evidence, explanations, and implications. *Journal of Child Language*, 20, 61 – 84.
- Nielsen, M. (2006). Copying actions and copying outcomes: social learning through the second year. *Developmental Psychology*, 42(3), 555.
- 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.
- 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.
- Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and social psychology review*, 9(3), 184-211.
- Olineck, K. M., & Poulin-Dubois, D. (2009). Infants' understanding of intention from 10 to 14 months: Interrelations among violation of expectancy and imitation tasks. *Infant Behavior and Development*, 32(4), 404-415.
- Onishi, K. H., Baillargeon, R., & Leslie, A. M. (2007). 15-month-old infants detect violations in pretend scenarios. *Acta Psychologica*, 124(1), 106-128.

- 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.
- Over, H., & Gattis, M. (2010). Verbal imitation is based on intention understanding. *Cognitive Development*, 25(1), 46-55.
- Perner, J., & Rafetseder, E. Counterfactual and Other Forms of Conditional Reasoning. (2011). In C. Hoerl, T. McCormack & S. R. Beck (Eds.). (2011). *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology*. Oxford University Press.
- Perner, J., Sprung, M., & Steinkogler, B. (2004). Counterfactual conditionals and false belief: A developmental dissociation. *Cognitive Development*, 19(2), 179-201.
- Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its Piagetian perspective: why alternative naming comes with understanding belief. *Cognitive Development*, 17(3), 1451-1472.
- Perner, J., & Lang, B. (2002). What causes 3-year-olds' difficulty on the dimensional change card sorting task? *Infant and Child Development*, 11(2), 93-105.
- Perra, O., & Gattis, M. (2008). Reducing the mapping between perception and action facilitates imitation. *British Journal of Developmental Psychology*, 26(1), 133-144.
- Peterson, C. C., & Wellman, H. M. (2009). From fancy to reason: Scaling deaf and hearing children's understanding of theory of mind and pretence. *British Journal of Developmental Psychology*, 27(2), 297-310.

- Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in Theory of Mind Development for Children With Deafness or Autism. *Child development*, 76(2), 502-517.
- Piaget, J., & Inhelder, B. (1948). *The child's conception of space*. London: Routledge and Kegan Paul.
- Piaget, J., & Inhelder, B. (1974). *The child's construction of quantities: Conservation and atomism*. London: Routledge & Kegan Paul.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39(4), 633-651.
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological bulletin*, 134(3), 427-459.
- Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in cognitive sciences*, 11(3), 105-110.
- Pine, J. M. (1992). How referential are "referential" children? Relationships between maternal-report and observational measures of vocabulary composition and usage. *Journal of Child Language*, 19, 75-86.
- Pinker, S., Lebeaux, D. S., & Frost, L. A. (1987). Productivity and constraints in the acquisition of the passive. *Cognition*, 26(3), 195-267.
- Poulin-Dubois, D., & Forbes, J. N. (2006). Words, intentions, and action: A two-tiered model of action word learning. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. New York: Oxford University Press.

- Poulin-Dubois, D., & Forbes, J. N. (2002). Toddlers' attention to intention-in-action in learning novel action words. *Developmental Psychology*, 38, 104 – 114.
- Pruden, S., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Development*, 77, 266 – 280.
- Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: Developing a sense of “nearest possible world”. *Child Development*, 81, 376–389.
- Rafetseder, E., & Perner, J. (2010). Is reasoning from counterfactual antecedents evidence for counterfactual reasoning? *Thinking and Reasoning*, 16, 131–155.
- Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, 114, 389-404.
- Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, 13, 73-90.
- Roberts Jr, R. J., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental neuropsychology*, 12(1), 105-126.
- Robinson, E. J., & Beck, S. (2000). What is difficult about counterfactual reasoning? In P. Mitchell & K. J. Riggs (Eds.), *Children's Reasoning and the Mind*. Hove, England: Psychology Press/Taylor & Francis (UK).
- Ryalls, B. O., Gul, R. E., & Ryalls, K. R. (2000). Infant imitation of peer and adult models: Evidence for a peer model advantage. *Merrill-Palmer Quarterly*, 188-202.

- Sakkalou, E., Ellis-Davies, K., Fowler, N. C., Hilbrink, E. E., & Gattis, M. (2013).  
 Infants show stability of goal-directed imitation. *Journal of experimental child psychology*, 114(1), 1-9.
- Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development*, 2(1), 27-49.
- Scott, F. J., Baron-Cohen, S., & Leslie, A. (1999). 'If pigs could fly': A test of counterfactual reasoning and pretence in children with autism. *British Journal of Developmental Psychology*, 7, 349-362.
- Scott, R. M., He, Z., Baillargeon, R., & Cummins, D. (2012). False-belief understanding in 2.5-year-olds: Evidence from two novel verbal spontaneous-response tasks. *Developmental Science*, 15(2), 181-193.
- Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science*, 10(3), 322-332.
- Schulz, L., Kushnir, T., & Gopnik, A. (2007). Learning from doing: Intervention and causal inference. *Causal learning: Psychology, philosophy, and computation*, 67-85.
- Seehagen, S., & Herbert, J. S. (2011). Infant imitation from televised peer and adult models. *Infancy*, 16(2), 113-136.
- Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology*, 47(5), 1239.

- Shimpi, P. M., Gámez, P. B., Huttenlocher, J., & Vasilyeva, M. (2007). Syntactic priming in 3-and 4-year-old children: Evidence for abstract representations of transitive and dative forms. *Developmental Psychology*, 43(6), 1334.
- Shipley, E. F., Smith, C. S., & Gleitman, L. R. (1969). A study in the acquisition of language: Free responses to commands. *Language*, 322-342.
- Simoneau, M., & Markovits, H. (2003). Reasoning With Premises That Are Not Empirically True: Evidence for the Role of Inhibition and Retrieval. *Developmental Psychology*, 39(6). 964-975
- Simpson, A., & Riggs, K. J. (2011). Three-and 4-year-olds encode modeled actions in two ways leading to immediate imitation and delayed emulation. *Developmental Psychology*, 47(3), 834.
- Slobin, D. I. (1963). *Grammatical transformations in childhood and adulthood*. Unpubl. doct. Dissert., Harvard University.
- Slobin, D. I. (1968). Recall of full and truncated passive sentences in connected discourse. *Journal of Verbal Learning and Verbal Behavior*, 7(5), 876-881.
- Smallman, R., & Roese, N. J. (2009). Counterfactual thinking facilitates behavioral intentions. *Journal of Experimental Social Psychology*, 45, 845-852.
- Smetana, J. G., Jambon, M., Conry-Murray, C., & Sturge-Apple, M. L. (2012). Reciprocal associations between young children's developing moral judgments and theory of mind. *Developmental Psychology*, 48(4), 1144-1155.
- Smith, M., Apperly, I. A., & White, V. (2003). False belief reasoning and the acquisition of relative clause sentences. *Child Development*, 74(4), 1709–1719.
- Smith, M., & Wheeldon, L. (2000). Syntactic priming in spoken sentence production— an online study. *Cognition*, 78(2), 123-164.

- Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive science*, 28(3), 303-333.
- Spoor, J., & Williams, K. D. (2007). The evolution of an ostracism detection system. *The evolution of the social mind: Evolutionary psychology and social cognition*, 279-292.
- Subiaul, F., Anderson, S., Brandt, J., & Elkins, J. (2012). Multiple imitation mechanisms in children. *Developmental Psychology*, 48(4), 1165-1179.
- Sudhalter, V., & Braine, M. D. (1985). How does comprehension of passives develop? A comparison of actional and experiential verbs. *Journal of Child Language*, 12(02), 455-470.
- Svartvik, J. (1996). *Corpora are becoming mainstream. Using corpora for language research*. London and New York: Longman.
- Talmy, L. (1985). Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and the lexicon*, Vol. III: Grammatical categories and the lexicon. Cambridge, UK: Cambridge University Press.
- Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, 10(7), 309-318.
- Thatcher, K., Branigan, H., McLean, J., & Sorace, A. (2008). Children's early acquisition of the passive: Evidence from syntactic priming. *2007 Child Language Seminar*, City University, London, 187-196.
- Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition*, 108(1), 51-68.
- Tomasello, M. (1992). *First verbs*. Cambridge, UK Cambridge University Press.



- Tomasello, M. (1995). Pragmatic contexts for early verb learning. In M. Tomasello & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive linguistics*, 11(1/2), 61-82.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences*, 28(05), 675-691.
- Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language*, 19, 311-333.
- Tomasello, M., Strosberg, R., & Akhtar, N. (1996). Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language*, 23, 157 – 176.
- Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, 4, 197 – 212.
- Towse, J. N., Redbond, J., Houston-Price, C. M., & Cook, S. (2000). Understanding the dimensional change card sort: Perspectives from task success and failure. *Cognitive Development*, 15(3), 347-365.
- Turner, A.E., & Rommetveit, R. (1968). Focus of attention in recall of active and passive sentences. *Journal of Verbal Learning and Verbal Behavior*, 7(2), 543-548.
- Uzgiris, I. C. (1981). Two functions of imitation during infancy. *International Journal of Behavioral Development*, 4(1), 1-12.

- VanderBorgh, M., & Jaswal, V. K. (2009). Who knows best? Preschoolers sometimes prefer child informants over adult informants. *Infant and child development*, 18(1), 61-71.
- van der Wel, R. P., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition*, 130(1), 128-133.
- Vasilyeva, M., Huttenlocher, J., & Waterfall, H. (2006). Effects of language intervention on syntactic skill levels in preschoolers. *Developmental Psychology*, 42(1), 164.
- De Villiers, J. G., & De Villiers, P. A. (1978). *Language acquisition*. Harvard University Press.
- Wagner, L., & Carey, S. (2005). 12-month old infants represent probably endings of motion events. *Infancy*, 7(1), 73 – 85.
- Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology*, 43, 217 – 242.
- Wellman, H. M., & Brandone, A. C. (2009). Early intention understandings that are common to primates predict children's later theory of mind. *Current opinion in neurobiology*, 19(1), 57-62.
- 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.
- Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychological Science*, 17(12), 1075-1081.

- Wellman, H. M., Lane, J. D., LaBounty, J., & Olson, S. L. (2011). Observant, nonaggressive temperament predicts theory-of-mind development. *Developmental Science*, 14(2), 319-326.
- Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development*, 75(2), 523-541.
- Wellman, H. M., Lopez-Duran, S., LaBounty, J., & Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Developmental Psychology*, 44(2), 618-623.
- Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology*, 7(2), 131-149.
- Whitehurst, G. J., Ironsmith, M., & Goldfein, M. (1975). Selective imitation of the passive construction through modeling. *Journal of Experimental Child Psychology*, 17, 288-302.
- Whitehurst, G. J., & Novak, G. (1973). Modeling, imitation training, and the acquisition of sentence phrases. *Journal of Experimental Child Psychology*, 16(2), 332-345.
- Whiten, A., & Brown, J. (1998). Imitation and the reading of other minds: Perspectives from the study of autism, normal children and non-human primates. *Intersubjective communication and emotion in early ontogeny*, 260-280.
- Williamson, R. A., & Markman, E. M. (2006). Precision of imitation as a function of preschoolers' understanding of the goal of the demonstration. *Developmental Psychology*, 42(4), 723-731.

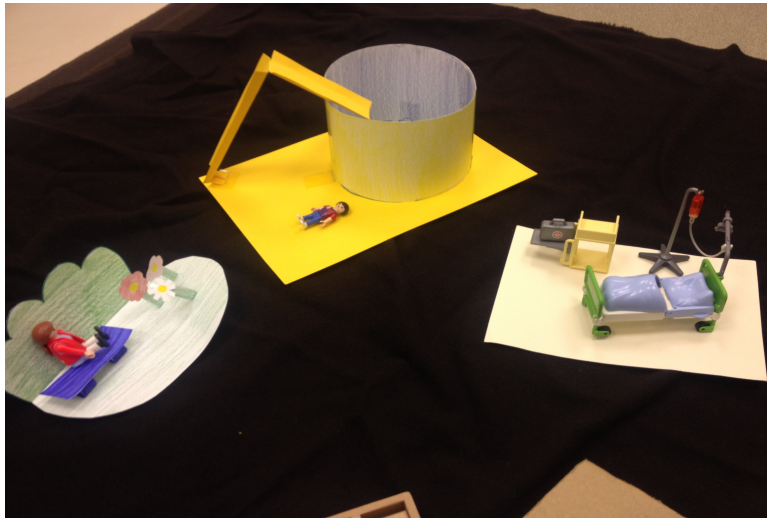
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103-128.
- Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Copy me or copy you? The effect of prior experience on social learning. *Cognition*, 127(2), 203-213.
- Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69, 1 – 34.
- Zabelina, D. L., & Robinson, M. D. (2010). Creativity as flexible cognitive control. *Psychology of Aesthetics, Creativity, and the Arts*, 4(3), 136.
- Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH toolbox cognition battery (CB): measuring executive function and attention. *Monographs of the Society for Research in Child Development*, 78(4), 16-33.
- Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: II. The development of executive function in childhood. *Current Directions in Psychological Science*, 121-126.
- Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, 11(1), 37-63.
- Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., ... & Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the society for research in child development*, i-151.
- Zimmer, C. (1986). Entrepreneurship through social networks. *The art and science of entrepreneurship*. Ballinger, Cambridge, MA, 3-23.

Zmyj, N., Daum, M. M., Prinz, W., Nielsen, M., & Aschersleben, G. (2012). Fourteen Month Olds' Imitation of Differently Aged Models. *Infant and Child Development, 21*(3), 250-266.

Zmyj, N., & Seehagen, S. (2013). The Role of a Model's Age for Young Children's Imitation: A Research Review. *Infant and Child Development, 22*(6), 622-641.

## Appendix A

### Counterfactual Reasoning Story Task



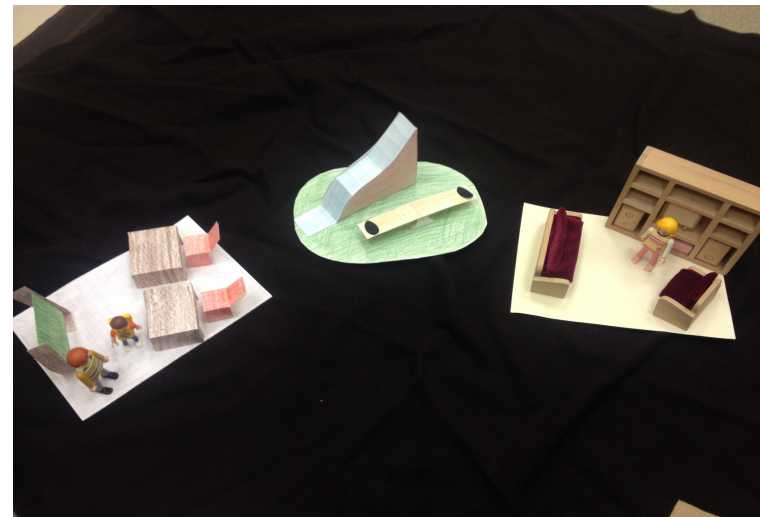
Doctor Story



Policeman Story



Fireman Story



Teacher Story

## Counterfactual Stories

### 1. Typical condition

#### *Doctor*

- Swimming Pool
- Hospital
- Park

Before work a Doctor was sitting in the park enjoying the beautiful weather. She left the park to go to work in the hospital. When she got there he received an emergency phone call. Look, there has been an accident at the swimming pool and little Jacob has slipped and hurt himself. Jacob needs a Doctor. The Doctor lifts her emergency first aid case and walks from the hospital to the swimming pool to help Jacob.

Control – **Where is the Doctor now?**

CF – **What is Jacob hadn't slipped and hurt himself, where would the Doctor be now?**

Forced choice – **Would the Doctor be at the park or at the swimming pool? / Would the Doctor be at the swimming pool or at the park?**

Example of a correct answer: She would be at the hospital.

Example of an incorrect answer: She would be at the swimming pool/park.

#### *Fireman*

- Living Room
- Fire Station
- Forest Fire

A Fireman was sitting at home in his living room before it was time to go to work. He then left there to start work in the fire station. When he got there he received a 999. Look there is a fire in the local forest! The fire brigade will need to come quickly. The Fireman got into his fire engine and drove to the forest to help fight the fire.

Control – **Where is the Fireman now?**

CF – **If the fire hadn't started, where would the fireman be now?**

Forced choice – **Would the Fireman be in his living room or would he be in the fire station? / Would the Fireman be in the fire station or would he be in his living room?**

Example of a correct answer: He would be at the fire station.

Example of an incorrect answer: He would be at the forest/home.

#### *Teacher*

- Playground

- Classroom
- Ben's Living Room

Before school started a teacher was in the playground looking after the school children playing. When the bell rang he went inside to the classroom to begin the lesson. Ben, one of the school children wasn't feeling very well so the teacher needed to bring him home to his mother. The teacher left the classroom and took Ben home.

Control – **Where is the teacher now?**

CF – **If Ben hadn't been feeling well where would the teacher be?**

Forced Choice – **Would the teacher be in the classroom or the playground? / Would the teacher be in the playground or in the classroom?**

Example of a correct answer: He would be in the classroom.

Example of an incorrect answer: He would be in the playground/Ben's living room.

### *Policeman*

- Shop
- Police Station
- Car Park

At lunchtime a Policeman was in the local shop buying some fruit and vegetables. Once he had finished he went to the police station where he worked. After he got to the police station he heard on the police radio that there had been a car accident. The Policeman got into his car and drove to the car park to take care of the accident.

Control – **Where is the Policeman now?**

CF – **If the car hadn't had an accident, where would the Policeman be now?**

Forced Choice – **Would the Policeman be in the local shop or would he be at the car park? / Would the Policeman be in the car park or would he be in the local shop?**

Example of a correct answer: He would be at the police station.

Example of an incorrect answer: He would be at the car park/shop

## **2. Atypical condition**

### *Doctor*

- Swimming Pool
- Hospital
- Park

After finishing work a Doctor was sitting in the park enjoying the beautiful weather. Suddenly she received an emergency phone call. Look, there has been an accident at the swimming pool and little Jacob has slipped and hurt himself. Jacob needs a



Doctor. The Doctor left the park to go to the hospital. The Doctor lifts her emergency first aid case and walks from the hospital to the swimming pool to help Jacob.

Control – **Where is the Doctor now?**

CF – **What is Jacob hadn't slipped and hurt himself, where would the Doctor be now?**

Forced choice – **Would the Doctor be at the park or at the swimming pool? / Would the Doctor be at the swimming pool or at the park?**

Example of a correct answer: She would be at the park.

Example of an incorrect answer: She would be at swimming pool/hospital.

### *Fireman*

- Living Room
- Fire Station
- Forest Fire

A Fireman was sitting at home in his living room after he had finished work. He received a 999 call. Look there is a fire in the local forest! The fire brigade will need to come quickly. He then left home to go to the fire station. The Fireman got into his fire engine and drove to the forest to help fight the fire.

Control – **Where is the Fireman now?**

CF – **If the fire hadn't started, where would the fireman be now?**

Forced choice – **Would the Fireman be in his living room or would he be in the fire station? / Would the Fireman be in the fire station or would he be in his living room?**

Example of a correct answer: He would be at home.

Example of an incorrect answer: He would be at the forest/ fire station

### *Teacher*

- Playground
- Classroom
- Ben's Living Room

After school had finished for the day a teacher was in the playground looking after some of the school children playing. Suddenly the Principal sent him a message that Ben, a boy in the homework club was feeling ill. The teacher needed to bring him home to his mother. The teacher left the playground to collect Ben from his classroom and then took him home.

Control – **Where is the teacher now?**

CF – **If Ben hadn't been feeling well where would the teacher be?**

Forced Choice – **Would the teacher be in the classroom or the playground? / Would the teacher be in the playground or in the classroom?**

Example of a correct answer: He would be in the playground.

Example of an incorrect answer: He would be at Ben's house/the classroom

### *Policeman*

- Shop
- Police Station
- Car Park

After he had finished work a Policeman was in the local shop buying some fruit and vegetables. However suddenly on his radio he heard there had been a car accident in a nearby car park. Quickly he went to the police station where he worked. Once he got to the police station the Policeman got into his car and drove to the car park to help take care of the accident.

Control – **Where is the Policeman now?**

CF – **If the car hadn't had an accident, where would the Policeman be now?**

Forced Choice – **Would the Policeman be in the local shop or would he be at the car park? / Would the Policeman be in the car park or would he be in the local shop?**

Example of a correct answer: He would be at the shop.

Example of an incorrect answer: He would be at the car park/ police station.

## Appendix B

### Location Change Counterfactual (Pipe) Task



Counterfactual pipe target apparatus

## Appendix C

### Excerpt from NIH Toolbox Scoring and Interpretation Guide

© 2006-2012 National Institutes of Health and Northwestern University.

#### NIH Toolbox Flanker Inhibitory Control and Attention Test (Flanker)

Description: The Flanker task measures both a participant's attention and inhibitory control. The test requires the participant to focus on a given stimulus while inhibiting attention to stimuli (fish for ages 3- 7 or arrows for ages 8-85) flanking it. Sometimes the middle stimulus is pointing in the same direction as the "flankers" (congruent) and sometimes in the opposite direction (incongruent). Twenty trials are conducted for ages 8-85; for ages 3-7, if a participant scores  $\geq 90\%$  on the fish stimuli, 20 additional trials with arrows are presented. The test takes approximately three minutes to administer. This test is recommended for ages 3-85.

Scoring Process: Scoring is based on a combination of accuracy and reaction time and is identical for both the Flanker and DCCS measures (described below). A 2-vector scoring method is employed that uses accuracy and reaction time, where each of these "vectors" ranges in value between 0 and 5, and the computed score, combining each vector score, ranges in value from 0-10. For any given individual, accuracy is considered first. If accuracy levels for the participant are less than or equal to 80%, the final "total" computed score is equal to the accuracy score. If accuracy levels for the participant reach more than 80%, the reaction time score and accuracy score are combined.

#### Accuracy Vector

There are 40 possible accuracy points:

· Flanker o Fish: 20 Points o Arrows: 20 Points

Individuals age eight and older automatically receive 20 accuracy points for the Fish Trials of the Flanker. (It was determined previously that they typically score at the ceiling on these trials.)

The accuracy score varies from 0 to 5 points. For every correct behavioral response, a participant receives a value of 0.125 (5 points divided by 40 trials) added to his/her score for Flanker:

Flanker Accuracy Score =  $0.125 * \text{Number of Correct Responses}$

#### Reaction Time Vector

The task-specific reaction time scores are generated using individuals' raw, incongruent median reaction time score from the Flanker. Median reaction time values are computed using only correct trials with reaction times greater than or equal to 100ms and reaction times no larger than 3 SDs away from the individual's mean (for respective trial type).

6

Like the accuracy score, the reaction time score ranges from 0 to 5 points. One issue regarding reaction time data is that it tends to have a positively skewed distribution. A log (Base 10) transformation is therefore applied to each participant's median reaction time score from the DCCS, creating a more normal distribution of scores. Based on the validation data, the minimum reaction time for scoring is set to 500ms and the maximum reaction time for scoring is 3,000ms. Participants with median reaction times that fall outside this range but within the allowable range of 100ms – 10,000ms are truncated (i.e., reaction times between 3,000ms and 10,000ms are set equal to 3,000ms) for the purpose of score calculation. Scoring of the validation data indicates that this truncation does not introduce any problems with regard to ceiling or floor effects. Log values are algebraically rescaled from a log(500)-log(3,000) range to a 0-5 range. Note that the rescaled scores are reversed; smaller reaction time log values are at the upper end of the 0-5 range while larger log values are at the lower end of the range. The formula for rescaling is:

$$\text{Reaction Time Score} = 5 - \frac{\log(\text{RT}) - \log(500)}{\log(3000) - \log(500)}$$

Once these reaction time scores are obtained, they are added to the accuracy scores for participants who achieved the accuracy criterion of better than 80%. For participants who fail to reach this criterion, only accuracy scores are used. This combination score is then converted to a scale score with mean of 100 and SD of 15.

Interpretation: The Flanker is a measure of executive function, specifically tapping inhibitory control and attention. It is considered a “fluid ability” – the capacity for new learning and information processing in novel situations – measure, in which performance reaches a peak in early adulthood, then tends to decline across the life span (based on health and individual factors, of course). To interpret individual performance, one can evaluate all three types of scale scores in which higher scores indicate higher levels of ability to attend to relevant stimuli and inhibit attention from irrelevant stimuli. In addition to the three scale scores provided, the Flanker Computed score provides a way of gauging raw improvement or decline from Time 1 to Time 2 (or subsequent assessments). This computed score ranges from 0-10, but if the score is between 0 and 5, it indicates that the participant did not score high enough in accuracy (80 percent correct or less). A change in the participant's score from Time 1 to Time 2 represents real change in the level of performance for that individual since the previous assessment. One can also put such a score in a different context by

comparing scale scores from Time 1 to Time 2, which will show the participant's performance relative to others (specific comparisons depending on which scale score is used).

### **NIH Toolbox Dimensional Change Card Sort Test (DCCS)**

Description: DCCS is a measure of cognitive flexibility. Two target pictures are presented that vary along two dimensions (e.g., shape and color). Participants are asked to match a series of bivalent test pictures (e.g., yellow balls and blue trucks) to the target pictures, first according to one dimension (e.g., color) and then, after a number of trials, according to the other dimension (e.g., shape). "Switch" trials are also employed, in which the participant must change the dimension being matched. For example, after four straight trials matching on shape, the participant may be asked to match on color on the next trial and then go back to shape, thus requiring the cognitive flexibility to quickly choose the correct stimulus. This test takes approximately four minutes to administer and is recommended for ages 3-85.

---

èë ûø

7

Scoring Process: Scoring is based on a combination of accuracy and reaction time. A 2-vector scoring method is employed that uses accuracy and reaction time, where each of these "vectors" ranges in value between 0 and 5, and the computed score, combining each vector score, ranges in value from 0- 10. For any given individual, accuracy is considered first. If accuracy levels for the participant are less than or equal to 80%, the final "total" computed score is equal to the accuracy score. If accuracy levels for the participant reach more than 80%, the reaction time score and accuracy score are combined.

#### Accuracy Vector

There are 40 possible accuracy points:

· DCCS o Pre-Switch (before changing to the other dimension): 5 Points o Post-Switch: 5 Points o Mixed Trials: 30 Points

Individuals age 8 and older automatically receive 10 accuracy points for the Pre-Switch and Post-Switch trials of the DCCS.

The accuracy score will vary from 0 to 5 points. For every correct behavioral response, a participant receives a value of 0.125 (5 points divided by 40 trials) added to his/her score for DCCS:

DCCS Accuracy Score = 0.125 \* Number of Correct Responses

The task-specific reaction time scores are generated using individuals' raw, non-dominant dimension (the one cued less frequently for sorting) median reaction time score from the DCCS.

#### Reaction Time Vector

The task-specific reaction time scores are generated using individuals' raw, non-dominant dimension (the one cued less frequently for sorting) median reaction time score from the DCCS. Median reaction time values are computed using only correct trials with reaction times greater than or equal to 100ms and reaction times no larger than 3 SDs away from the individual's mean (for respective trial type).

Like the accuracy score, the reaction time score ranges from 0 to 5 points. Reaction time data tends to have a positively skewed distribution. A log (Base 10) transformation is therefore applied to each participant's median reaction time score from the DCCS and Flanker, creating a more normal distribution of scores. Based on the validation data, the minimum reaction time for scoring is set to 500ms and the maximum reaction time for scoring is 3,000ms. Participants with median reaction times that fall outside this range but within the allowable range of 100ms – 10,000ms will be truncated (i.e., reaction times between 3,000ms and 10,000ms will be set equal to 3,000ms) for the purpose of score calculation. Scoring of the validation data does not indicate that this truncation introduces any problems with regard to ceiling or floor effects. Log values will be algebraically rescaled from a  $\log(500)$ - $\log(3,000)$  range to a 0-5 range. Note that the rescaled scores will be reversed; smaller reaction time log values will be at the upper end of the 0-5 range while larger log values will be at the lower end of the range. The formula for rescaling is:

$$\text{Reaction Time Score} = 5 - \frac{\log(\text{RT}) - \log(500)}{\log(3000) - \log(500)}$$

---

8

Once these reaction time scores are obtained, they are added to the accuracy scores for participants who achieved the accuracy criterion of better than 80%. For participants who fail to reach this criterion, only accuracy scores are used. This combination score is then converted to a scale score with mean of 100 and SD of 15.

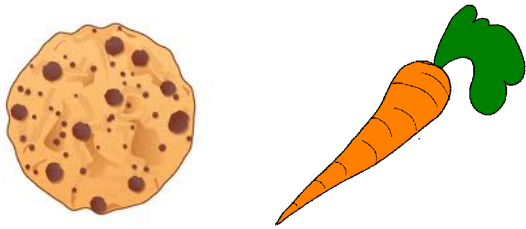
Interpretation: The DCCS is a measure of executive function, specifically tapping cognitive flexibility. It is considered a “fluid ability” measure, like Flanker, with performance generally increasing through childhood and then declining across the adult age span. To interpret individual performance, one can evaluate all three types of scale scores, where higher scores indicate higher levels of cognitive flexibility. In

addition to the three scale scores provided, the DCCS Computed score provides a way of gauging raw improvement or decline from Time 1 to Time 2 (or subsequent assessments). This computed score ranges from 0-10, but if the score is between 0 and 5, it indicates that the participant did not score high enough in accuracy (80 percent correct or less). A change in the participant's score from Time 1 to Time 2 represents real change in the level of performance for that individual since the previous assessment. One can also put such a score in a different context by comparing scale scores from Time 1 to Time 2, which will show the participant's performance relative to others (specific comparisons depending on which scale score is used).

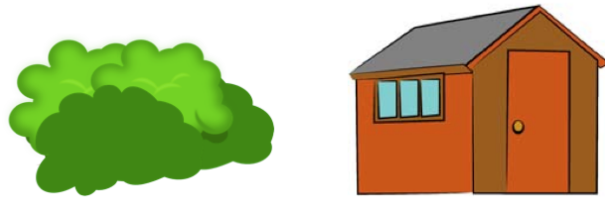


Appendix D

Mental State Understanding Resources



Question 1



Question 2



Question 4



Question 4



Question 5 Protagonist



Question 5 Emotion Faces

## Mental State Understanding Script

### 1-Diverse Desires

1. “Here’s Mr. Jones. It’s break time, so, Mr. Jones wants a snack to eat. Here are two different snacks: a carrot and a cookie. Which snack would you like most? Would you like a carrot or a cookie most?”
2. “Well that is a good choice, but Mr. Jones really likes cookies. He doesn’t like carrots. What he likes most is cookies.”
3. Target Q - “So now it’s time to eat. Mr. Jones can only choose one snack, just one. **Which snack will Mr. Jones choose? A cookie or a carrot?**”

### 2-Diverse Beliefs

1. “Here’s Linda. Linda wants to find her cat. Her cat might be hiding in the bushes or it might be hiding in the garage. Where do you think the cat is? In the bushes or in the garage?”
2. “Well, that’s a good idea, but Linda thinks her cat is in the garage. She thinks her cat is in the garage.”
3. Target Q - “**So where will Linda look for her cat? In the bushes or in the garage?**”

### 3-Knowledge Access

1. “Here’s a box. What do you think is inside the box?”
2. “Let’s see .... it’s really a train inside!”
3. “Okay, what is in the box?”
4. “Polly has never ever seen inside this drawer. Now here comes Polly. **So, does Polly know what is in the drawer?**”
5. Memory Q - “Did Polly see inside this drawer?”

### 4-Content False Belief

1. “Here’s a Juice Bottle. What do you think is inside the juice bottle?”
2. “Let’s see..... it’s really a pen inside!”
3. Container closed “Okay, what is in the juice bottle?”

4.Target Q -“Peter has never ever seen inside Juice bottle. Now here comes Peter. **So, what does Peter think is in the case? Juice or a pen?**

5.Memory Q - ‘Did Peter see inside this case?’” **5-Real-Apparent Emotion**

1.Check Q’s – “Can you point out which is the happy face, which is the sad face, which is the normal face?”

2.“This story is about a boy. I’m going to ask you about how the boy really feels inside and how he looks on his face. He might really feel one way inside but look a different way on his face. Or, he might really feel the same way inside as he looks on his face. I want you to tell me how he really feels inside and how he looks on his face.”

3.“This story is about Matt. Matt’s friends were playing together and telling jokes. One of the older children, Rosie, told a mean joke about Matt and everyone laughed. Everyone thought it was very funny, but not Matt. But, Matt didn’t want the other children to see how he felt about the joke, because they would call him a baby. So, Matt tried to hide how he felt.”

4.Memory checks:

**“What did the other children do when Rosie told a mean joke about Matt?”**

**“In the story, what would the other children do if they knew how Matt felt?”**

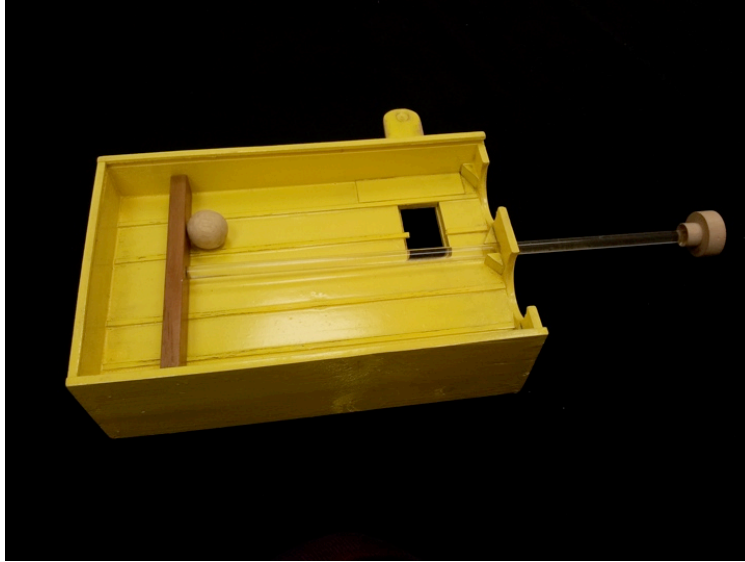
Pointing to the three emotion pictures:

**“So, how did Matt really feel, when everyone laughed? Did he feel happy, sad, or okay?”**

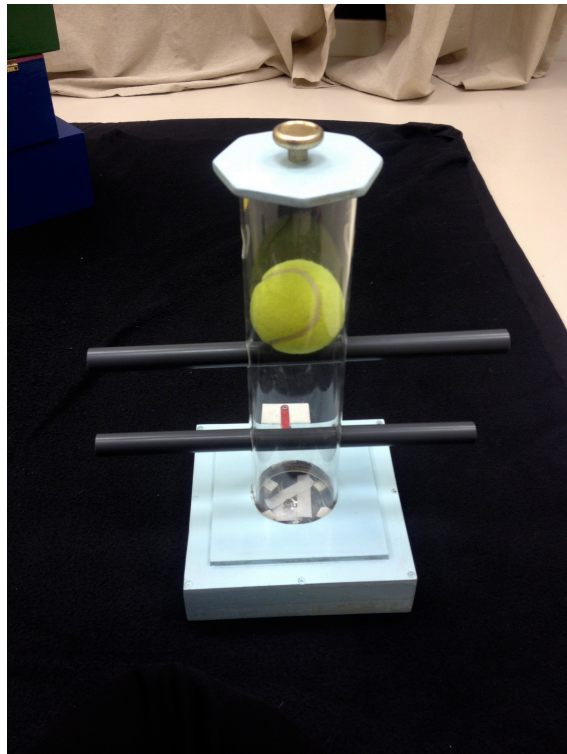
**“How did Matt try to look on his face, when everyone laughed? Did he look happy, sad, or okay?”**

## Appendix E

### Object Imitation Apparatus



Apparatus 1. Rake



Apparatus 2. Tower

## Appendix F

### Verbal Imitation Sentences

---

| <b>Grammatical Repetitions</b> | <b>Ungrammatical Repetitions</b>      |
|--------------------------------|---------------------------------------|
| He was in a dark dark house    | It was dark dark outside              |
| There there don't cry          | There there is the park               |
| He said bye bye to his horse   | He went to buy buy a horse            |
| Sam was a big big cat          | The cat was too big big for the chair |

---

| <b>Filler Sentences</b> |                        |
|-------------------------|------------------------|
| I like to eat cake      | Katie found an old toy |
| Where is my blue hat    | She is a tall girl     |
| I have a new friend     |                        |

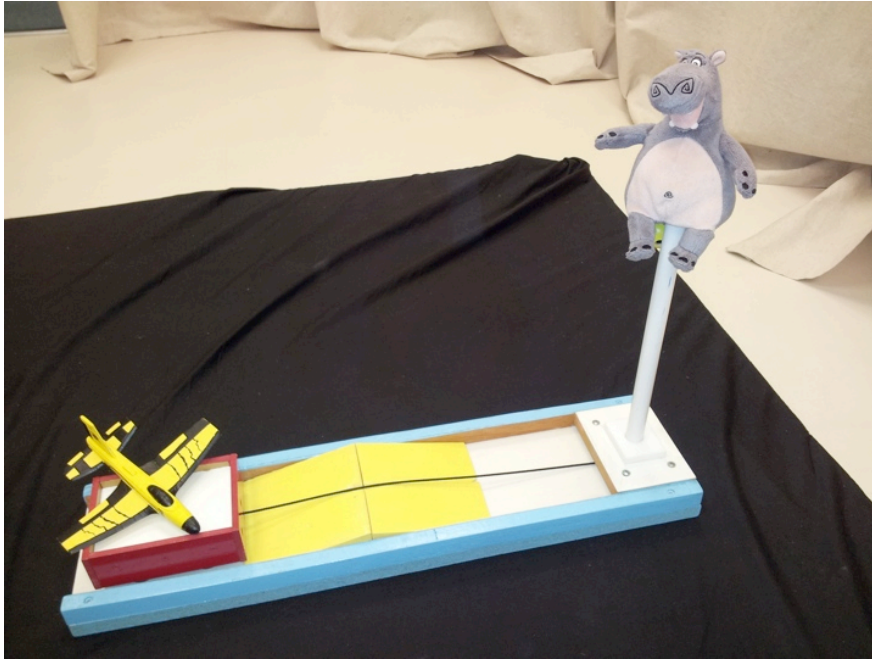
---

| <b>Practice Sentences</b> |             |
|---------------------------|-------------|
| Red                       | Door        |
| Small mouse               | Blue car    |
| A brown cat               | A big apple |

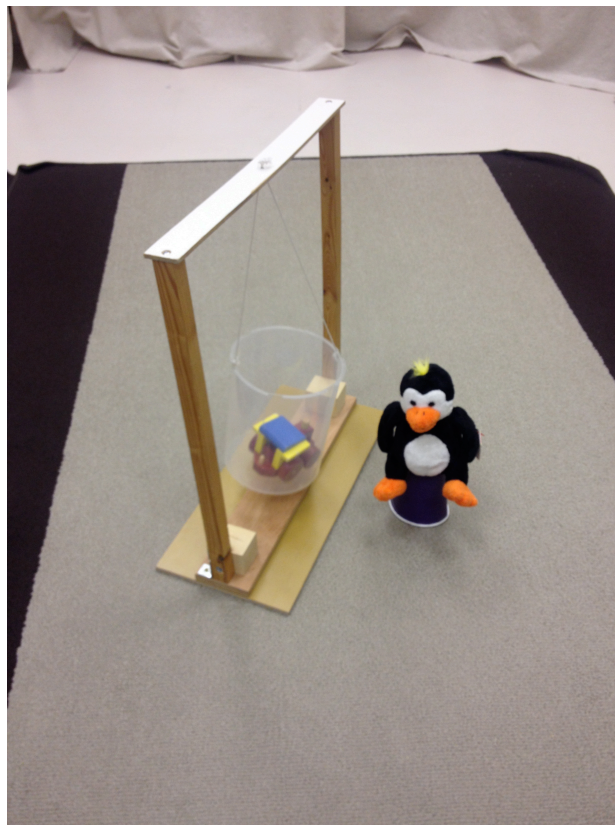
---

## Appendix G

### Verbal Imitation Apparatus



Trolley



Swing

## Verbal Imitation Example Training Script

| Passive construction   | Active Construction  |
|--|--|
| Look, the hammer is going to get meeked                                  | Look the horse is going to meek something                      |
| The hammer is going to get meeked by the horse                           | The horse is going to meek the hammer                          |
| What's going to get meeked? (Pointing to the hammer)                     | Who's going to meek the hammer? (Pointing to the horse)        |
| That's right, the hammer is going to get meeked                          | That's right, the horse is going to meek the hammer            |
| The hammer is going to get meeked by who? (Pointing to the horse)        | The horse is going to meek what? (Pointing to the hammer)      |
| Yes, the hammer is getting meeked by the horse. (While modelling action) | Yes, the horse is meeking the hammer. (While modelling action) |
| Did you see what got meeked by the horse?                                | Did you see who meeked the horse?                              |
| Exactly! The hammer got meeked by the horse.                             | Exactly! The horse meeked the hammer.                          |