Imitation: Evaluating neurophysiological signatures and clinical significance

Hannah Hobson

University College

University of Oxford

Thesis submitted for the degree of Doctor of Philosophy

at the University of Oxford

Trinity Term 2016

Department of Experimental Psychology

South Parks Road, Oxford, OX1 3UD

Abstract

Imitation and its neural basis has been the subject of great interest, particularly for researchers interested in autism spectrum disorders (ASD). Research has reported imitation deficits in ASD, which have been theoretically linked to the condition's sociocommunicative difficulties. Abnormalities in the mirror neuron system (MNS), as measured using EEG techniques, have also been reported in autism, and linked to imitation skills. In this thesis, I examine questions both around behavioural imitation deficits, and the validity of EEG techniques used to investigate the MNS. I consider whether children with other communication disorders show similar imitation deficits to children with autism, examining verbal and motor imitation. I report data from non-autistic children with and without language problems, and children with autism and normal language (ANL) and children with autism and language-impairment (ALI). In contrast to previous studies, I find similar verbal imitation performance in children with ALI and non-autistic language impaired children, but limited motor imitation impairments in the disorder groups. I also examine the evidence from mu suppression studies for the role of the MNS in language, social processes and the development of ASD. I present findings from a preregistered study that show that methodological factors such as baseline condition have important effects on mu suppression studies' results, and question its validity as an MNS measure. I also examine how imitation skills, self-reported communication abilities and autistic traits relate to individual differences in mu suppression. I find limited support for the theory that mu suppression correlated with behaviours and traits proposed to be underpinned by the MNS. I discuss the implications of these two projects for future research and clinical interventions for children with communication disorders.

Acknowledgments

The work in this thesis was supported by a Medical Sciences Division Studentship, provided by the Medical Research Council. I wish to thank the MRC, University College, the Post-graduate Affairs Group of the British Psychological Society, and the West Sussex Oxford University Society for the financial support I received to undertake and present my research at conferences, throughout my studies.

I am hugely grateful to all the children, families and schools that took part, and to all the participants of my adult study. The research could not have happened without them.

I would like to thank my supervisor, Dorothy Bishop, for her constant guidance and support throughout my DPhil studies. I have been very lucky to be surrounded by many inspiring female academics during my time at Oxford, and I thank all of them for giving me such good role-models to attempt to emulate.

I wish to thank Clara Grabitz and Nimarta Narang for their patient and diligent coding of my video data. I also thank Sally Rogers and Raphe Bernier for providing a copy of the MIT, and Tom Loucas for the algorithms and macros used for calculating Levenshtein Distance. My thanks also to Mihaela Duta for her assistance coding my mu suppression presentation scripts, and to Mervyn Hardiman for training me in conducting EEG. I wish to send special thanks also to the Oxford Science Adventures team!

Finally, I wish to thank my family and friends, who have encouraged me throughout my time at Oxford. I thank especially my husband Andy, and who has always been ready with a glass of wine when things were tough. Andy could quite probably have written a half-decent chapter on mu suppression by himself, after listening to me talk about it for so long. Thank you for keeping me going.

Table of Contents

PART TWO: Mu suppression as a neurophysiological measure of the MNS

List of publications

The following publications relate to research presented in this thesis:

Chapter 2:

Hobson, H.M. (submitted). Motor Imitation and Language in Typical and Atypical Development - Current Theories and Future Directions.

Chapter 5:

Hobson, H.M., & Bishop, D.V.M. (in prep). Mu suppression in the past, present and future.

Chapter 6:

Hobson, H. M., & Bishop, D.V.M (in press). Mu suppression – a good measure of the human mirror neuron system? *Cortex*.

List of Tables

List of Figures

List of abbreviations

Abbreviation	Meaning
3di	The developmental, dimensional and diagnostic interview
ADI	Autism Diagnostic Interview
ADOS	Autism Diagnostic Observation Schedule
ALI	Autism with Language Impairment
ANL	Autism with Normal Language
ASD	Autism Spectrum Disorder
AQ	Autism Spectrum Quotient
CCC	Children's Communication Checklist
CC-SR	Communication-Checklist Self Report
CELF	Clinical Evaluation of Language Fundamentals
CNRep	Children's Test of Nonword Repetition
DSM	Diagnostic and Statistical Manual of Mental Disorders
EEG	Electroencephalography
EMG	Electromyography
HFA	High Functioning Autism
HNO	Hand (No object)
HO	Hand (with object)
IQ	Intelligence Quotient
LFA	Low Functioning Autism
LI	Language Impaired
MIT	Mature Imitation Task
MNS	Mirror Neuron System
NEPSY	Developmental NEuroPSYchological Assessment

NHS	National Health Service
NWR	Nonword Repetition
PDD-NOS	Pervasive Developmental Disorder – Not Otherwise Specified
PLI	Pragmatic Language Impairment
PSTM	Phonological Short-Term Memory
SEN	Special Educational Needs
SENCo	Special Educational Needs Coordinators
SLCN	Speech Language and Communication Needs
SLI	Specific Language Impairment
TD	Typically Developing
WS	Williams Syndrome

CHAPTER ONE: Introduction

The study of imitation has spanned many diverse fields in psychology and cognitive science. In typical development, imitation represents a key tool for children in the acquisition of numerous skills, including language and understanding the behaviour of others (Meltzoff, 1999; Tomasello, 1992). Imitation has also been studied in atypical development, particularly in autism spectrum disorders (ASD)[1], in which imitation deficits have been viewed as a core feature, and may present a target for intervention (Ingersoll & Lalonde, 2010; Rogers & Pennington, 1991; Williams, Whiten, Suddendorf, & Perrett, 2001). Beyond child development, adult neuropsychological models have been developed to suggest distinct routes for different types of imitation (Rumiati et al., 2005). In recent years, the study of imitation has encompassed both behaviour and neuroimaging, with increased interest in the brain systems that underlie these skills, and how they may develop abnormally in autism (Iacoboni, 2005; Williams et al., 2001).

Across these fields, imitation has been linked, directly or indirectly, to language and communication in a variety of theories. Rogers and Pennington (1991) proposed that deficits in self-other mapping underlie both the imitation and broader sociocommunicative difficulties in ASD. Recent investigations into interventions for autism target imitation, in an attempt to improve the children's social and communicative symptoms (Ingersoll, 2010). In Tomasello's (1992) account on the

[1] Throughout the thesis, the terms "autism" and "autism spectrum disorder" are used interchangeably. This reflects the diagnostic label adopted in the DSM-V (American Psychiatric Association., 2013), which replaces a range of previous diagnostic terms on the autism spectrum, including autistic disorder, PDD-NOS, and Asperger's Syndrome (although for many clinicians and families, these labels still retain much significance and utility), with "autism spectrum disorder".

social bases for language development, imitation is a key process in children's acquisition of language, revealing an understanding of intentionality that scaffolds children's emerging communicative skills. Rizzolatti and Arbib (1998) suggest that mirror neurons, the substrate proposed to underlie action understanding and imitation, could be the biological level of explanation for the motor theory of speech perception (Liberman & Mattingly, 1985). Indeed, they posit that early communication in our ancestors may have depended on an imitation-based system – language thus evolved *out* of imitation. Oberman and Ramachandran (2007) extend the notion that a mirror neuron system (MNS) could underlie imitation and communication, arguing that a *broken* MNS could account for many of the socio-communicative deficits in autism, including both imitation and language difficulties. Others have challenged the notion that a deficit in observation-execution matching, underpinned by the mirror neuron system, is the reason for imitation difficulties in ASD (Leighton, Bird, Charman, & Heyes, 2008; Southgate & Hamilton, 2008). Nevertheless, the study of mirror neuron systems has been argued to be a key research need for other disorders of communication, not just autism (Le Bel, Pineda, & Sharma, 2009).

Despite the diverse and extensive research that has been done on imitation, there remain a number of gaps in our knowledge, for both imitation as a behaviour and its neurological substrates, and both in typical and atypical populations. Indeed, while the focus on imitation deficits in clinical groups has dominantly been on autism, other groups of children also have weaknesses in imitation, including children with Specific Language Impairment (SLI). While in autism, imitation research has centred on motor imitation, in SLI much interest has been paid to these children's verbal imitation skills, in particular their ability to repeat nonsense words in nonword repetition tasks. Over the last few years, researchers have used nonword repetition tasks as a means to explore

whether language problems in autism and SLI could have shared causal pathways, or arise for separate aetiological reasons. There are outstanding questions about how these two groups of children compare on verbal and motor imitation tasks, and whether imitation deficits may represent a site of behavioural or cognitive overlap. Is the imitation deficit specific to ASD, or is imitation a weakness in a broad range of communicative disorders, including children with SLI? Could imitation problems explain the nonword repetition deficits shown in some language subtypes in autism? Examining imitation across a range of disorders of communication would allow the relationship between imitation and language to be unpicked better, and could help inform interventions for children with communication difficulties.

Behavioural imitation deficits in autism have been attributed to abnormalities in the MNS (Oberman & Ramachandran, 2007; Williams et al., 2001). However, this notion remains contentious. EEG studies have played a central role in the debate, particularly investigations using an EEG phenomenon known as mu suppression. The human EEG can be decomposed into a number of different frequency bands – mu is a frequency band that has been associated with the activation of the motor cortex (Kuhlman, 1978; Schoppenhorst & Brauer, 1980). Following studies showing that mu-band activity can be altered both by participants' own movement and by observing the actions of others, researchers have suggested that changes in mu might be capturing changes in MNS activity (Muthukumaraswamy & Johnson, 2004).

If so, using mu suppression designs to investigate MNS activity would have a number of advantages; EEG is better tolerated by participants with autism than other imaging methods such as fMRI, and is comparably cheap and easy to administer. Indeed, given its ease and affordability, it is unsurprising that researchers have already used mu suppression designs to test a wide range of hypotheses about the role of the

MNS in language and social processes. Among these investigations, mu suppression has been used to show abnormal MNS functioning in ASD, and it has been argued that individual differences in mirror neuron functioning (as measured by mu suppression) relate to individual differences in imitation (Bernier, Aaronson, & McPartland, 2013; Bernier, Dawson, Webb, & Murias, 2007; Oberman et al., 2005). Mu suppression has also been proposed as a target for neurofeedback therapy in autism. If mu suppression could be used this way, it would satisfy the appetite for a biologically-based treatment for autism that is not pharmacological. Other researchers have also called for mu suppression investigations to be done with other communication disorder groups to examine the role of mirroring systems in communication problems beyond autism (Le Bel, Pineda, & Sharma, 2009), yet, as far as I am aware, no work has been undertaken thus far.

Thus, mu suppression may appear to offer the opportunity to examine neurophysiological questions about the systems thought to underlie imitation and autism. However, the validity and reliability of mu suppression as a measure of the human mirror neuron system is questionable. In particular, it is unclear to what extent mu suppression may be confounded by attentional factors, and how certain methodological decisions may make this confound even worse.

1.1 Structure of the thesis

This thesis touches on a number of questions associated with imitation and autism, through two distinct and quite different projects, explored in the two parts of the thesis. Part 1 considers behavioural imitation and the links to language in clinical samples – children with autism and non-autistic children with language impairments. Chapter 2 explores the current evidence for a relationship between imitation and

language, reviewing previous correlational evidence in typical and atypical development, and considers whether the imitation deficit might not be specific to autism but found in SLI as well. Cross-disorder comparisons of verbal and motor imitation are examined in Chapters 3 and 4.

Part 2 examines the electrophysiological phenomenon of mu suppression which has been argued to measure the activity of the mirror neuron system, and has been used to evidence a "broken" MNS in autism. As explored in a review in Chapter 5, mu suppression has been widely used in the study of mirror neuron involvement in processes such as language and social understanding. However an examination of what such studies have found proves rather discouraging, and there are consistent methodological flaws in mu suppression studies. Mu suppression captured my interest as a means to explore the underlying neurological causes of the imitation deficit in autism, and potentially to consider as yet unexplored questions about mirroring systems in non-autistic children who have communication impairments, such as children with Specific Language Impairment (SLI). However, both early piloting and an extensive review of the measure suggested some fundamental issues with mu suppression designs. Instead of using this technique with clinical populations, a thorough methodological study was conducted with typical adults. Chapter 6 contains the findings of a study examining methodological issues with mu suppression. Chapter 7 extends this to consider whether mu suppression relates to individual differences in autism traits, communication skills and imitation skills, as would be predicted by previous theories about the MNS and studies with clinical groups.

1.2 Brief overview of autism spectrum disorders and specific language impairment

Rather than a traditional general literature review in the introductory chapter of the thesis, two chapters (Chapters 2 and 5 respectively) contain in-depth reviews specific to the behavioural links between motor imitation and language in typical and atypical development, and mu suppression in the study of mirror neuron systems and language and social processes. However, the original motivation for the thesis was to better understand the relationships between ASD and SLI, and Part 1 is largely concerned with these two conditions. Thus, a brief overview of both disorders is given here.

First, it is useful to set current thinking about autism and SLI in the context of broader ideas about the nature of many developmental disorders. In the fields of autism, SLI and other developmental disorders research, there has been a gradual shift away from considering these conditions as qualitatively different from typical development, and arising from single cognitive deficits. Instead, it has been argued that SLI and ASD represent extreme ends on a continuous spectrum, and arise due to a range of potential and possibly co-occurring deficits (Bishop, 2006; Happé, Ronald, & Plomin, 2006). This change in the behavioural work on these disorders has been paralleled by similarly changing theories in genetics, which is moving away from searches for "the autism gene", and single-gene accounts of common developmental disorders, to accepting that the genetic aetiology of these conditions likely arises from multiple genetic and environmental effects (with potentially interactive and epistatic effects), which may also be shared between conditions (Plomin & Kovas, 2005). Research into developmental disorders generally – not just autism and SLI – is thus moving into a new era with increased acceptance of the complexity of the causes of these conditions, and that behaviourally defined conditions may in truth be made up of children with deficits arising from diverse neurobiological underpinnings.

1.2.1 Autism Spectrum Disorders

Autism Spectrum Disorders (ASD), or autism, is a condition in which a child shows markedly poor social and communication skills, and restricted and repetitive interests (American Psychiatric Association, 2013). It is a neurodevelopmental condition with a strong genetic basis (Rutter & Thapar, 2014). Some genetic syndromes are known to have a higher rate of ASD - for example, Fragile X Syndrome (Hatton & Sideris, 2006) – however, in many cases of autism there is no known genetic origin, and the exact genes that increase the risk of autism are still unclear. The prevalence of ASD is thought to be around 1% (Baird et al., 2006), with increasing prevalence over the last few decades thought to be largely due to changes in diagnostic criteria and increased awareness of the condition (Gernsbacher, Dawson, & Goldsmith, 2005; Matson & Kozlowski, 2011). It is considered to be a lifelong condition, but the prognosis of a child diagnosed with autism is hugely varied. Many individuals on the spectrum also have intellectual disability, but many others will have normal or excellent intellectual functioning. Furthermore, while many autistic children do learn to speak – and may even develop precocious language skills – many others remain nonverbal, or develop only very limited language. In a study of 43 autistic participants, Sigman and McGovern (2005) found that 49% of their sample had little to no functional language by adolescence. In a similarly sized study, Kjelgaard and Tager-Flusberg (2001) found only a quarter of their sample of autistic children had language skills within the normal range for their age (and the authors note that their selection criteria may have biased their sample towards children with at least some verbal abilities). Language level in early childhood is predictive of later general social and cognitive functioning (Gillberg & Steffenburg, 1987; Szatmari, Bryson, Boyle, Streiner, & Duku, 2003), and poor or absent language is a source of concern for many parents of autistic children, even after a

diagnosis has been made (Gray, 1994). Understanding language development and difficulties, and what causes the large array of language outcomes in autism, remains an important area of research.

Over the decades, there have been many suggested core deficits in autism, including impaired theory of mind, weak central coherence, and executive dysfunction (Baron-Cohen, Tager-Flusberg, & Cohen, 1994; Frith, 1989; Russell, 1997). More recent theories about autism include abnormal levels of "neural noise" (e.g. Dinstein et al., 2012), and abnormalities in the mirror neuron system (Williams et al., 2001). However, classic cognitive theories of ASD have struggled to explain all the difficulties associated with the condition, and autism's vast heterogeneity. This has led some to call for a move away from single-deficit explanations of autism, and embrace the notion that the full behavioural and cognitive profile of autism may be underpinned by multiple co-occurring deficits, each with their own genetic aetiology (Happé et al., 2006).

1.2.2 Structural language abilities in ASD

Investigations into language ability in ASD has been largely concerned with "high functioning"[2] groups, while individuals with very poor language or intellectual disability (ID) have been largely neglected by research (Boucher, Mayes, & Bigham, 2008). Furthermore, structural language in ASD has been little researched, despite evidence that structural language is frequently if not always atypical in autism

[2] The terms high functioning and low functioning (HFA and LFA) are widely used in the autism literature and are referred to in this thesis. However, it is important to acknowledge that these terms are problematic, and not necessarily welcomed by the autistic community. The definitions of HFA and LFA are variable, with different researchers opting for different thresholds to divide children or adult participants into these groups. Furthermore, "functioning" is usually defined by how a participant scores on an IQ test, but outside of research studies is often conflated with severity of autistic symptoms. For further discussion about the problems of using the terms "high functioning" and "low functioning", see the insightful blogpost by Dawson (2010).

(Boucher, 2012). Pragmatic problems (difficulties with the appropriate use of language in context) are a defining feature of autism and common to all children on the autism spectrum. Structural language describes other aspects of language, including grammar, phonology and word learning. In her review of the evidence on structural language abilities in ASD, Boucher (2012) proposes an 'ASD-typical' language profile, and that differences between the language subtypes ALI (Autism with Language Impairment) and ANL (Autism with Normal Language) are largely in severity, rather than the pattern of strengths and weaknesses. Most prominent in this profile are weak comprehension and semantic skills, and a general difficulty with the processing of meaning. Many autistic children exhibiting poor language in the preschool years (including those whose language problems appear to resolve later), show additional problems with phonology and grammar, but articulatory and syntactic impairments tend to resolve, with a few morphological errors remaining more persistent. It is comprehension, semantics and the processing of meaning that is generally reliably and severely affected in persistent cases of ALI. In cases of ANL, when language is never delayed or catches up, Boucher (2012) argues that language abilities are still uneven, showing the same relative strengths and weaknesses as found in ALI: weaker comprehension compared to expression (though even in expressive abilities there are some anomalies, with words and phrases sometimes used idiosyncratically). Despite vocabulary performance within the normal range, semantic processing is argued to exhibit some peculiarities, suggesting semantic weaknesses milder but of a similar kind to those in ALI.

1.2.3 Specific Language Impairment

There is currently a lack of consensus around the *strict* definition, criteria and terminology for Specific Language Impairment (SLI), and these matters are currently the topic of much discussion in the field of childhood language disorders (Bishop,

2014). Broadly, the term SLI describes a child in which there is a mismatch between verbal and nonverbal skills, which cannot be attributed to a known physical or environmental cause (including deafness, or cognitive impairments such as Down Syndrome, or social deprivation). In effect, SLI can be thought of as a condition in which a child displays poor language for no obvious reason. Like ASD, there is evidence to suggest SLI is a heritable condition, though the rate of heritability depends on how cases of SLI are defined and selected (Bishop & Hayiou-Thomas, 2008).

The nature of the language difficulties in SLI varies greatly, affecting expressive abilities, receptive abilities or both, and children exhibiting one profile do not necessarily retain it throughout development, meaning that a child's most prominent difficulties may change over time (Botting, 2007; Leonard, 2014). Children with SLI can be difficult to understand, as they may have quite unintelligible speech, or word finding problems (Nation, 2008). They may also have grammatical difficulties, omitting certain morphemes when they talk (e.g. "She walk to school yesterday" instead of "She walk*ed* to school yesterday"), and struggle to understand grammatically complex sentences (Botting, 2007; Rice & Wexler, 1996). In general, most attention has been paid to these children's *structural* language problems, but children with SLI can also show difficulties with the appropriate use of language, even when they seem able to produce intelligible language (Nation, 2008).

SLI is thought to affect roughly 7% of school-aged children (Tomblin et al., 1997). Despite being more common than autism, SLI seems less well known by the general public, and there is much less research funding for SLI than there is for ASD (Bishop, 2010b). The prognosis of children with SLI is not well understood, with few studies following children with SLI into adulthood. Studies that have investigated the outcomes of adults who were considered to have SLI as children are discussed in

Bishop (2014), and suggest that these populations are at higher risk of poor academic and social outcomes. Some young children who exhibit language difficulties do appear to catch up as they grow older (Bishop & Edmundson, 1987). However research has highlighted that some of these may be cases of illusory recovery, and early language problems can leave a lasting legacy on literacy and academic achievement (Stothard, Snowling, Chipchase, & Kaplan, 1998).

Theoretical accounts attempting to explain the primary deficits in SLI can be broadly divided into two groups – those that propose the underlying deficit in SLI is specific to the domain of language, and those that suggest a domain-general deficit, encompassing others areas of cognition as well. One influential domain-specific theory was proposed by Rice, Wexler, and Cleave (1995), which argued that SLI arises due to the delayed maturation of a modular system involved in grammatical development. In this account (known as the 'Extended Optional Infinitive' theory), children with SLI do not progress at the usual rate from a particular stage in linguistic development. This theory draws support from the pattern of grammatical errors children with SLI commonly make, particularly their difficulties with appropriately marking tense.

Domain-general theories include arguments that children with SLI have low-level auditory processing difficulties, or an impairment in procedural learning (Tallal & Piercy, 1973; Ullman & Pierpont, 2005). One theory that straddles the domain-specific/domain-general divide is the notion that children with SLI have a deficit in phonological short term memory (PSTM) (Gathercole & Baddeley, 1990). A deficit in memory could arguably be considered to be a domain-general account, but the proponents of this theory have argued that children show a limitation in phonological memory specifically. Phonological short term memory, or the phonological loop, is a component of memory thought to have a key role in language development. Baddeley,

Gathercole, and Papagno (1998) argued that humans possess a specialised PSTM system, evolved to facilitate language learning, and that poor PSTM limits children's vocabulary acquisition and development of syntax. Support for deficient PSTM in SLI comes from experiments of nonword repetition, a task that children with SLI show marked difficulties with.

Given the competing evidence for each of these different deficits, there have been attempts to unite them into more parsimonious accounts. For instance, could nonword repetition difficulties actually be a manifestation of a low-level auditory processing problem (Bishop et al., 1999)? However, considering how different deficits in SLI relate to one another has presented researchers with a more complex picture. In one group of twins (including twin pairs with and without language impairment), Bishop et al. (1999) noted that the language-impaired children showed deficits in both auditory processing and phonological short term memory. Despite this, the correlation between these two deficits was rather weak, and showed evidence for distinct causes – nonword repetition performance showed strong heritability, while auditory processing tasks did not. These findings seem difficult to square with a single-deficit account. Still, incorporating a developmental perspective, Bishop et al. (1999) suggested that perhaps children with SLI do have auditory processing difficulties initially, but that these may resolve over time, leaving a lasting impact on children's phonological short term memory abilities, which continue to limit their language development.

In another experiment, Bishop, Adams, and Norbury (2006) examined the evidence that grammatical deficits and phonological short term memory deficits share their genetic aetiologies. In fact, while again the language-impaired children did show deficits in both tasks, analysis of data acquired from twins suggested that their genetic underpinnings were quite different. Again, these findings are difficult to reconcile with

a single-deficit, or single-gene, approach. Instead, they can be interpreted within a framework that considers SLI to arise due to multiple deficits, with different genetic causes, that work together to increase the risk that a given child will exhibit language difficulties. Such an account is fitting with observations that, by and large, children's language acquisition is remarkably robust. Bishop (2006) suggests that language acquisition may be able to take multiple routes, and that children suffering one very specific deficit may get by, showing limited impairments in their language. However, when multiple impairments occur, possible alternative routes to successful language acquisition are further limited, and language impairment is made more likely. This account also fits with the notion that SLI is rarely caused by a single simple genetic effect, but more likely due to a constellation of genetic risk factors.

1.2.3.1 How "specific" is SLI?

A common approach in developmental disorders research is to attempt to exclude children who exhibit comorbid symptoms of other conditions. This "pure" sample is considered to have more theoretical meaning. However, SLI exhibits a high degree of comorbidity with other developmental disorders, such as dyslexia, dyspraxia and ADHD (Catts, Adlof, Hogan, & Weismer, 2005; Gillberg, 2010) and it is unclear how representative a sample of "pure" SLI cases is, given that clinically SLI rarely appears without symptoms of other conditions. The point stretches beyond SLI, and indeed developmental disorders – it has been shown that screening for comorbid conditions when researching psychopathology in general can have dramatic effects on the results obtained (Newman, Moffitt, Caspi, & Silva, 1998). In the case of SLI however, common comorbidities raise a larger nosological issue – the *specificity* of language problems in SLI have arguably been considered this condition's defining

feature. Some have criticized the view that the difficulties in SLI are "specific" to language, given the high prevalence of these additional problems in SLI (Hill, 2001).

Another difficulty in SLI research is whether there is a need to show that weaker nonverbal skills could not account for language problems. That is, could a child with poor language skills simply be considered to have *global* deficits in *all* domains, not just language? Such a case would appear contrary to the very definition of SLI, but there is debate about whether to be classified as SLI a child would need to demonstrate skills in the normal range for their age, or to demonstrate a mismatch between verbal and nonverbal skills. This discrepancy approach is sometimes referred to as 'cognitive referencing'. However, this notion is poorly supported by the evidence; Tomblin's (2008) longitudinal study found limited differences in language problems, academic and social outcomes between cases of language impairment with and without nonverbal cognitive difficulties. Furthermore, this stance runs contradictory to the view that language and cognitive development are overlapping and dynamically integrated, and it seems quite plausible that having a language impairment could impede certain aspects of cognitive development. Indeed, a longitudinal study by Botting (2005) showed that as children with SLI grew up, their intelligence appeared to fall. Of course, as children get older and intelligence tests become suitably harder (in order to be appropriate for the average child of a given age), not having sufficiently good language skills may limit a child's reasoning skills, even on seemingly nonverbal tasks.

1.2.3.2 Pragmatic Language Impairment

There exist a third group of children, who appear to present with relatively intact structural language (unlike SLI), but very poor pragmatic language skills, while not meeting the criteria for an ASD. These children have been described in several ways,

including “semantic pragmatic disorder”, “pragmatic language impairment” (PLI), and “social communication disorder”. These children may have expressive language that is fluent, but the use of this language is abnormal; they are verbose, have problems understanding and producing a connected discourse, and give conversational responses that are socially inappropriate. The recent release of the DSM-V included a new diagnostic label “Social Communication Disorder”, and PLI is forecast for inclusion in the ICD-11.

PLI/SCD is not the focus of this thesis, but the existence of such a group of children holds implications for how we consider the social communication problems in SLI and ASD. Some have suggested that these cases might be thought of as being intermediate between ASD and SLI (Bishop, 2000). The existence of children with PLI would seem to suggest that pragmatic communication skills develop at least partially independently from structural language skills and autism, per se. However, as discussed by Bishop (2000), while some children with quite pure pragmatic deficits *can* be found, many children with SLI do show pragmatic language problems, and there is a correlation between pragmatic and structural language skills, which undermines a strong dissociative account.

1.2.4 Why study ASD and SLI together?

While text-book cases of ASD and SLI would be considered quite distinct, evidence suggests the boundaries between them are not so clear. While language ability is extremely diverse in autism, some children with ASD and poor language skills appear to show comparable difficulties to children with SLI, including poor performance on tasks regarded as clinical markers for SLI (Kjelgaard & Tager-Flusberg, 2001). Conversely, despite a focus on children’s structural language abilities, some children

with SLI show broader communication and social difficulties, and a significant minority develop symptoms of autism spectrum disorder, or even meet the diagnostic criteria for autism (Conti-Ramsden, Simkin, & Botting, 2006). While we may expect there will be a small number of children who inherit risk genes for both autism and SLI by chance, and thus manifest both disorders independently, it is clear this account cannot explain the number of seemingly overlapping cases. Multiplying the prevalence rates of autism and SLI together, the prevalence rate of a combined presentation would be 0.07%, and would thus be unlikely to have come to the research community's notice. Indeed, family studies and investigations into the genetic causes of either condition have also suggested there may be shared genetic risk factors between SLI and ASD (Folstein & Rutter, 1977; Folstein et al., 1999; Le Couteur et al., 1996; Tomblin, Hafeman, & O'Brien, 2003 - but see Bishop, 2010 and Szatmari, Jones, & Tuff, 1993).

Research on the relationships between ASD and SLI could yield theoretical insights, both for understanding these conditions in their own right, and for broader theories about language and social development. Aside from these scientific gains, evidence of similarity and difference between ASD and SLI also has important applications for clinical and educational practices, in both diagnosis and provision for these groups of children. Early language difficulties are often considered a "red flag" for autism as well as SLI, and some have suggested that disentangling cases of autism and cases of language difficulty in young children can be especially challenging, which is problematic in light of a drive for even earlier diagnosis and intervention (Charman & Baird, 2002). In fact, there is evidence that over the decades diagnostic substitution has taken place – cases of language disorder diagnosed a few years ago would be more likely to receive a diagnosis of ASD now (Bishop, Whitehouse, Watt, & Line, 2008). The diagnostic label given to a child has non-negligible ramifications for the support

that the child will receive; for example, in a report for the Department of Education, Dockrell, Ricketts, Charman, and Lindsay (2010) noted that children with ASD were more likely to receive additional resources than children with language impairment, including in subjects such as English, in which children with language impairments showed poorer attainment than children with autism. Accurate diagnosis, intervention and information about likely prognosis is key to supporting these children during their school years and beyond, and appreciating how these two disorders relate to each other has an important role in this understanding.

CHAPTER TWO: Motor imitation and language in typical and atypical development

2.1 Introduction

Imitation, the copying of a behaviour demonstrated by a model, has long been a topic of interest in developmental psychology; Piaget wrote about the development of imitation sixty years ago (Piaget, 1951), and Meltzoff and Moore's seminal work on neonatal imitation is now thirty years old (Meltzoff & Moore, 1983). Imitation has been linked, directly and indirectly, to language in multiple accounts. Tomasello (1992) describes imitative learning as a key process in children's acquisition of words and grammar. Imitation features in the behavioural checklists of clinicians making a diagnosis of autism spectrum disorder (ASD), a condition in which poor social and communicative skills are arguably the most prominent deficits. Poor imitation in autism has been argued to reflect problems with self-other mapping, which could lead to a cascade of social and communicative deficits. These mapping problems have been suggested to be the result of abnormalities in mirror neuron systems, which could underpin both the social and language difficulties seen in ASD (Oberman & Ramachandran, 2007; Rogers & Pennington, 1991). Therapies targeting imitation for children with ASD are being investigated, in part based on the notion that imitation is causally related to communicative development, and that increasing a child's imitation will have positive effects on their social and language development (Ingersoll & Lalonde, 2010).

However, despite these links and despite several reports suggesting imitation to be predictive of language development, there is yet no definitive model for how imitation and language are related, in typical development or in developmental

disorders. Indeed, few investigations have attempted to examine the mechanisms behind this relationship directly. This chapter considered the association between motor imitation and language – is there a special relationship between the two, and if so what could underlie it?

2.1.1 The taxonomy of imitation and language

Imitation is not considered a unitary phenomenon, but has been subtyped in many ways. Furthermore, it has also been suggested that language, and different aspects of language, may be differentially related to different types of imitation (Stone, Ousley, & Littleford, 1997). Considering the structure of both imitation and language is thus a central issue to how the two are related.

Neuropsychological models of imitation have been developed out of work on apraxia (Tessari, Canessa, Ukmar, & Rumiati, 2007; Tessari & Rumiati, 2004) and from social psychology and imitation performance in autism (Hamilton, 2008). Both models suggest that there are two routes by which imitation can operate - a direct and an indirect route. The route by which imitation takes place depends on what is being imitated. Meaningful or goal-oriented actions are served by the indirect route, in which visual input activates the meaning or goal of the action, and this in turn then activates the motor sequences involved in imitating it. Imitating meaningless actions or tasks without goals is served by a direct connection between visual input and motor sequence outputs. While these models have arisen largely from studies of adults, similar divisions between meaningful and meaningless actions have been made in the developmental literature. For example, the direct route has been argued to be selectively impaired in autism, and previous reports have suggested that body imitation (likely the direct route) is more strongly related to language skills than object-based imitation (likely the indirect route) (Hamilton, 2013; Stone, Ousley, & Littleford, 1997).

Furthermore, when considering how imitation relates to language, it is important to recognise that "language" encompasses a wide range of skills. Language is often separated into sub-categories, such as structural versus pragmatic, or expressive versus receptive. Structural language skills include phonology, syntax and vocabulary, while pragmatic skills refer to the individual's use of appropriate use of language in context. These sub-categories can also be described as expressive (being able to share one's thoughts, ideas and feelings) or receptive (understanding what others are saying). Speech is not the same as language; speech is related to the articulation and fluency of words expressed, and it is possible for an individual with a speech difficulty to have intact language skills.

2.1.2 Using clinical populations to understand imitation and language

Studies of typically-developing children provide some useful information, however they may be limited when it comes to distinguishing between certain theories, as most skills tend to be intercorrelated in normal development. Clinical groups where skills become dissociated can allow investigation into what skills go together (see Bishop & Norbury, 2005). Indeed, groups of children with different communication and language difficulties arguably provide a natural means to examine the relationships between imitation and different aspects of language. These conditions include children with autism spectrum disorders (ASD) and children with Specific Language Impairment (SLI).

A large proportion of the research that has been conducted on imitation in developmental populations has been on children with autism, and imitation is known to be an area of deficit (see a recent meta analysis by Edwards, 2014). Older studies of ASD focused on low-functioning or very young children, where language is not well-developed. However, language skills are considerably variable in autism, with some

children displaying good and even precocious structural language skills. In fact, some have argued that language subtypes exist within autism, with some showing structural language difficulties comparable to those seen in SLI (Kjelgaard & Tager-Flusberg, 2001). Previous reviews of imitation in autism bewailed the lack of studies that had controlled for language level (Smith & Bryson, 1994), but since then much more work on imitation in autism has been undertaken, including studies that had control groups with comparable verbal IQ. It would seem of considerable interest whether these children also have spared skills in at least some aspects of imitation, and given that imitation has been shown to predict language skills in autism, we may expect to see differences between groups of autistic children with different language levels (Stone & Yoder, 2001). Is the imitation deficit evident throughout the autistic spectrum, including children with poor and good language?

As well as children with ASD, children with other communication difficulties could shed light on the relationships between imitation and language, including children with SLI. As described in Chapter 1, SLI is a condition in which a child fails to develop appropriate language skills, despite adequate opportunity, and no apparent hearing loss or cognitive impairment that could account for poor language. While pragmatic problems are usually the focus in ASD, children with SLI are usually considered to have spared wider socio-communicative skills but poor structural language. Imitation deficits are certainly not considered typical of SLI as they are in ASD, but preserved or impaired imitative skills in this population may point to which aspects of language imitation is most strongly linked to. It could be that imitation relates more to pragmatic than structural language development. If so, we may predict that imitation may be affected in all cases of ASD (as all children with ASD show pragmatic deficits regardless of structural language ability), but that there will be spared imitation skills in

SLI. If imitation is more strongly related to structural language problems, imitation in SLI can be expected to be affected, and imitation problems in autism will pattern with the language status of the children.

2.2 Reported correlations between language and imitation in typical and atypical development

Correlational studies seem particularly pertinent to the question of how strongly imitation and language are related, and whether certain subtypes of language or imitation show special relationships. For studies examining correlations between imitation and language, forest plots of effect sizes and confidence intervals were created, and are shown in Figures 2.1 and 2.2. Summary tables describing these studies in brief are included in the Appendices A and B. These plots and tables include data on children who are typically developing, and children with an ASD (few correlational studies were found for language-impaired children – these are discussed later).

For studies of typically developing children, 13 studies that were found that reported correlations. Sample sizes ranged from just 12 to over 42,000 participants. Similarly, 13 studies were found that examined the relationship between language and imitation in autism spectrum disorders, and sample sizes ranged from 12 to 164. There were two studies that do not feature in the forest plots as their statistical methods were not comparable to the other studies. This includes a study by Vivanti et al. (2008) (*N*=13, for the typically developing group), who found that in their analyses that language level contributed to imitation performance of meaningless gestures in their typically developing 11-year-olds, but found no significant correlation between imitation of meaningful object-based actions or non-meaningful gestures and language level in their autistic sample (*N*=17). Freitag et al. (2006) found pragmatic language skills to be unrelated to imitation in ASD (*N*=15) and also considered these associations

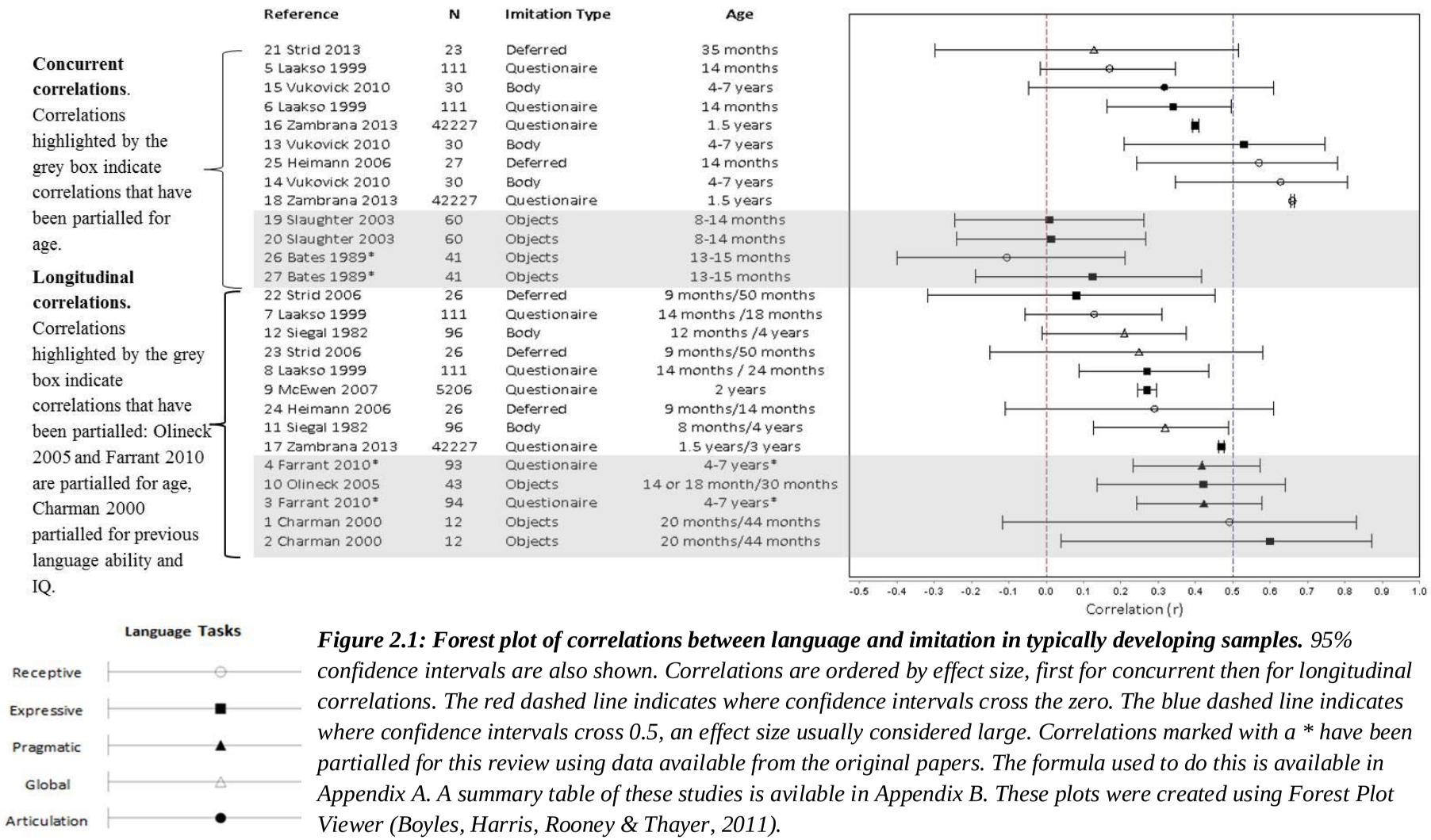


Figure 2.1: Forest plot of correlations between language and imitation in typically developing samples. *95% confidence intervals are also shown. Correlations are ordered by effect size, first for concurrent then for longitudinal correlations. The red dashed line indicates where confidence intervals cross the zero. The blue dashed line indicates where confidence intervals cross 0.5, an effect size usually considered large. Correlations marked with a * have been partialled for this review using data available from the original papers. The formula used to do this is available in Appendix A. A summary table of these studies is avilable in Appendix B. These plots were created using Forest Plot Viewer (Boyles, Harris, Rooney & Thayer, 2011).*

Concurrent correlations. Correlations highlighted by the grey box indicate correlations that have been partialled: Ingersoll 2011b and Rogers 2003 partialled out developmental level, nonverbal IQ was partialled out for Ingersoll 2011 and Luyster 2008, and age for Dawson 1984.

Longitudinal correlations. Correlations highlighted by the grey box indicate correlations that have been partialled: previous language ability was partialled by Stone and Yoder 2001, and IQ for Sallows 2005.

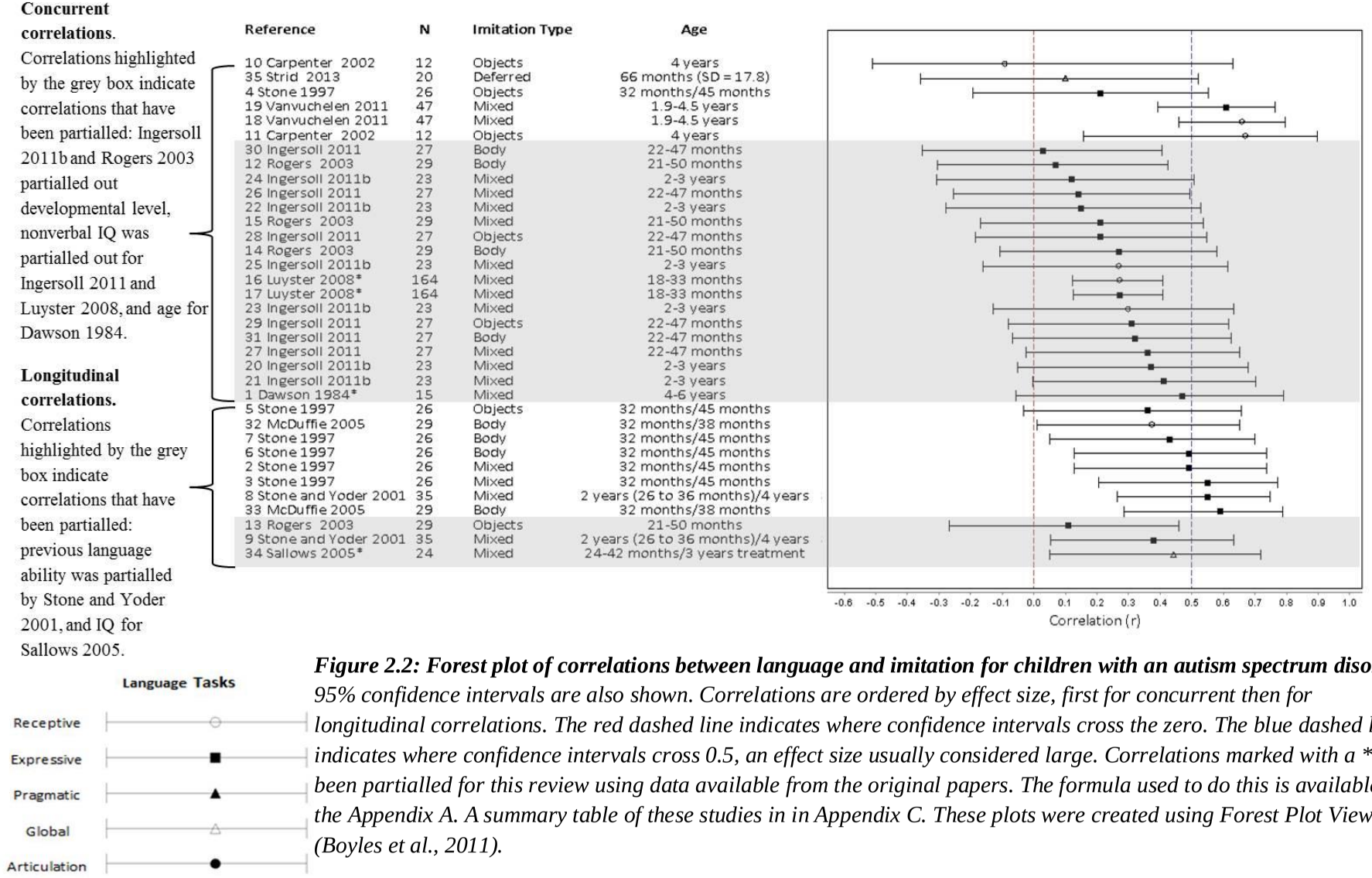


Figure 2.2: Forest plot of correlations between language and imitation for children with an autism spectrum disorder. *95% confidence intervals are also shown. Correlations are ordered by effect size, first for concurrent then for longitudinal correlations. The red dashed line indicates where confidence intervals cross the zero. The blue dashed line indicates where confidence intervals cross 0.5, an effect size usually considered large. Correlations marked with a * have been partialled for this review using data available from the original papers. The formula used to do this is available in the Appendix A. A summary table of these studies in in Appendix C. These plots were created using Forest Plot Viewer (Boyles et al., 2011).*

for their typical control group (*N*=29). Overall, few convincing correlations emerged, and none would survive Bonferroni corrections.

Of the studies with typically developing groups, five (Bates, Thal, Whitesell, Fenson, & Oakes, 1989; Freitag, Kleser, von Gontard, & von Gontardf, 2006; Slaughter & McConnell, 2003; Strid, Heimann, & Tjus, 2013; Strid, Tjus, Smith, Meltzoff, & Heimann, 2006) found no significant correlations between imitation and language on any measures. For studies of children with ASD, four found no correlations (Freitag et al., 2006; Rogers, Hepburn, Stackhouse, & Wehner, 2003; Strid et al., 2013; Vivanti, Nadig, Ozonoff, & Rogers, 2008). The others reported small to moderate correlations with at least one language measure.

Evidence from across the lifespan appears quite lacking, and the focus of study in this field has been early in development. The majority of these studies were conducted with quite young children and toddlers, and where correlations were available for older typical children, they were primarily controls for participants with SLI or ASD, suggesting limited study of imitation in later typical development in its own right. In investigations with autistic children, the samples are slightly older than in many of the studies with typically developing children – this is likely due to the age at which autism can be reliably diagnosed.

It is difficult to see any methodological factor that can readily explain variations in findings between studies. It has been argued that there are stronger relationships between imitation (in particular body imitation) and vocabulary or expressive abilities than with comprehension (Charman, Baron-Cohen, et al., 2000; Stone & Yoder, 2001). However, many studies *only* tested vocabulary, making it difficult to tell whether there are differential relationships between imitation and different components of language.

Nonetheless, in many studies with separate measures for expressive and receptive language no systematic pairing between expressive abilities and imitation is usually observed (e.g. Bates et al., 1989; Vukovic, Vukovic, & Stojanovik, 2010; Zambrana, Ystrom, Schjolberg, & Pons, 2013).

Considering the pattern of results for samples of typical developing children, it does seem that concurrent studies partialled for age rarely yield significant results – longitudinal correlations however seem more robust, even when partialled. This would seem to suggest that *early* imitation skills are indeed correlated with *later* language. However, the partialled concurrent correlations shown on the forest plot represents data drawn from just 2 studies, and these were of infants aged 8 to 15 months, an age at which language skills were probably only just emerging in these children. This could reduce the variance of language skills in these young populations, potentially limiting the degree to which correlations can even be observed.

Only three studies (Brookman, McDonald, McDonald, & Bishop, 2013; Vukovic et al., 2010; Wray, Norbury, & Alcock, 2015) were found that tested the association between imitation and language in language impaired children, and two of these studies' correlations are actually drawn from across the whole sample of language-impaired and typical controls (Brookman et al., 2013; Wray et al., 2015). Two studies report positive correlations between imitation and language (Brookman et al., 2013; Vukovic et al., 2010). While primarily an investigation of children's gesture, Wray et al (2015) used a hand imitation task as a measure of motor ability, and performance on this task was not found to correlate with receptive or expressive language abilities (or gesture). However, as noted by the authors, the language-impaired sample did not score in the impaired range on language assessment – potentially this sample did not have sufficient language impairment for clear imitation problems to

manifest (though the variance reported for the hand imitation task would suggest sufficient variation for a correlation to arise). Other studies were found that reported poor imitation in language impaired children compared to typically developing groups (Dohmen, Chiat, & Roy, 2013; Marton, 2009), but these studies did not look for correlations between these abilities.

Overall, the correlational studies between imitation and language, both for typically developing children and children with communication disorders, suggest that an association between imitation and language is commonly observed. However, the strength of this relationship is difficult to ascertain, given that many of these studies used small sample sizes, leading to very wide confidence intervals (the studies with the tightest confidence intervals represent those that used questionnaire-based methods, which naturally allow researchers to collect much larger samples). No obvious patterning between specific imitation tasks or language skills emerges, but as a number of studies measured only certain types of imitation or language skills the evidence to draw upon to is restricted.

The reason for the association between imitation and language was not the focus of direct study in many of these studies, and hence ideas expressed about what drives this correlation are speculative. Nonetheless, despite these limitations, it is still possible to explore several theories for the mechanisms that link imitation and language, drawing both on these correlational investigations but also wider evidence and theories from related fields of study. Five different theories linking imitation and language are explored in the rest of this chapter. The first theory considers a simple causal link between imitation and language – that perhaps imitation is the means by which we develop language. The second theory is that imitation and language are related only through very domain-general factors, such as IQ or developmental level. The final three

theories consider specific points of cognitive overlap that could explain the association between imitation and language. For each theory, evidence from typical development and children with communication difficulties is considered.

2.3 Is language dependent upon imitation?

First of all, could there be a simple, direct causative link between imitation and language? An extreme theory relating these two skills is that language is learned through imitation, and so language is totally dependent upon it. Children who are good imitators (assuming that verbal and nonverbal imitation are highly related) therefore develop better language skills. This perspective harks back to old behaviorist ideas about language, penned by Skinner (1957), and hotly rejected by Chomsky (1959). Imitative learning is a key process in the account given by Tomasello (1992) on the social basis of language development. This account would also arguably predict that the relationship between language and imitation will change across the lifespan; the two skills will likely show the closest relationship early in development, but that this association may peter out as language becomes more developed, and children perhaps stop relying on imitation as a dominant means to acquire new language.

Considering groups of children who have very limited motor skill allows us to consider how much language development is dependent upon imitation. Overall, the literature suggests that the notion that language development is *totally* dependent upon imitation can quickly be rejected; explicit imitation is evidently not required, because children who have severe motor impairments and who cannot imitate at all can still learn to comprehend language. Investigations of groups of children with cerebral palsy who are unable to speak, or find speaking very difficult, have highlighted that being able to articulate speech sounds is not a prerequisite for all areas of language

development (Bishop & Robson, 1989). Similarly, children with spinal muscular atrophy have been found to have precocious language skills, despite their motor impairments (Bénony & Bénony, 2005; Sieratzki & Woll, 2002). These findings suggest that a simplistic, direct link between imitation (or motor skills generally) and language cannot be right.

These ideas are discussed by Bishop (1998), which argues that, provided children have unimpaired hearing and intelligence, children who cannot physically produce speech can still learn to communicate very well. Arguably, considering again Tomasello's (1992) account, despite their motor impairments these children may still have intact intention understanding, a skill that is crucial to Tomasello's argument, more so than imitation. Indeed, Tomasello notes that it is impaired intention reading that may underlie children with autism's often abnormal language development; a child with autism may parrot what they hear around them (in fact echolalia is quite commonly observed in ASD), but without intersubjective participation and without an understanding of intention. Such mimicking is argued not to support full and proper language development.

Furthermore, this account would seem to depend on the notion that verbal and motor imitation are highly related to one another, and that a general imitation factor underlies them both, linking motor imitation to language. However, evidence suggests that verbal and gestural imitation are not that well inter-related (Rodgon & Kurdek, 1977). What would be ideal is evidence that motor and verbal imitation both predict language development, but that the predictive power of motor imitation and that of verbal imitation are independent, or that verbal imitation mediates the correlation between motor imitation and language skills – as yet, it is unclear to what extent verbal and motor imitation are strongly related.

In sum, motor imitation appears neither necessary nor sufficient to underpin language development, and as motor imitation and verbal imitation are not necessarily strongly related, a direct model in which imitation underlies language development does not seem adequate to explain the association between these skills.

2.4 Do general developmental factors mediate the relationship between imitation and language?

Another key question is how far the correlation between imitation and language may be simply due to a general maturational factor ('developmental level'), just as foot size and language will be correlated in developing samples. Both imitation and language rely on general skills, such as memory and attention, which affect most cognitive processes. They may both be proxies for general developmental level, or intelligence - imitation may be a predictor of general intelligence in a preverbal infant, and language may be a good indicator of intelligence once the child learns to speak. If we assume that intelligence is stable (i.e. infants who had good cognitive ability grow up to be children with good cognitive ability), then we would expect imitation and language to correlate. Thus, it may be that both abilities, like almost all other skills in development, improve with age and maturation, and there is nothing particularly special about the imitation and language relationship.

For the correlational studies on typically developing children, only one was found that partialled out IQ (and previous language ability) (Charman et al., 2000), but the relationship between imitation and language remained significant. This does suggest that the association between imitation and language is not just part of a general factor reflecting intellectual development. Indeed, imitation problems in SLI would also seem to run counter to the notion that language and imitation are linked through some general development factor. A diagnosis of SLI requires that children’s nonverbal IQ skills be

within the normal range, and strict definitions of SLI would aim to rule out intellectual problems. Arguably then, imitation problems in these groups of children would seem to provide evidence that imitation deficits can arise in children without general cognitive difficulties. However, this point must be tempered by the fact that evidence suggests that in reality some nonverbal weaknesses do typically accompany language problems in SLI (Botting, 2005).

In studies of children with autism, partialling out or statistically controlling for intelligence, age and general developmental factors goes some way to reducing the association between imitation and language, potentially to the point of insignificance (Ingersoll & Meyer, 2011b; Rogers, Hepburn, Stackhouse, & Wehner, 2003). Nonetheless, given that a strong relationship between imitation and intelligence has not always been demonstrated, it is unlikely the association between imitation and language is *only* a reflection of general ability. Imitation deficits appear throughout the autistic spectrum, even in individuals with normal to high IQ, and imitation has not always been found to correlate with intelligence (Vanvuchelen, Roeyers, & De Weerdt, 2007; Williams, Whiten, Suddendorf, & Perrett, 2001). Indeed, it may be that for some children a general cognitive problem will limit their imitation ability, but for children with normal to high intelligence who still demonstrate poor imitation ability other factors must be at play. Vanvuchelen et al. (2011) found that for children with autism and low IQ, imitation ability was in line with their mental age, but imitation problems were apparent in children with an IQ above 80, and children with an IQ above 100 showed imitation ability that was significantly below their nonverbal mental ages. This pattern of results suggests that adequate cognitive ability is certainly necessary for imitation, but even once children with autism have sufficient general cognitive skills, some other deficit becomes the limiting factor. These findings also raise implications

for studies examining only "low functioning" or "high functioning" groups of children – findings from selective groups will not necessarily generate conclusions that can be generalised across the autistic population.

Of course, it is important to note that we cannot assume the direction of causality between intelligence and imitation. Imitation also represents an important learning tool, and it could be that reduced imitation results in fewer skills and abilities, including lower IQ scores. By statistically controlling for IQ, we may actually "control away" the very skill we are interested in (Rogers & Williams, 2006). Longitudinal data may help to disentangle the correlations between imitation and development. Indeed, social communication behaviours in infancy have been found to predict both communication skills *and* intellectual functioning in childhood (Poon, Watson, Baranek, & Poe, 2012). This makes sense when one considers that social learning and social experiences gate many learning opportunities for children, and suggests that imitation has long term predictions for not only language but also more general cognitive functioning.

2.5 Do representational skills mediate the relationship between language and imitation?

Considering even the seemingly relatively straightforward model of general factors mediating imitation and language highlights a number of gaps in our knowledge. Nonetheless, a simple model linking imitation and language through very general developmental factors cannot account for the full pattern of results. Potential specific cognitive skills that could provide a site of overlap between imitation and language should therefore be considered. The first candidate proposed is symbolic or representational skills.

The emergence of deferred and facial imitation in particular have been suggested to mark a transition in infant development, away from being largely dependent upon sensorimotor domains and towards a representational cognitive system (Piaget, 1951). These higher-level imitations were linked by Piaget to the emergence of an infant's first words, and to symbolic play. Later evidence indicated that infants are capable of representational thought much earlier than Piaget previously suggested (Meltzoff & Moore, 2002), but symbolic representation may still be a key prerequisite for language development, and has been considered the link between children's gesture and speech (Bates et al., 1989), and recent studies on language and imitation still appeal to this framework (e.g. Zambrana, Ystrom, Schjolberg, & Pons, 2013). In the literature on imitation in autism, it has been argued that tasks that do not involve objects, or tasks in which objects must be used in an unconventional way (i.e. as a symbol for another object), are particularly impaired (Hamilton, 2008; Smith & Bryson, 1994). This may lend some credence to the idea that representation is a more primary deficit than imitation, and that representation problems might account for imitation deficits as well as difficulties with theory of mind tasks (see Leslie, 1987; Baron-Cohen 1988).

This account would predict particular imitation tasks should correlate with language, especially if the imitation task is symbolic (such as pantomime tasks, or gestures with meaning), but this pattern has not been demonstrated (Vivanti et al., 2008). Given that deferred imitation is supposed to tap higher level representational skills than immediate imitation, one plausible hypothesis is that deferred imitation should show a stronger correlation to language than immediate imitation. Very few empirical studies have explored the relationship between deferred imitation and language (let alone compared immediate and deferred imitation as correlates). While one study on children with autism did find that deferred imitation was predictive of

language development (Toth, Munson, Meltzoff, & Dawson, 2006), other studies have found weak correlations between deferred imitation and language (Strid et al., 2013, 2006). As noted by Rogers, Young, Cook, Giolzetti, and Ozonoff (2008), deferred imitation has received much less research interest than immediate imitation in autism, though their own study suggested no difference in the performance of deferred and immediate imitation in young children with autism. No articles were found that had investigated deferred imitation in SLI (this is no doubt because deferred imitation is a method suitable for quite young children, and therefore it is unlikely that children young enough to be meaningfully tested with deferred imitation would have received a diagnosis of language impairment).

Overall, at present there is limited evidence for symbolic skills having a large role in mediating the relationship between language and imitation, although there were few studies to draw upon.

2.6 Do motor skills mediate the relationship between imitation and language?

Motor skills may also be a key mediator between language and imitation skills. Language production requires good control over the articulators, and success on typical imitation tasks relies on good motor control over the hands, arms and face. Furthermore, motor difficulties have been proposed to underlie the imitation deficits in autism (Vanvuchelen et al., 2007), motor impairments have been found in children with speech and language impairment (Hill, 2001), and twin studies have highlighted that genes affecting communication development also affect motor development (Bishop, 2002).

A motor account would predict associations between imitation and motor ability (and imitation and articulation skills in particular) – in typical development, this patterning has not always been found (Vivanti et al., 2008; Vukovic et al., 2010).

However, the findings from studies with autistic participants have suggested a key role for motor skills – particularly praxis problems – in the imitation deficit in ASD. Praxis tasks include imitating gestures, but also performing gestures to verbal command and tool use. Difficulties with praxis are not inherently the same as basic motor problems. As pointed out by several authors, while in adult neuropsychological models the term "dyspraxia" really describes individuals who show impaired skilled motor performance *despite normal motor dexterity*, this is not always the case in developmental studies (Dziuk et al., 2007; Mostofsky et al., 2006). Imitation problems have been the focus in autism, but reports highlight broader motor and praxis problems (Dewey, Cantell, & Crawford, 2007; Dziuk et al., 2007; Green et al., 2009; Leighton, Bird, Charman, & Heyes, 2008; Mostofsky et al., 2006), and strong correlations between imitation and motor skills in studies of children with ASD have been reported (Luyster, Kadlec, Carter, & Tager-Flusberg, 2008; Vanvuchelen et al., 2007; Vanvuchelen, Roeyers, & De Weerdt, 2011).

The difficulties autistic individuals show with praxis tasks extend *beyond* basic motor problems, and children with autism show imitation deficits even compared to groups of children with dyspraxia (Dewey et al., 2007; Dziuk et al., 2007). Praxis problems that include and extend beyond imitation problems suggest a broad difficulty with action representation and execution. Praxis difficulties have also been shown to correlated with autism symptomatology, leading some to suggest these impairments may be core to the disorder (Dziuk et al., 2007). Recently it has been speculated that even the suggested dysfunctions in the mirror neuron system in autism could have a root in poor planning and representation of individuals' *own* actions, not just a problem mapping others to the self (Fabbri-Destro, Cattaneo, Boria, & Rizzolatti, 2009). This account would suggest that the difficulties autistic individuals have with imitating (and

indeed understanding) others arises from lacking good representations for actions in themselves.

Could a primary deficit in action representation also explain the relationship between imitation and language? The potential role for language in the praxis deficits demonstrated by their sample is discussed briefly by Dewey et al. (2007), but otherwise work on the praxis deficits in autism has not considered relationships to language skill. Given that praxis deficits have been associated with greater autism symptomatology (Dziuk et al., 2007), it may be that poorer language skills would be seen in children with the worst praxis, on account of having the most severe autism. Alternatively, given that accounts arguing for the role of the mirror neuron system in language have drawn upon the motor theory of speech perception, perhaps the notion of a deficit in action representation could be broadened to include poor representation of articulatory gestures. This would seem problematic however, as weak representations of articulatory gestures would be expected to manifest in poor articulation – but articulation is commonly considered a relative strength in language in autism (Kjelgaard & Tager-Flusberg, 2001).

Motor deficits have also been found in children with SLI, and appear to be a common co-occurrence in children with language problems (Hill, 2001). Studies that have found imitation deficits in SLI have also noted deficits in motor coordination, and motor coordination has been found to predict children with SLI's imitation performance (Marton, 2009; Vukovic et al., 2010). However, imitation deficits in language-impaired groups have been found even when compared to children who did not differ from the clinical group in terms of motor skills (Dohmen et al., 2013).

In sum, while motor skills likely play an important role in the observed imitation problems, in both ASD and SLI, problems with imitation (or praxis more broadly) seem to extend beyond basic motor problems, in autism at least. Examining the mediating role of motor abilities in the relationship between language and imitation will require a distinction between basic motor and praxis skills. Indeed, doing so may lead us to reframe the relationship between imitation and language to one of *praxis*, or action representation and language. Including broader praxis batteries (such as those used by Dewey et al., 2007, Dziuk et al., 2007, and Mostofsky et al., 2006) would allow for better understanding of whether language and imitation show a special relationship, or if language is actually related to skilled actions and action representation in general.

2.7 Does social engagement mediate the relationship between imitation and language?

Alternatively, social engagement may be the link between a child's imitation and language abilities. Socially engaged children imitate more, and are rated by their parents as having better conversational skills (Farrant, Maybery, & Fletcher, 2010). Dohmen et al. (2013) suggested that imitation difficulties in children with language delay were indicative of a problem connecting socio-emotionally with the model. Children with autism are characterised by difficulties with social engagement and demonstrate abnormalities in orienting to social stimuli (Hobson & Lee, 1998; Riby, Brown, Jones, & Hanley, 2012). Social motivation problems in ASD have been proposed as a primary deficit, and feasibly imitation difficulties or reduced imitation may be a consequence or symptom of reduced social engagement (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012). Poor social engagement may also have negative consequences for an autistic child's language development through reduced joint attention, and indeed joint attention is predictive of both language and imitation in

children with autism (Charman et al., 2000; Pickard & Ingersoll, 2014; Toth et al., 2006).

Very few studies have directly examined the effects of social engagement on imitation, in typical or atypical samples. However, in a recent review, Etten and Carver (2015) examined the potential importance of social engagement in the imitation deficits in ASD, highlighting that while few studies have empirically tested the role of social motivation in imitation deficits, motivation could explain the pattern of imitation difficulties typically reported in ASD. For example, copying instrumental actions better than meaningless gestures, or failing to copy the style of the actor, could both be interpreted as a lack of motivation to engage and connect with the model (intriguingly this is a very similar interpretation to that of Dohmen et al. (2013) in their study of toddlers with speech and language delay).

The potential role of social engagement on imitation skills and language development raises some important points about how imitation tasks are commonly administered and scored. The performance of a child on an imitation task is not only about how well they imitate, but also whether they *choose* to imitate. Vivanti (2014) notes that *propensity* to imitate and *accuracy* of imitation are not the same thing, and propensity to imitate (how often does the child choose to engage in imitative behaviour) may be more related to social motivation than accuracy is. This calls into question the tradition of imitation tasks being primarily elicited imitation in experimental settings, marked on accuracy.

Considering the idea that social motivation could be an important factor in autistic imitation deficits, how does this square with findings of reduced imitation accuracy in older sample of children with autism? It could be argued that this account would predict reduced spontaneous imitation and increased refusals, rather than less

accurate performance. Deficits in intrinsic motivation may be more likely to affect younger, less cognitively able participants than older and more cognitively able participants. In imitation experiments with older "high-functioning" autistic participants, is there really evidence that these participants are less motivated?

In fact, having older participants happy and willing to do imitation tasks but performing them badly is not in disagreement with *early* deficits in social motivation. Potentially, a deficit in social motivation early in life may alter the trajectory of a child's imitation development. Indeed, recent proposals for the development of mirroring (mirroring here really pertaining to the neural phenomenon rather than behavioural mirroring) have taken a neuroconstructivist approach, arguing that early experiences interact with an experience-expectant system, leading to the specialisation of certain regions to support mirroring (Quadrelli & Turati, 2015). (This account has many parallels with the development of specialised face areas in the visual system, and perceptual narrowing in early language development.) An early deficit that interferes with this process (such as a deficit in motivation that reduces the appropriate input to a developing mirroring system) could arguably prevent the proper development of these mirroring systems. Thus, even if later on in development a child is willing and motivated to do an imitation task, this disruption could leave lasting effects on the systems underlying imitation, limiting the ability to accurately imitate. The notion that a primary deficit in social motivation could underpin the abnormal development of mirror neuron systems and imitation in autism is also in keeping with other, more generalist accounts of mirror neuron development, such as the associative sequence learning theory (see Heyes, 2016 for an associative sequence learning account of imitation). In autism, a motivational deficit may prevent the developing system receiving the usual

input enjoyed by typically developing children, and thus imitation fails to be established in the normal way (Heyes, 2001).

At present, there is tentative support for the notion that the amount of spontaneous imitation has associations with language development. Laakso et al. (1999) found that typically developing children's spontaneous imitation of mothers' object directed actions predicted language production longitudinally. Farrant et al. (2010) found that parental reports of imitation were correlated with reports of conversational skills, in two separate samples of typically developing children. Their analyses suggest that imitation mediates the relationship between social engagement and language, the opposite of the usual account, in which social engagement mediates between imitation and language. Possibly, more socially engaged children are more imitative, and children who imitate more invite more linguistic input from their caregivers, particularly linguistic input that relates to what the child is doing or seeing ("Are you cleaning like Mummy? Are you wearing a hat like Dadda?"). In studies of children with autism, behaviours associated with social engagement do seem to predict both imitation and language. For example, one study examined whether joint attention skills predicted language and imitation skills - high-level joint attention initiations were found to predict both, while low-level initiations did not (Pickard & Ingersoll, 2014). The authors suggest that these findings fit with the social motivation hypothesis of ASD, as high-level initiations are more reflective of social motivation than low-level initiations.

Considering this account, if social skills or social motivation provide the link between imitation and language, why do children with SLI show evidence of imitation difficulties, as these children are not traditionally considered to have the social deficits seen in ASD? While children with SLI have demonstrated reduced theory of mind abilities, problems with emotion recognition, and difficulties making and retaining

friends, language impairment could be reasoned to underlie all of these issues (Taylor, Maybery, & Whitehouse, 2012). However, it has been debated to what the extent difficulties these children experience socially can be directly put down to their limited language skills (Marton, Abramoff, & Rosenzweig, 2005). An alternative hypothesis could be that children with SLI actually have subtle social problems, independent of their language issues. Indeed, many imitation tasks would seem to have limited language demands, and could provide a window into these additional social problems, without the confounding effects of language difficulty. Furthermore, if one reflects on the predictions outlined earlier by the combined social motivation and neuroconstructivist accounts, the observed pattern of refusals by children with language delay in Dohmen et al. (2013) would suggest the development of mirroring systems in these children could be set down an abnormal trajectory, regardless of what causes these refusals in the first place.

In sum, despite limited work examining the role of social engagement and social skills in imitation and language directly, this account offers a promising framework for the interpretation of current data, and for future research questions.

2.8 Imitation and language in the brain – the mirror neuron system

The above theories take a cognitive approach to the link between imitation and language. An alternative view is that there are no real shared cognitive aspects of imitation and language, but imitation and language may have overlapping or adjacent neural substrates. Certainly there is evidence for anatomical links between imitation, gesture, motor skills and language. Both gesture and language seem to be supported largely by the left hemisphere. Left hemisphere damage is associated with apraxia and aphasia more so than right hemisphere damage (Renzi, Motti, & Nichelli, 1980).

Participants with unusual cerebral lateralisation for language show the same pattern of lateralisation for gesture understanding (Króliczak, Piper, & Frey, 2011).

Undoubtedly researchers' ongoing fascination with mirror neurons, and broader mirroring systems, continues to fuel interest into imitation. Mirror neurons, discovered originally in the macaque, were originally proposed to underlie action understanding, but the role of the mirror neuron system has expanded to encompass numerous higher level social and communicative processes, including speech and language (Gallese, 2008; Oberman & Ramachandran, 2007; Rizzolatti & Arbib, 1998; Rizzolatti et al., 1988). Mirror neurons have been explicitly linked to the motor theory of speech perception – the notion that speech sounds are perceived by listeners mapping the sounds they hear onto the motoric gestures used to produce sounds themselves (Liberman & Mattingly, 1985). In fact, mirror neurons have not just been suggested to underpin the perception of speech sounds, but also language comprehension - these systems may allow for an "embodied" view of language (Fischer & Zwaan, 2008; Gallese, 2008). Furthermore, a dysfunction in the MNS has been suggested to underlie some of the core features of ASD, including *both* language problems and imitation (Oberman & Ramachandran, 2007). Research using transcranial magnetic stimulation also suggests Broca's area, traditionally considered a speech area, may also have a role in imitation (Heiser, Iacoboni, Maeda, Marcus, & Mazziotta, 2003). Further discussion about the theories linking the mirror neuron system to language processes is in Chapter 5.

Investigating such hypotheses is more suited to neuroimaging methods than the behavioural studies reviewed here. However some predications can be made about what we expect from the behavioural data, if this theory were true. It could be predicted that if imitation and language occupy neighbouring regions of the brain, but are not

cognitively overlapping, then little association should be observed in typical development beyond the effects of general brain maturation. Close links may be observed in cases of neurodevelopmental disorders, in which the brain does not develop typically, and in cases of brain damage. Certainly some researchers have suggested that in ASD or SLI imitation and language show different correlations than typical controls (Vukovic et al., 2010). Furthermore, if imitation and language were tightly coupled in the brain, we might also expect to see patterns of association – when a given disorder shows a deficit in one skill, the other will be poor as well. Double dissociations on the other hand would not be expected (however, see Bishop (1997) for a discussion on some of the issues of dissociation in developmental disorders).

2.9 Challenges for the field of imitation and language

There are several potential lines of enquiry for future research into the mechanisms linking imitation and language development - however, a number of issues that will need to be addressed were also noted in the literature, largely relating to how imitation is operationalised.

The first issue relates to how much of the association between imitation and language can be accounted for by broad developmental factors, such as age and IQ. As noted earlier, imitation does not always correlate strongly with intelligence, and so it seems unlikely that these maturational factors can explain the whole story. Nonetheless, while many developmental studies use standardised measures of language (which mean that the effects of age on differences between individual children's language skills are controlled for), standardised imitation batteries do not exist, and therefore measures of imitation do not take into account the potential effects of age. Where studies have recruited participants from a narrow age range, this is less problematic, but the confounding effects of age and development are amplified in studies with very wide age

ranges. Also, in early development even age ranges of a few months can leave a study open to the confounding effects of age, as young children show steep developmental trajectories in their early years. Some of the correlational studies had very wide age ranges in their participants, which can make studies more prone to the effects of general development.

The development of normed tests for children's imitation would be an extremely useful tool for future work in this area. Indeed, a related issue is the lack of study of imitation in typical development in its own right. Understanding how imitation normally develops, and how it usually relates to language and other areas of development, is key if we are to understand how to interpret the imitation deficits in atypical groups. The recent development of the PIPS (Preschool Imitation and Praxis Scale) is an excellent step forward in this area, although this scale does not currently extend up to school-age children (Vanvuchelen et al., 2011).

Another difficulty in evaluating the role of motor skills in imitation is that many motor measures involve imitation to some extent – the experimenter may demonstrate a motor task and then the participant will perform it. Disentangling the relationship between motor and imitation skills will require selection of motor tests and batteries that are not confounded by reliance on imitation. This point is particularly pertinent to the motor theory, but it is interesting to note that while researchers have suggested a role for social engagement and representational skills in imitation, imitation itself has also been used as a measure of these things; deferred imitation is used as a measure of children's mnemonic abilities, and imitation is often included in scales of young children's social development. Arguably, imitation's diverse usage as a developmental research tool is testament to the number of sub-skills it clearly depends upon.

Furthermore, as discussed previously, historically the study of imitation and how it relates to other skills has largely been concerned with the accuracy of a child's elicited imitation performance. Theories such the social motivation hypothesis would seem to suggest that the amount of spontaneous imitation is also extremely important, and poses a challenge to the ways imitation tasks are currently administered and measured.

Finally, a very pervasive issue in developmental research on imitation is the field's broad definition of the concept. While most would agree that imitation refers to cases where an individual copies the behaviour of a model, cases where the behaviour of two individuals match are not *necessarily* imitation. For example, Tomasello (1996) outlines the process of stimulus or local enhancement - seeing a model act with a particular location or object in the environment guides another's attention to these locations or objects, and on attending to these locations or objects the observer may interact with them. This would lead to matching behaviour between the model and the observer, but strict definitions of imitation would not call this "true" imitation. As noted by Sevlever and Gillis (2010), copied behaviour in a child does not necessarily imply that the child is capable of imitation, and while comparative psychology has prised apart true imitation from low level copying, human imitation research has rarely made this distinction (and the imitation deficits in autism are not usually scrutinised at this level). In the majority of studies considered in this review, the operationalisation of imitation could really be argued to refer to broad behaviour matching, not "true" imitation as defined above. It is worth considering how far the proposed relationship between imitation and language really reflects a relationship between broader *behaviour matching* and language.

2.10 Imitation deficits outside of the autism spectrum

Studying children with communication difficulties could provide a means to tease apart some of the potential nuances of the relationship between these language and imitation. This review specifically considered the literature from two atypical groups – children with ASD (who have very heterogeneous language outcomes) and children with SLI. The evidence from investigations of children with ASD or SLI was selected because while both these groups of children have communication difficulties, arguably the proposed deficits are quite different. However, rather than imitation difficulties patterning with particular communication impairments, imitation problems are seen in all groups, including children with selective structural or pragmatic language problems, and children with both. This leaves a puzzle; if imitation was largely related to pragmatic language skills, or broader nonverbal communication skills, why are imitation deficits seen in SLI, children for whom these skills are supposed to be relatively intact? Equally, if imitation is related to structural language skills, why do we see imitation deficits across the spectrum, not just in children with autism and impaired language skills? This would seem to raise the possibility that imitation and language are related at their broadest levels. Alternatively, perhaps imitation is quite vulnerable in development, though if so we would expect to see imitation difficulties in a range of developmental conditions – comparisons between ASD and a wide variety of other developmentally disordered groups have suggested imitation difficulties to be relatively specific (e.g. Dewey et al., 2007, Rogers et al., 2003).

It could be that using SLI and ASD with and without language difficulties as a means to contrast structural versus broader pragmatic language problems is problematic because in reality the boundaries between these two groups of children is not clear cut. While historically distinct groups, it has been proposed that these two conditions may share some cognitive deficits or genetic risk factors (Tager-Flusberg & Joseph, 2003).

However, others have argued that similarity between ASD and SLI may be relatively superficial, rather than stemming from a shared aetiology (Bishop, 2010). More detailed characterization of the imitation difficulties in children with language difficulties and comparisons with children with ASD could assist in understanding the relationship between these two conditions further.

Specifically, based on emerging evidence that children with SLI face continued social difficulties, it may be that shared imitation difficulties in these groups reflect some overlapping social abnormalities. A number of children diagnosed with SLI later meet the criteria for an autism spectrum disorder (Conti-Ramsden, Simkin, & Botting, 2006), and some children with SLI also show deficits in nonverbal communication and social cognition (Bishop, Chan, Adams, Hartley, & Weir, 2000; Marton et al., 2005). Traditionally, the social difficulties in SLI have not been considered central to the disorder, but instead social problems are usually considered secondary to structural language problems. However, it has also been argued that these difficulties are not simply a result of language problems but represent an additional and independent deficit (Marton et al., 2005). Potentially, the difficulties with imitation in SLI could be related to elevated levels of autism symptomatology and social difficulties.

Alternatively, imitation deficits are in keeping with previous reports of common comorbid motor weaknesses in these children (Hill, 2001). However, this same theory has also been investigated in ASD, and several reports have highlighted that basic motor difficulties cannot account for the full picture – thus, we should be careful not to write off imitation impairments in SLI as simply a manifestation of motor problems.

Working memory has also been a focus of investigation in SLI (e.g. Archibald & Gathercole, 2006), particularly phonological working memory, assessed using nonword repetition tasks. Complex imitation can be expected to place demands on working

memory, as the imitator has to remember the model's actions well enough to copy them. While Broca's area has been considered a site underpinning imitation, gesture and language, it also has been suggested to underpin subvocal rehearsal (Baddeley, 2003), so perhaps even the neural overlap between imitation and language could be in part explained by working memory. Indeed, working memory has been linked to imitation deficits in schizophrenia (Matthews, Gold, Sekuler, & Park, 2013). The hypothesis that working memory deficits may underlie imitation difficulties in SLI was tested by Marton (2009), but it was not found to contribute significantly to language-impaired children's imitation. However, nonverbal IQ and working memory were highly related, so the correlation between imitation and working memory was essentially partialled out when the authors controlled for nonverbal IQ. Motor imitation in SLI has still received very little research, but considering these children's working memory abilities as a potential factor in their imitation skills could yield interesting results in future work, particularly given the extent to which working memory has been investigated in these children's verbal imitation skills.

Given the evidence considered here, examining imitation abilities in other groups with interesting communication profiles could help further our understanding of the relationship between imitation and language. Specifically, children with Social Communication Disorder (SCD) or Pragmatic Language Impairment (PLI) demonstrate difficulty with the pragmatics, but do not meet criteria for autism spectrum conditions. Comparing imitation in SCD and ASD would seem to allow a means to investigate whether imitation difficulties relate to autism symptomatology or more specifically social-communicative deficits. Another group that would be interesting to consider would be children with Williams Syndrome (WS). Children with WS are usually regarded as having very good language skills, considering their delays in many other

aspects of development, and have very high levels of social interest (hypersociability), although their social skills are still abnormal (Laws & Bishop, 2004). Children with WS exhibit uneven language profiles, with relatively good concrete vocabulary but poor referential language, grammar and pragmatics (Mervis & Velleman, 2011), and therefore the study of imitation skills in relation to language in WS could be particularly interesting. Furthermore, given theories that imitation problems in autism may have a root in social motivation deficits, examining imitation in WS could yield interesting results, as these children could be argued to have abnormally high levels of social motivation.

2.11 Summary

In sum, associations between imitation and language have been consistently documented, but as yet there is little definitive evidence for the mechanisms that link imitation and language. Few studies have attempted to directly examine what mechanisms link imitation and language. Nonetheless, from reviewing the literature there are several lines of enquiry that can be pursued, and also several issues that will need to be addressed, pertaining especially to the measurement of imitation. Currently, it seems that very general factors are insufficient to explain the association between these abilities, and there are likely shared cognitive aspects between the two skills. Considering children's praxis and social abilities seem like promising lines of enquiry. Imitation difficulties are not exclusive to ASD, but have been documented in children with other forms of communication impairment. While indisputably imitation deficits in autism are theoretically interesting and have the potential for being clinically useful, considering the deficit as unique to autism maybe problematic for fully understanding the deficit both in autistic populations and in other populations. For children with SLI, poor imitation could be providing vital evidence about other underlying difficulties. Our

knowledge about imitation and language currently is lacking, both for typically developing children, and children for whom imitation has been proposed to be a viable target for intervention. There are many fruitful avenues for future research on imitation and language, both for uncovering the mechanisms behind this association, and for better understanding populations for whom imitation and language present difficulties. There has been limited direct comparison of motor imitation skills in autism and SLI to date, though current evidence suggests that this may be an area of overlap between the conditions. Understanding the links between language and imitation may be enhanced by comparing these groups. In subsequent Chapters 3 and 4, I report a cross-disorder comparison of verbal and motor imitation skills in autistic and non-autistic language-impaired children.

CHAPTER THREE: Nonword Repetition in Autism Spectrum Disorders and Language Impaired Children

3.1 Introduction

As described in Chapters 1 and 2, Autism Spectrum Disorders (ASD) and Specific Language Impairment (SLI) could be conceived as two very different conditions. Broadly, "SLI" describes the problems of a child when there is a mismatch between verbal and nonverbal skills, a discrepancy which cannot be attributed to a known physical or environmental cause. In effect, SLI can be thought of as condition in which a child displays poor language for no obvious reason. Previously, children with SLI had been thought to be particularly impaired in the structural aspects of language (phonology, morphology, and syntax), but to be otherwise normally developing communicators, with intact nonverbal communication skills. ASD, or autism, is a condition in which a child shows markedly poor social and communication skills, and restricted and repetitive interests. The language prognosis of a child diagnosed with autism is hugely varied; many are nonverbal, or have very limited language (Sigman & McGovern, 2005). While the emphasis in SLI research has traditionally been focused on these children's structural language problems, children with ASD are seen to be particularly impaired on the pragmatic aspects of language, and to have pervasive communication problems, beyond structural language difficulties.

Until the latest publication of the DSM-V, one of the exclusionary criteria when diagnosing SLI was the presence of an autism spectrum disorder, making clinical comorbidity technically impossible (American Psychiatric Associaiton, 1994). However, three broad lines of evidence suggest there are similarities between ASD and SLI. Firstly, it has been argued that some children with an ASD exhibit language

difficulties very similar to those with SLI (Kjelgaard & Tager-Flusberg, 2001). Secondly, children diagnosed originally with SLI sometimes go on to display features of autism, especially social and pragmatic difficulties (Conti-Ramsden, Simkin, & Botting, 2006; Leyfer, Tager-Flusberg, Dowd, Tomblin, & Folstein, 2008). Finally, family studies have suggested that there is a higher incidence of language problems in relatives of autistic probands, and a higher incidence of autism in relatives of language-impaired probands, which would suggest there are shared genetic risk factors for these disorders (Folstein & Rutter, 1977; Folstein et al., 1999; Le Couteur et al., 1996; Tomblin, Hafeman, & O'Brien, 2003 - but see Bishop, 2010 and Szatmari, Jones, & Tuff, 1993).

3.1.1 Nonword repetition performance in SLI and ASD – the phenomimicry account

Of particular importance in the debate about the relationship between SLI and ASD have been examinations of these children's performances on nonword repetition tasks. Nonword repetition tasks simply involve a child listening to a nonsense word and repeating it back immediately to the experimenter. These tasks have been widely used in the study of SLI, and are considered to be clinical markers for the condition (Bishop, North, & Donlan, 1996). Typically developing children show decreased performance on nonword repetition tasks as words become longer, but word length has a particularly strong effect on the performance of children with SLI. This has been taken as evidence for poor phonological short term memory (PSTM) abilities in these children, a deficit that could provide a causal account for children's language impairments (Gathercole & Baddeley, 1990).

Previous reports argued that children with autism and language impairment (ALI) also performed poorly on nonword repetition tasks, fuelling theories that a genetic subtype of autism related to SLI exists (Kjelgaard & Tager-Flusberg, 2001).

However, closer examination of nonword repetition performance has suggested that while both ALI and SLI do indeed show poor nonword repetition performance, the underlying causes of these difficulties may be quite different. Whitehouse, Barry, and Bishop (2008) examined nonword repetition in ALI and SLI, and found children with SLI exhibited a stronger length effect than children with ALI, a finding later replicated by Riches, Loucas, Baird, Charman, and Simonoff (2011). Specifically, both these studies found an interaction between group (ALI and SLI) and word length, and when qualifying this interaction found the groups performed the same at words of two and three syllables in length, showed a trend towards different performance at four syllable words, and showed significantly different performance when repeating five syllable words.

ALI children have thus shown comparatively flat nonword repetition performance compared to children with SLI – the increasing length of the nonwords appears to affect them less. This has led to the suggestion that perhaps while repetition deficits in SLI are dependent on poor PSTM, this is not the case for ASD. Indeed, Williams, Payne and Marshall (2012) point out that short term memory is meant to be a relative *strength* in ASD. Echolalia, and the repetition of phrases that children do not understand, could be essentially conceived as nonword repetition. Poor PSTM is not in keeping with these commonly observed behaviours in ASD. Williams et al. (2012) conducted the first study of nonword repetition in ALI and SLI to include a group of typically developing controls matched to the disorder groups for verbal mental age. The ALI group showed no more of a deficit in nonword repetition than this younger control group; they did not differ on overall nonword repetition accuracy, the task's relation to

children's other memory abilities, effects of consonant clusters or length[3]. Given the similarity of the ALI and verbal age matched controls, Williams et al. (2012) argue that the language impaired children with ASD are developmentally *delayed*, but not *deviant*, in their nonword repetition performance. Children with SLI on the other hand are indeed deviant, showing deficits compared to the verbal mental age matched children.

Together, these findings support a phenomimicry account of the relationship between SLI and ASD. This account argues that the two disorders may show superficial similarities in their language abilities, but that language problems arise in these two groups of children for very different reasons (Bishop, 2010). The investigations into nonword repetition have suggested that while children with ALI do indeed exhibit poor nonword repetition, the pattern of these children's errors lacks the strong length effect shown by children with language difficulties (Riches et al., 2011; Whitehouse et al., 2008.However, stronger length effects were observed for the ALI group in Williams et al., 2012). This has been taken as evidence that the nonword repetition difficulties – and thus *language* difficulties – in autism have a different aetiology to those in SLI.

The source of nonword repetition difficulties in ALI is still unclear. Williams et al. (2012) argue that these difficulties in autism are quite straightforwardly related to language delay (thought they provide limited suggestion for what the primary cause of the language delay is in autism). Others have suggested that the flat performance in ALI could be due to broad imitation problems, attentional difficulties, or autism severity (Whitehouse et al., 2008). All of these hypotheses would be aided by comparing

[3] Although these authors only examined performance for three- and four-syllable words, they still found a significant length by group interaction. However, length was found to affect the performance of the SLI group *less* than the ALI and typical control groups, the opposite to what previous investigations found. This was argued to be due to the poor performance of SLI even when words were only three-syllables long.

children with ALI and children with autism and normal language (ANL). If nonword repetition performance is appropriate for children's language level, then children with ANL should not be expected to show different patterns of nonword repetition compared to typically developing children. Neither Riches et al. (2011) nor Williams et al. (2012) included an ANL group, and Whitehouse et al. (2008) do not examine the error patterns of their children with ANL, compared to children with SLI or ALI (and even if they had done, they had no typically developing controls to compare the performance of the ANL group to, meaning it would not have been possible to examine whether the nonword repetition abilities of the ANL group were normal). An investigation into language skills in relatives of ALI, ANL and SLI did find that ALI and SLI probands show nonword repetition deficits comparative to an ANL group, but no analysis is presented of the effects of word length (Lindgren, Folstein, Tomblin, & Tager–Flusberg, 2009) (and as language-impairment was in some cases classified by individuals showing nonword repetition deficits, these differences could arguably be due to the somewhat circular analyses).

Only one recent study of nonword repetition in autism and SLI has included an ANL comparison group. Hill, van Santen, Gorman, Langhorst, & Fombonne (2015) considered nonword repetition, and the verbal and nonverbal memory skills in their groups. Intriguingly, they found that the largest differences between the ALI and SLI groups were at the *shortest* word lengths, rather than the longer lengths found in previous studies. The authors argued that these differences supported the notion that the deficits in these two groups had separate aetiologies. However, compared to the ANL group, both children with SLI and ALI exhibited verbal memory difficulties (as assessed by digit span tests and narrative memory tests), and verbal memory difficulties were argued to play a role in the language impairments of the ALI group. Given that

verbal memory difficulties are usually used to explain the nonword repetition problems (and indeed *language* problems) in SLI, finding comparable memory difficulties in ALI does not fit well with the notion that ALI and SLI do not share verbal memory problems, and these verbal memory problems could not similarly underlie NWR problems in ALI.

While useful in understanding how nonword repetition patterns with language difficulties in autism, the paper by Hill et al. (2015) does not include a group of typically developing children, which limits the conclusions that can be made about whether children with autism and normal or good language show any abnormalities in their nonword repetition abilities. A useful comparison would be to have four groups, representing children with language problems but without autism, language problems and autism, autism and normal language skills, and children without autism or language problems.

3.1.2 Nonword repetition and social motivation

Another factor that could underpin poor but flat nonword repetition in ASD is a lack of motivation to engage in such a social game, or a difficulty or unease with social stimuli. This could lead to reduced performance, regardless of word length. Intriguingly, for typically developing children it may be that increasing the socialness of verbal imitation tasks improves performance - the manual of the much-used Children's Test of Nonword Repetition (CNRep) advises against live administered nonword repetition tasks (rather than using a pre-recorded tape), because this may inflate children's performance (Gathercole, Willis, Baddeley, & Emslie, 1994). This effect has been presumed to be because of acoustic differences between live and recorded performance (during live administration the mouth of the examiner is covered up, to prevent extra

visual cues boosting nonword repetition performance), however it is possible that better performance is a result of social factors related to live interaction. Indeed, theories about early word learning suggest that social interaction has a critical role in gating language development (Kuhl, 2007). These theories pertain to language development in infancy, but it is possible that social interaction continues to have important effects later childhood. A live social partner may be more engaging, boosting attention and motivation, and therefore performance.

While social interaction may confer benefits to language learning and other aspects of development in typical children, these benefits may be lacking in ASD. The social motivation hypothesis outlined by Chevallier, Kohls, Troiani, Brodkin, and Schultz (2012) suggests that an underlying difference in how motivated children with autism are to attend and respond to social stimuli could underpin the social difficulties seen in the condition. This account cites evidence of reduced attention to speech sounds (but not non-speech sounds), selective social anhedonia, and reduced audience effects on the behaviour of ASD individuals. All of these observations would seem relevant social factors that could influence children's performances on nonword repetition tasks, and could explain reduced performance in autism.

3.1.3 Aims of the current project

Given the centrality of nonword repetition in the SLI literature, understanding nonword repetition deficits in autism better could inform our understanding of language problems in either condition. The primary aim of this project was to replicate findings that children with ALI show a different pattern of errors in verbal imitation tasks to children with SLI. As well as replicating previous work on nonword repetition, it was hoped that this project could illuminate some of the underlying causes of nonword

repetition difficulties in ALI. The project aimed to compare children with autism and language problems and those with normal language for their age, and consider how nonword repetition patterns with language in ASD. Unlike previous studies, four groups of children were included – ALI and ANL, non-autistic language-impaired children, and a group of non-autistic children with normal language.

A further aim of this project was to examine children's performance in different versions of nonword repetition tasks, manipulating the amount of social engagement required. Demonstrating such effects would highlight that autistic children's difficulties with this task are not to do with phonological working memory problems, but more to do with broad motivational or attentional factors. Several potential predictions could rationally be made about the effects of manipulating social task demands. Firstly, it may be that while typically developing children benefit from increasing social interaction, children with autism would show no difference between a social and non-social nonword repetition game. Alternatively, perhaps a discomfort with social interaction in some children with autism could be detrimental to nonword repetition performance; it could be expected that nonword repetition tasks that require further attention to social stimuli – such as a live social partner – should further compound these children's difficulties with nonword repetition. For children with language impairment, it was predicted that these children would show similar effects to typically developing children, given that these children are not broadly considered to have the social attention and motivation difficulties considered of ASD.

3.2 Method

3.2.1 Data collection

There were three primary methods of data collection. In the majority of cases, data was collected by visiting the children at school, during school hours. For a few children, data was collected at home, during holidays or weekends. For a proportion of the typically developing children, data was collected during visits from families to the laboratory, during the half term holidays as part of an activity week organised the at Oxford University's Experimental Psychology Department, "Oxford Science Adventures". Much of the recruitment was done by distributing information to mainstream and specialist schools. Interested families sent back their contact details to the researcher, and after a telephone interview with parents to ensure the children were suitable for the study and to arrange obtaining written proof of informed consent, children were visited at school or home to conduct the behavioural tasks. For children who were recruited through a half term activity week, the week was advertised through email through parent mail lists, and University child care emailing lists. Flyers were also distributed at local schools and social groups. Families signed up for a morning or afternoon session of activities at the department, including developmental psychology experiments, and neuroscience themed crafts and games.

3.2.2 Participants

Children aged between 6 and 12 years were recruited for the study. A total of 115 children were seen. Six children were excluded at the testing stage; two children with ASD and two with LI were excluded due to noncompliance, and two further children with LI were seen, but excluded as these children were very globally impaired and could not do the majority of the tasks. This then left 59 typically developing (TD) children, 22 LI, 22 ASD. Five children (4 TDs and 1 ASD) did not complete all of the nonword repetition tasks, and are therefore not included in the analyses (however, these children are included in a later chapter of this thesis examining motor imitation skills).

To be considered "language impaired" children had to demonstrate a score of less than 1.5 standard deviations (SD) below the average norm for their age on at least 2 of the language tasks. This equates to a *t*-score of 35 or lower, or a scaled score of 6 or lower. The language tasks included sentence repetition, vocabulary, or concepts and directions subtests (see Section 3.4.3 *Procedure and Task Battery* for details of the tasks). The CNRep was not included as a means to classify the children, as nonword repetition was one of the study's dependent variables so group classification based on this measure would be circular. As a large number of the language-impaired children showed evidence of nonverbal cognitive weakness (nine had *t*-scores of 35 or lower on the nonverbal skill task), this group is referred to as the "LI" group rather than the "SLI" group.

With these thresholds, of the original 22 children grouped as "language impaired" at the recruitment stage of the study, eight did not show sufficient language difficulties to meet these criteria. However, three children who entered the study as part of the typically developing control group *did* show these deficits. For these three children, the descriptions given by the families at the initial screening questionnaire were indicative of subtle language difficulties, and literacy problems. One had a formal diagnosis of dyslexia, and for another the family had concerns about dyslexia and reported that their child had been very late to start reading and writing. The third had a history of speech and language therapy and delayed speech development, although a formal diagnosis was never made, and had needed continued help with reading in school. It was deemed appropriate to regroup these three children as language impaired.

To be classified as in the autism group, children needed to have been diagnosed or be in the process of being diagnosed with autism. In addition, families of children in the autism group completed the 3di autism interview (see Section: *Procedure and Task*

Battery for more details on the 3di). Two of the children previously diagnosed with autism by a clinician did not meet the criteria for autism according to the 3di (one ALI participant and one ANL participant). However, these children are retained in our analyses. We applied the same criteria for classifying children as language impaired to our autism group (two language subtests 1.5 SD below the age norm), splitting our autism group into ANL (normal language, $N = 13$) and ALI (language impaired, $N = 8$).

As a whole, the remaining 52 children in the typically developing group significantly differed from the LI group on nonverbal ability (as measured by the WASI matrix task). Therefore, to extract a group of TD children with comparable nonverbal abilities, a screen was used to select only TD children who achieved a raw score of 18 or lower on this task (this corresponds to the mean plus 1 SD of the LI group). This left a group of 14 TD children, who did not significantly differ from the LI group in terms of age or nonverbal ability (both raw and standardised), at the group level. The characteristics of the final four groups are shown in Table 3.1. There was one sentence repetition task missing from 1 TD participant and 1 LI participants, due to technical problems. Two children with ASD (1 ANL and 1 ALI) did WASI-2 as part of separate project very close to the time of data collection – it was felt more appropriate to use the standard scores from this task rather than to re-administer a highly similar task the children would have recently done. The raw scores for these cases were not entered into the mean raw score calculation for these groups. Table 3.2 shows the 3di scores for social reciprocity, communication, nonverbal communication and repetitive behaviours for the two autism groups. The two groups differed on their restrictive and repetitive behaviours (t (20) = -2.18, p = .042,) but this difference does not survive correction for multiple comparisons. The ALI and ANL groups are therefore assumed to have comparable severity of autism symptoms.

Table 3.1: Child sample characteristics

Variables		*Participant group*			
		TD (*N*=14)	**ANL (*N*=13)**	**ALI (*N*=8)**	**LI (*N*=17)**
Age (Months)		112.36 (7.61)	105.08 (21.24)	112.00 (27.20)	111.35 (20.51)
Gender		11 F, 3 M	12 M, 1 F	3 F, 5 M	12 M, 5 F
WASI Matrix	*Raw scores*	15.71 (3.71)	16.69 (8.17)	11.00 (6.85)	11.29 (7.19)
	Standard scores	45.29 (7.12)	49.54 (12.82)	38.38 (19.91)	39.38 (11.10)
WASI Vocab.	*Raw scores*	37.29 (8.47)	32.00 (11.89)	16.38 (6.05)	18.82 (7.21)
	Standard scores	54.00 (9.55)	51.69 (11.51)	27.00 (5.55)	30.69 (8.14)
CELF Concepts & Directions		9.57 (2.21)	8.69 (3.17)	1.50 (0.93)	3.12 (2.03)
NEPSY Sentence Repetition		10.79 (2.72)	10.00 (3.08)	3.25 (2.55)	3.81 (2.01)

Table 3.1: *Characteristics of the final sample. With the exception of gender, this table shows the means and standard deviations for each group on the measures shown.*

Table 3.2: 3di Scores

3di Scale	*Participant group*	
	ANL (*N*=13)	**ALI (*N*=8)**
Social Reciprocity *(Min. with clinical significance = 10)*	15.3 (4.28)	15.7 (4.52)
Communication *(Min. with clinical significance = 8)*	16.0 (2.69)	14.0 (4.34)
Nonverbal Communication *(Min. with clinical significance = 7)*	8.7 (2.12)	7.6 (3.87)
Repetitive Behaviours *(Min. with clinical significance =3)*	7.8 (1.96)	5.6 (2.77)

Table 3.2: *Mean and standard deviations for the 3di scores of the two ASD groups.*

3.2.3 Procedure and Task Battery

This study received approval from the ethics committee at the University of Oxford (Medical Sciences Interdisciplinary Research Ethics Committee Code: C2-2013-02). Table 3.3 summarises the tasks the children did. For children whose data were collected during school, the tasks were conducted over a period of two to three days to avoid disruption to the children's learning. Data collected during half term activity weeks were collected in an hour's session with the researcher. The task battery was kept to a minimum in order to retain children's attention. Several task subtests were drawn from standardised assessment batteries including the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999), the Clinical Evaluation of Language Fundamentals (CELF -IV) (Semel, Wiig, & Secord, 2006), and the NEPSY (A Developmental NEuroPSYchological Assessment) (Schmitt & Wodrich, 2004).

Table 3.3: Task battery

Language tasks	*Nonword Repetition Tasks*	*Other Tasks*
WASI Vocabulary NEPSY Sentence Repetition CELF Concepts and Directions CCC *(Completed by parents)*	Intercom Game (CNRep words) Co-Spy game Guard game	WASI Matrix NEPSY Hand imitation NEPSY Fingertapping Audiogram Mature Imitation Task 3di (*ASD groups only – completed by parents*)

Table 3.3*: Tasks completed by the children.*

3.2.3.1 Language tasks

Expressive vocabulary skills were assessed using the WASI Vocabulary task, a task in which children are asked to define words ("Can you tell me what a "car" is?"). Comprehension was assessed using the Concepts and Directions subtest from the CELF. In this task, children are shown an array of pictures, and asked to point to the pictures in

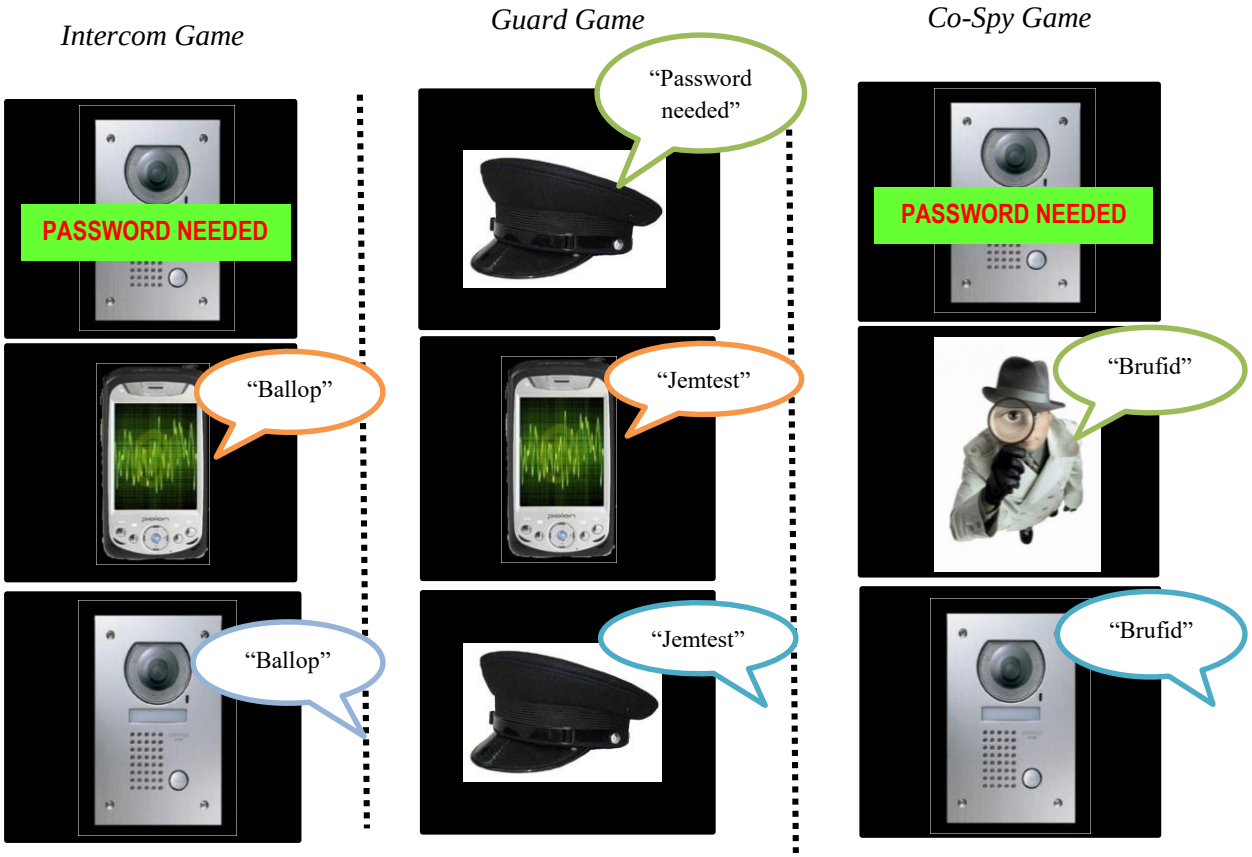
the order the experimenter says. A sentence repetition task (a subtest taken from the NEPSY) was also included. This task is designed to assess children's abilities to repeat sentences of increasing complexity and length. Typically, this test is administered with the experimenter saying the sentences, however as the majority of the nonword repetition tasks were presented with recorded stimuli, for parity the sentence repetition test was also presented using pre-recorded stimuli. Parents of the children also filled in the Children's Communication Checklist (CCC -V1), which yields an index of the children's pragmatic skills (Bishop, 1998). The CCC was originally included as a means to exclude any children in the typically developing group who exhibited abnormal scores on their communication skills. However, the rate of return for the paper questionnaire was not good, and the data from it are thus not presented.

3.2.3.2 Nonword repetition tasks – the Spy Password Games

The nonword repetition tasks required the children to listen to a nonword, either presented by the laptop, or live from the researcher. They then repeated the word, either to the researcher, or to the laptop. As each list of non-words was 40 items long, and there were three contexts being examined, it was felt that the task would need to be sufficiently engaging to avoid the children becoming bored, not trying and lowering their performance. Therefore, these tasks were presented in the context of an engaging spy password game (a theme picked because it was felt it would engage both younger and older children in the sample). In these games, the children were told they must repeat passwords (nonwords) to open safes and unlock pieces of a top-secret map.

There were three versions of this game. The "Intercom" version of the game used the original words from the CNRep, as this is the version that best matches the conditions under which the CNRep is normally administered. In this game, the child

was told to listen to the passwords (the nonwords) told to them by the computer, and then repeat them to the computer. There was effectively no social interaction in this game. In the "Guard" version of the game, the child listened to the passwords the computer says and told them to the "guard" (the experimenter). In the "Co-Spy" version of the game, the child listened to the passwords their co-spy (the experimenter) says,



and tells them to the computer. Figure 3.1 displays these three tasks.

Figure 3.1: The three "Spy Password" Games. *The left column displays the intercom game, the centre column the guard game and the right column the co-spy game. Children either heard the password from their spy device (in the intercom and guard games), or from their co-spy (in the co-spy game). They then either repeated the password to the intercom (in the intercom and co-spy games), or the "guard" (in the guard game). Stimuli or prompts spoken by the experimenter are in green speech bubbles, those from the computer in orange, and the responses from the children in blue.*

As repeating the same words used in the CNRep seemed likely to lead to practice effects, two new nonword lists were constructed to closely match the CNRep. The nonword lists can be found in Appendix A: *Nonword lists used in verbal imitation tasks*. They were closely matched on syllable length, cluster position, embedded words, and stress pattern. To check that the average word likeness of the lists did not differ, the word likeness of the each of the words on the new lists and on the CNRep was also rated by 3 independent raters, who were not familiar with the CNRep. No significant differences were found between the average word likeness ratings of the lists. In addition, to check baseline performance on the 3 lists did not differ, 22 typical adults were also tested on the three nonword lists in a nonword repetition paradigm. Because it was anticipated that most adults would be at ceiling on this task, pink noise was added to the recordings to make the task harder. There was no significant difference in overall performance between the 3 lists, and no significant interaction effects between list and syllable (indicating that lists were matched overall, and the 4 syllable lengths within each list were also sufficiently matched). These analyses can be found in Appendix B: *Adult Nonword Repetition Pilot Analyses.*

3.2.3.3 Coding children's nonword repetition

Investigations of children's verbal repetition commonly code children's responses in a binary fashion – correct or incorrect. However, coding children's repetitions this way does not allow for distinction between children who get a repetition mostly right but perhaps repeat one phoneme incorrectly (e.g. "*dopelate*" repeated as "*topelate*") and children who misrepeat many or most phonemes (e.g. "*dopelate*" repeated as "*godebane*"). This may be of particular importance considering previous findings that children with ASD have a comparatively "flat" performance compared to children with SLI.

The Levenshtein Distance (LD) algorithm is able to calculate the "distance" between the correct response and given response by comparing responses phoneme-by-phoneme basis. The LD algorithm compares two strings and calculates the minimum numbers of operations (e.g. substitutions, additions or omissions) needed to transform one string to the other. This algorithm has been used by previous authors to achieve a better picture of children's errors (Riches et al., 2011). Each nonword repetition for each child was transcribed into phonetic script, and the LD algorithm applied to calculate a LD value for each response. Table 3.4 shows examples of the different LD values the algorithm gives depending on a child's repetition.

Table 3.4: Levenshtein Distance

	Correct Response	*Child A's Response*	*Child B's Response*
Orthographic Script	"Glistering"	"Glisterin"	"Giserin"
Phonetic Script	ɡlɪstɜrɪŋ	ɡlɪstɜrɪn	ɡɪsɜrɪn
Levenshtein Distance (compared to correct response)	0	1	3

Table 3.4*: Demonstration of outputs of Levenshtein Distance algorithm on individual nonword repetitions.*

3.2.3.4 Other tasks

Nonverbal reasoning and fine motor skills were also assessed, and a hearing screen conducted (500, 1000, 2000, and 4000 Hz at 25 dB). The WASI Matrix task was used as a measure of the children's nonverbal reasoning. In this task, children are shown a series of pictures, and asked to select a picture to complete the pattern. Children's fine motor skills were assessed using the NEPSY Fingertapping task, and tests of children's motor imitation were also administered, as an additional investigation with this sample was on motor imitation and language skills. Descriptions of these tasks can be found in

Chapter 4. To confirm that the children in the autism groups had current clinically significant levels of autism symptoms, the 3di ("The developmental, dimensional and diagnostic interview") (Skuse et al., 2004) was administered via telephone following behavioural testing.

3.2.4 Counterbalancing nonword lists and conditions

In the original sample, lists and conditions were broadly counterbalanced. However following removing and reclassifying some children, this was no longer the case. The lists were created to be close matches on a number of key features, but nonetheless the potential effect of list on nonword repetition was considered, using a non-parametric one-way ANOVA (Kruskal-Wallis test), as nonword repetition performance was not normally distributed. List was not found to have an effect on nonword repetition performance, both considering the overall performance for each condition, and at the level of the different syllable lengths. Thus, any differences that emerged between conditions were assumed not to be due to any differences between the word lists.

3.3 Results

The analyses of the nonword repetition performance took a two-stage approach; the effects of the different nonword repetition conditions for the different participant groups are considered first, followed by the effects of word length and participant groups. As the group sizes were uneven, multi-way ANOVAs were inappropriate, as uneven group sizes can lead to confounding of effects (i.e. in a two-way ANOVA with group membership and game condition or word length as factors, the effects of belonging to the ALI/ANL/LI/TD group and the effects of the different conditions or word lengths will be confounded). Because of this, both the effects of game condition

and word length are assessed using one-way ANOVAs. Analyses were conducted on children's sum of LD (for each condition and word length) rather than their average LD.

3.3.1 Effects of social task demands in the different nonword repetition games

The effects of the different game conditions (Intercom, Guard and CoSpy) are first considered. Boxplots in Figure 3.2 show the distribution of the sums of LD for the different conditions, for the different groups of children. As the group sizes were unequal, any ANOVA other than a one-way is not appropriate so instead, potential effects of condition were assessed using two one-way ANOVAs, on the differences between conditions. For each child, the sum of LD from the CoSpy game was subtracted from the Intercom game, and the sum of LD from the Guard game subtracted from the CoSpy game. One-way ANOVAs were then run on these differences to assess the effect of group membership of differences between games. As the difference variables were non-normal, non-parametric Kruskal-Wallis ANOVA were used. Neither found any group effects on differences between these conditions.

While these ANOVAs showed no group effects on the differences between game condition, they did not test a potential main effect of condition, across all groups. This was assessed using a Friedman's ANOVA. There was no main effect of condition.

Following this analysis, as no condition effects had been found, the nonword repetition performance for each child was pooled across the three game conditions. This meant that for the analyses on length effects, for each child there were 120 nonword repetitions, 30 at each syllable length.

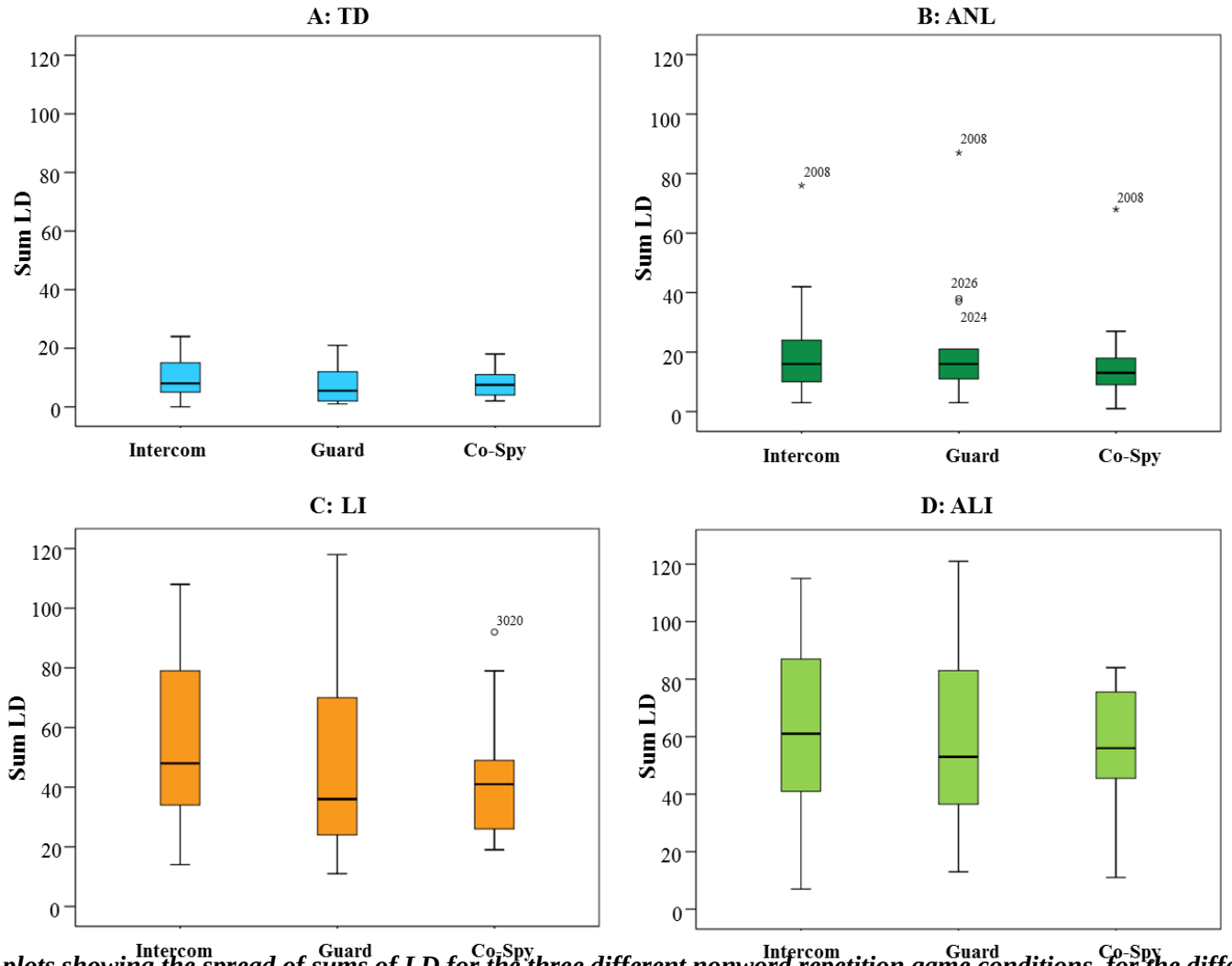


Figure 3.2: Box plots showing the spread of sums of LD for the three different nonword repetition game conditions, for the different groups. *The top two plots (A and B) include the groups with normal language status, and the bottom plots (C and D) include the two groups with language impairment. The left plots (A and C) are groups that are non-autistic, while the plots on the right are the autistic groups (B and D). These plots have been made to the same scale. Outliers are labelled with participant numbers.*

3.3.2 Effect of word length

Boxplots in Figure 3.3 show the performance of the different groups at different syllable lengths (pooled across game condition). Figure 3.4 shows the individual participants' performances at each syllable length. Table 3.5 shows the mean scores at each word length for each group. There were clear floor effects and a resultant positive skew to the data. A correct repetition results in an LD of zero, and the high number of zeros skewed the data. Indeed, there were a high number of children in the non-language impaired groups (TD and ANL) who could correctly repeat all of the two-syllable words, and many of the three syllable words (although the skew is clearly not limited to just these groups).

As a multi-way ANOVA was not appropriate, instead of entering word length as a factor with 4 levels (and group as another factor), for each participant the slope of the sum of LD was calculated. These slope values represent the gradient of the linear regression line that runs through the sum of LD for each word length. Figure 3.4 shows the individual plots for each participant – the slope value is effectively the regression line for each participant, running through the four values for the four different word lengths. These slope values were subjected to a one-way ANOVA. The variable slope and its residuals was significantly non-normal, and therefore a non-parametric ANOVA (Kruskal-Wallis Test) was used.

There was a significant effect of group on the slope of children's LD (χ^2 (3) = 27.33, p <.001). Post-hoc pairwise comparisons were done using Dunn-Bonferroni corrections, adjusted for multiple comparisons. These showed that the only significant group differences were TD versus LI (Dunn's z = -4.529, p <.001) and ALI (Dunn's z =

-4.109, p <.001). There was a trend for ANL to be significantly better than the LI (p =.073) and ALI (p =.081) groups.

Examining the plots of the individual children's data on Figure 3.4, one participant in the ANL group was noted to exhibit a particularly steep length effect, not comparable to the rest of the group. To assess whether this case may be an outlier, an outlier labelling rule was used, a version appropriate for non-normally distributed data (Iglewicz & Banerjee, 2001; Kimber, 1990). Using this rule, participants in the ANL group exhibiting a slope value above 19.95 would be considered outliers. The participant showed a slope of 36.7. No other participants, from ANL or other groups, were identified as outliers using this method.

Labelling and removing outliers from datasets (especially a dataset as small as this one) is contentious. However, because of the smallness of the dataset, an extreme outlier like this could exert a disproportionate effect on the results. This outlier case was therefore removed, and the analysis on the slope values re-run, to assess what effect this case could be having on the data.

Excluding this case did change the outcome of the post-hoc tests following the Kruskal-Wallis test of group effects on slope. The ANL group was now significantly different to both the LI (Dunn's z = -2.93, p=.021) and ALI groups (Dunn's z = -2.82, p =.029), after correcting for multiple comparisons. The overall effect of group was still significant: χ^2 (3) = 29.57, p <.001 (N=51).

Interestingly, the outlier case did have poor vocabulary, obtaining a t-score of 38. This did not meet the threshold used to classify impairment on a given task (t of 35 of lower), but suggested that perhaps vocabulary was a particularly important factor in determining the slope of a child's LD scores. As an exploratory analysis, the

correlations between the slope values and children's scores on the language tasks were compared, to see if vocabulary exhibited a special relationship with LD slope (using a technique described by Steiger, 1980). Standardized vocabulary scores correlated significantly with LD slope (r_s = -.676, p<.001), as did sentence repetition (r_s = -.764, p<.001), and performance on the concepts and directions task (r_s = -.639, p<.001). The correlations between slope and vocabulary, slope and comprehension, and slope and sentence repetition, did not significantly differ.

Table 3.5: Groups' nonword repetition performance

Word length	*Participant group*			
	TD	**ALI**	**LI**	**ANL**
Two syllables	1.86 (2.21)	10.75 (9.60)	9.88 (8.53)	3.77 (4.92)
Three syllables	4.21 (4.00)	29.88 (15.43)	24.76 (14.04)	7.77 (8.21)
Four syllables	6.71 (4.38)	58.88 (34.34)	45.06 (25.40)	17.31 (15.01)
Five syllables	12.79 (9.19)	79.00 (36.57)	68.41 (31.76)	31.62 (31.93)
Overall LD	25.57 (15.74)	178.50 (88.15)	148.12 (73.92)	60.46 (58.03)

Table 3.5*: Means and SDs for the sum of LD at each word length, for the four groups.*

3.3.3 Correlations between nonword repetition length effects and autistic symptoms

Previous investigations of language subtypes in ASD had suggested that children with ALI and poor nonword repetition could represent a group of children with more severe autism. The correlations between the 3di scores of the ASD groups and the slopes for nonword repetition performance were examined in an attempt to replicate previous findings. Spearman's rank was used to examine these correlations.

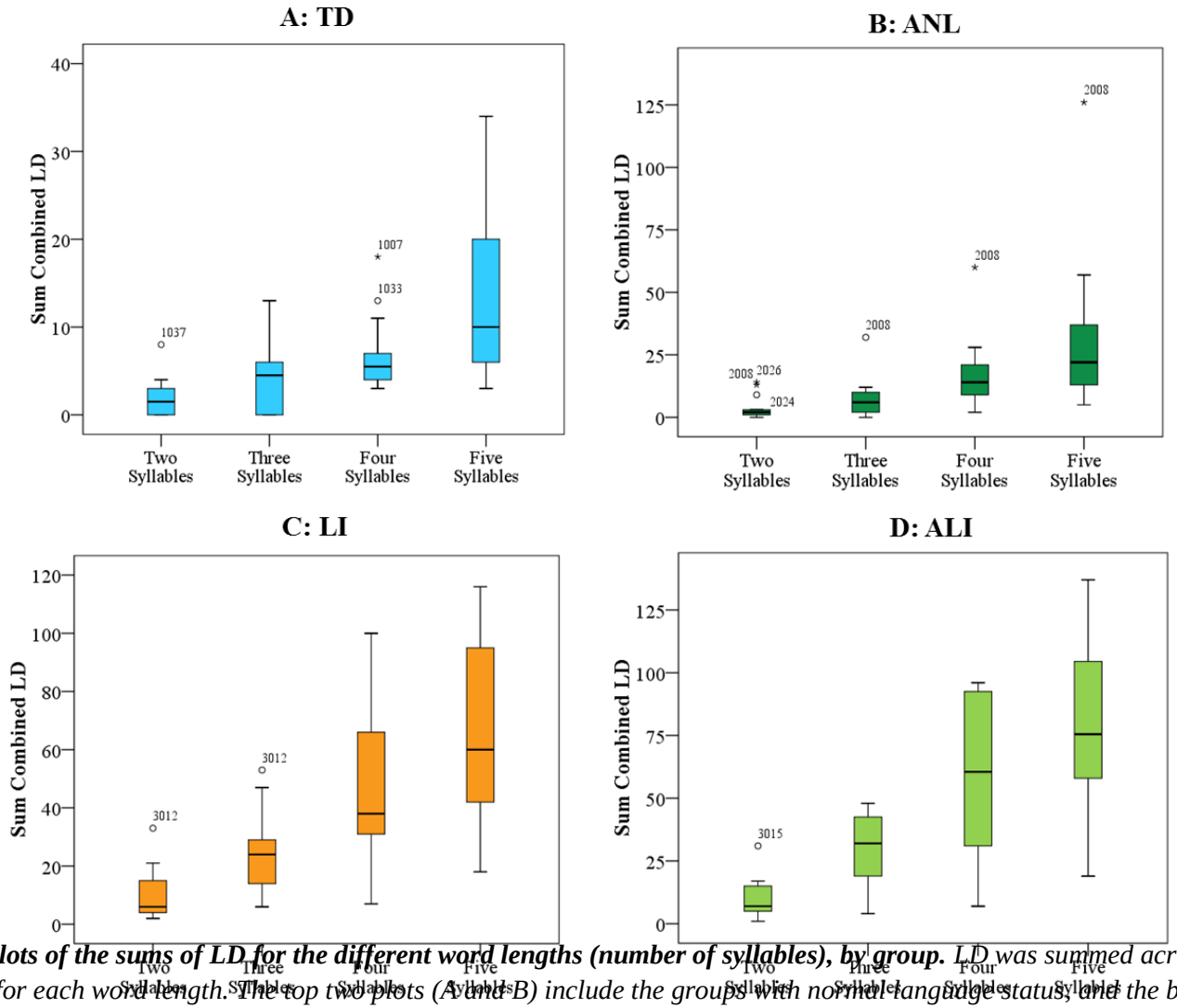


Figure 3.3: Box plots of the sums of LD for the different word lengths (number of syllables), by group. *LD was summed across the three nonword repetition games, for each word length. The top two plots (A and B) include the groups with normal language status, and the bottom plots (C and D) include the two groups with language impairment. The left plots (A and C) are groups that are non-autistic, while the plots on the right are the autistic groups (B and D). Note that the TD plot is on a different scale to the other three groups. Outliers are labelled with participant numbers.*

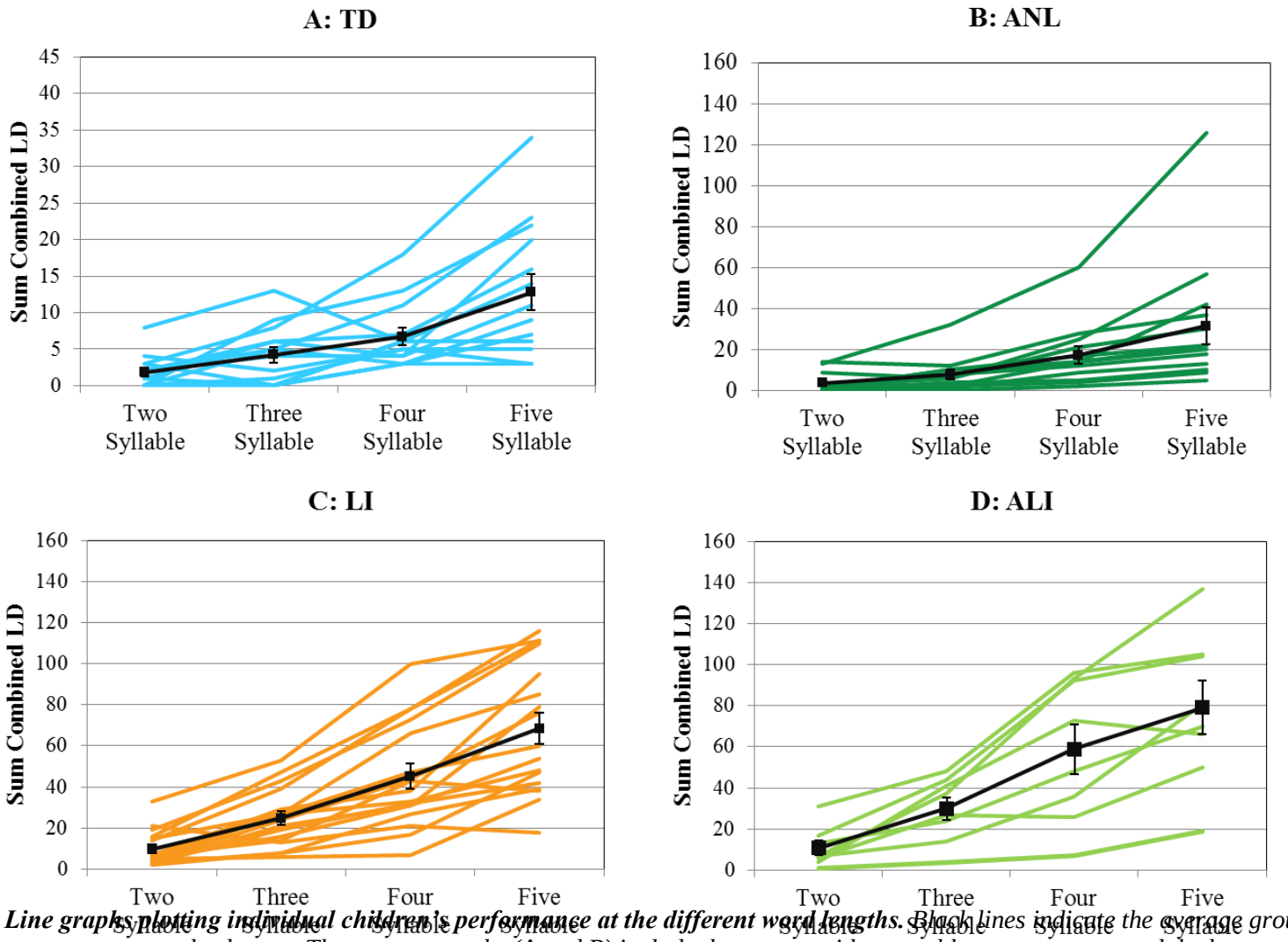


Figure 3.4: Line graphs plotting individual children's performance at the different word lengths. *Black lines indicate the average group performance, and error bars represent standard error. The top two graphs (A and B) include the groups with normal language status, and the bottom graphs (C and D) include the two groups with language impairment. The left graphs (A and C) are groups that are non-autistic, while the graphs on the right are the autistic groups (B and D). Note that the TD graph is on a different scale to the other three graphs.*

The four subscales of the 3di (social reciprocity, communication, nonverbal communication and repetitive behaviour) showed limited correlations with the slope of nonword repetition performance. However, repetitive behaviour did show a significant correlation (r_s = -.538, p =.012, N = 21), and this would survive corrections for multiple comparisons (.05/4= .0125). Nonetheless, this does not support the hypothesis that the ALI group show nonword repetition problems due to more severe autism symptoms, as the ANL group actually had higher repetitive behaviour 3di scores (see Table 3.2).

3.4 Discussion

Previous research had suggested that length effects in nonword repetition tasks in non-autistic language-impaired samples were stronger than for children with autism and language impairment, suggesting distinct aetiologies to nonword repetition problems and broader language difficulties. Yet in our sample of autistic, language impaired and typically developing children, length effects in nonword repetition patterned with language status, regardless of autism. Children in the LI and ALI groups showed no significant differences on the slope of their LD scores, and similarly children in the ANL group showed no differences from the typically developing group. The graphs for the TD and ANL groups' performances in Figure 3.4 (Graphs A and B) indicate that at the shortest words many children achieved LDs of zero or near zero, indicating perfect or near perfect performance. The language-impaired groups showed significantly greater slopes than the TD group, and the ANL group (following the removal of one outlier who showed an unusually strong length effect). These findings are in contrast to previous reports that find an interaction between ALI and LI groups, wherein autistic children show flatter performance and reduced effects of word length (Riches et al., 2011; Whitehouse et al., 2008; Williams, Payne, & Marshall, 2012).

The secondary hypothesis of this experiment was that manipulating the social task demands of nonword repetition tasks would affect the performance of groups differently. However, no main effects of game condition nor interactions between condition and participant group were observed. Hearing the nonwords live, or having to repeat the nonwords to a live listener, did not seem to confer any advantage or disadvantage to typical, autistic or language-impaired children.

Three potential factors could explain the different results for length effects obtained in this study compared to previous investigations. The first is the amount of nonword repetition data this study collected per child. Williams et al. (2012) collected data over 64 nonword repetitions per child, Riches et al. (2011) 68 nonwords per child, and Whitehouse et al. (2008) 13 nonwords. In this study, children performed 120 nonword repetitions. Potentially, this increased amount of data could have allowed for length effects to emerge in the ALI group.

The second factor that could underpin differences between this and previous studies is the nature of the children in the groups. While at the group level, there were no significant differences between the typically developing group and language impaired groups on the average performance of the nonverbal cognitive task, both the non-autistic language-impaired children and ALI children included in this study had nonverbal weaknesses. Of all of the language-impaired children originally seen, almost a third of the children did not show what could be regarded as sufficient language-impairments to be retained in this group. However, it was these children who showed nonverbal cognitive skills above the threshold usually used to exclude children with nonverbal weaknesses. Excluding children both for failing to show sufficient nonverbal skills and persistent language impairments would have severely reduced the number of children in the language impaired group. Thus, there was a dilemma – either children

who had weak nonverbal skills were excluded and children whose language impairments were not necessarily bad enough were retained, or children without enough evidence of a persistent language problem were excluded, even though these children had reasonable nonverbal skills. The latter option was selected, but other recent experiments of children with SLI have done the former (for example Wray, Norbury, & Alcock, 2015). The majority of the children grouped as ALI also showed nonverbal skills below the 1.5 SD threshold.

If this is the reason for these findings differing from previous accounts, this would mean that children with autism and language difficulties and children without autism and language difficulties look more similar on nonword repetition tasks, if their nonverbal skills are poor. By extension then, as nonverbal skills improve in the two groups, nonword repetition performance begins to differ. Logically, this would be caused either by the language-impaired children exhibiting flatter performance when these children have nonverbal weakness, or children with ALI exhibiting less flat performance when they have poorer nonverbal skills. Did the LI group in this study show flatter performance than would be expected for an *S*LI group? This seems unlikely to be the case, given the large slopes seen in this group. Could reduced nonverbal skills explain the length effects seen in the ALI group? Considering this idea, presumably for ALI performance to go from steeply length-affected to flat as nonverbal skills improve, ALI children with better nonverbal skills would perform better on shorter nonwords (as opposed to worse on long words – it is presumed better nonverbal skills would boost overall performance, not reduce it). It is difficult to conceive of a reason for why this would be. Usually, one would assume that poorer cognitive skill would pattern with flatter performance due to potential floor effects (and the flatter performance of the ALI groups in other studies certainly is not reflective of ceiling effects – the children are still

making errors, and showing nonword repetition deficits). Therefore, although a key difference between our groups and previous groups is nonverbal ability, it is hard to account for our pattern of findings based on this alone.

The third possibility is that broad motivational factors have affected the performance of the children with autism in previous investigations, leading to flat error patterns - but the embedding of nonword repetition tasks into engaging contexts in the current study has reduced this effect. The nonword repetition tasks used in this study were presented as games, designed to be appealing to children of this age group. Originally, the games served as a means to explore potential effects of increased or decreased social interaction on nonword repetition performance – the conditions were so highly similar in other respects, it was felt that presenting the task devoid of a creative story could have risked children becoming bored of nonword repetition tasks, and affecting performance. However, setting nonword repetition in this context may have boosted the performance of children who may otherwise have disengaged and shown flat performance. Given that all nonword repetition games were presented in these themed contexts, there is no "standard" nonword repetition task to compare these gamified versions too, so this hypothesis is admittedly speculative. However, the children did seem to enjoy the games, were happy to play the different versions (despite the games being highly similar). Indeed, given that these games were quite repetitive and predictable, they may actually have been particularly well-suited to children with autism.

What implications do these findings hold for theories concerning the relationship between autism and SLI? Previous reports had suggested that different error patterns in nonword repetition tasks provided evidence that nonword repetition difficulties arose from separate aetiologies in the two conditions. However, the results

of the present study – that the ALI and LI groups show comparable slopes in the number of phonemic errors as word length increases - instead indicate that these groups both have impaired phonological short term memory, and arguably support the notion that SLI and ALI have overlapping cognitive deficits.

An alternative framing of the nonword repetition problems in the language impaired groups would be that poor nonword repetition performance may be a *consequence* or *reflection* of a language problem, rather than pointing to the causal cognitive problem. Associations between language ability and nonword repetition performance have been traditionally interpreted thus; poor nonword repetition reveals a problem in phonological short term memory, and poor phonological short term memory impedes vocabulary and grammatical development (Gathercole & Baddeley, 1990). However, other authors have also highlighted that associations between language and nonword repetition or PSTM problems are not necessarily unidirectional. Vocabulary size has been shown to affect nonword repetition performance (Munson, Kurtz, & Windsor, 2005). It may be that a child with better vocabulary has finer phonological representations, leading to bolstered phonological processing and better performance on tasks that use it, including nonword repetition. Conversely, children with weak vocabulary have poorer phonological representations as a result, and perform worse on such tasks (Snowling, Chiat, & Hulme, 1991). In a similar line of enquiry, it has been shown that learning to read radically changes the way children process speech sounds, and that reading predicts performance on nonword repetition tasks, rather than the other way around (Nation & Hulme, 2011).

This account would seem to predict a special relationship between nonword repetition length effects and vocabulary size. As one outlier in the ANL group who showed quite a dramatic slope in their nonword repetition performance had weak

vocabulary, the correlations between slope and the different language subtests were compared. While the correlations between slope and the different language tasks were all strong, the correlations were not significantly different from each other. Thus, there was no evidence in this dataset of a special relationship between length effect and vocabulary skills, though slope was clearly related to language skills more broadly.

Taking this view, nonword repetition performance in the groups studied here could be argued to reflect simple differences and similarities in language level. The children with LI and ALI have poorer language skills, and as a result perform more poorly on nonword repetition tasks, compared to typically developing children and autistic children with intact language skills. This interpretation would still allow for children with ALI and SLI to have language impairments for quite distinct reasons, and could still fit with the phenomimicry account. What it would argue however is that examining nonword repetition in school-aged children with ALI and SLI may be problematic, as poor language development in these children may lead to these two groups appearing similar on nonword repetition tasks. Longitudinal work would be best placed to resolve these questions, such as that undertaken by Nation and Hulme (2011).

Considering now the lack of effects of social task demands on nonword repetition performance, it was somewhat surprising that no advantage for live administration was seen, even for typically developing children. While previous observations suggested that children tended to do better with live administration, our data suggest this does not lead to significant differences. It may also be the case that any advantage that might have been conferred by live administration was much smaller than the boosted performance of embedding nonword repetition assessment in an enjoyable game.

3.4.1 Limitations

The nature and size of the sample is this study's biggest weakness. The group sizes are clearly very small. Furthermore, the non-autistic language-impaired children had nonverbal cognitive weaknesses as well as language difficulties, meaning than not all researchers would agree that this sample represents a group of children with "*Specific* Language Impairment". Indeed, although the ALI and ANL groups do not differ in terms of their raw nonverbal cognitive scores, their standard scores are significantly different. Thus, do the ALI versus ANL groups in this study represent language impaired versus not language impaired, or far broader "low functioning" versus "high functioning" groups?

Another limitation with the current study is that ideally all families of children with LI should have undertaken the 3di interview, in order to ascertain whether similarities between the autistic and non-autistic language impaired groups could have been reflective of the fact that children with language impairment *not* diagnosed with autism nonetheless often exhibit high levels of autism symptoms, and are at increased likelihood of meeting the clinical diagnostic criteria for autism (Conti-Ramsden, Simkin, & Botting, 2006).

The use of slope to examine length effects in this study also presents some challenges when comparing the results of this study to previous ones. The use of the slope measures was adopted because it allowed the analysis to use a one-way ANOVA, which is robust to effects of unequal group sizes. Unequal group sizes in factorial ANOVAs can lead to confounding. However, previous investigations of nonword repetition performance of different group at different syllable lengths that had unequal

group sizes *have* used mixed ANOVAs to analyse their results (Riches et al., 2011; Whitehouse et al., 2008).

3.4.2 Summary

Overall, these findings suggest that children with language impairment and children with ALI show similar length effects on their nonword repetition performance, when these tests are presented in a way that is appealing to the children. The difference between our findings and those of previous investigations may have been due to motivational factors in the nonword repetition performance of children with ASD. No effects of social task demands were found on nonword repetition performance.

CHAPTER FOUR: Motor imitation in children with autism and language difficulties

4.1 Introduction

Chapter 3 considered the nonword repetition performance of children with autism and children with language impairment. To account for the nonword repetition problems seen in children with ALI, some have suggested that these children have a general difficulty with imitation (Whitehouse, Barry, & Bishop, 2008). If so, this would suggest verbal and motor imitation are quite strongly related, and jointly dependent upon some broader imitation factor. However, as noted in Chapter 2, there is limited evidence for whether these two types of imitation do indeed relate to one another. Rodgon and Kurdek (1977) studied gestural imitation (including object-based and body-based actions) and vocal imitation (including single and repeated consonant-vowel syllables, and single words) in young typically developing children aged 8, 14 and 20 months. They found no significant correlations both between and *within* types of imitation. This would suggest, at least at this early age and in typical development, that subtypes of imitation are relatively independent of one another, but how imitation subtypes are associated later in development, or in cases of abnormal development, is not clear. Potentially, verbal and gestural imitation could become more related over the course of development, or could exhibit stronger relations in disorders such as ASD, where imitation skills are impaired.

The review in Chapter 2 noted that while imitation is known to be an area of difficulty in autism, several investigations have suggested that imitation may also be poorer in children with SLI. It is unclear at present whether these difficulties are really comparable to those in autism, and whether poor imitation could represent an area of

shared deficit between these two overlapping conditions. While separate studies of imitation in ASD and SLI have demonstrated deficits, the only study identified that directly compared these groups of children on their imitation skills was Stone, Lemanek and Fishel (1990). Stone et al. (1990) found imitation deficits in ASD, but intact imitation in their group of language impaired children[4].

However, since this study, a number of different research groups have reported imitation problems in SLI (Brookman, McDonald, McDonald, & Bishop, 2013; Dohmen, Chiat, & Roy, 2013; Marton, 2009; Vukovic, Vukovic, & Stojanovik, 2010). In fact, some reports of imitation deficits in language-impaired samples bare strong resemblance to those in autism. Dohmen et al. (2013) reported poor imitation of meaningless body movements, but spared imitation of instrumental actions, a profile that has been reported in several studies of imitation in autism, and has been highlighted in several reviews and theories about the imitation deficits in the condition (Hamilton, 2013; Stone, Ousley, & Littleford, 1997; Williams, Whiten, Suddendorf, & Perrett, 2001). Hamilton (2013) suggests that the imitative deficits in autism fit with a neuropsychological model of imitation, in which there are two routes by which imitation can operate – a direct and an indirect route. In the indirect route, actions to be imitated are first parsed in terms of their goals and meaning, and then the action is reconstructed, based on achieving the same goal or meaning. Meaningless actions are supported by the direct route – the link between the visual information of the action being viewed is mapped directly to the motor commands needed to reproduce it. Hamilton (2013) suggests that goal-based actions are imitated properly by individuals with autism, but imitation dependent upon the direct route shows abnormalities. More

[4]While the language-impaired group are not explicitly called "SLI", the description of these children arguably fits the criteria– they were required to show IQ scores above 70, and not meets ASD criteria on the CARS.

specifically, she postulates that top-down control and social modulation of imitation supported by the direct route is impaired, pointing to the prevalence of echolalia in autism as evidence for an abnormality in the *control* of imitation, not its accuracy.

Another observation about imitation in autism has been that the styles of the actions are not always imitated, even if the action overall is reproduced (Hobson & Lee, 1998; Hobson & Hobson, 2008). Hobson and Hobson (2008) tested object-based imitation in children with and without autism. They found that even though there were no differences in whether children with autism imitated (or attempted to imitate) the action itself, the autism group imitated the style of the action (gentle or harsh) significantly less than the mental-age control group. Hobson and Hobson (2008) interpret these findings in terms of limited capacity for intersubjective engagement in autism. Arguably such interpretations are quite comparable to Dohmen et al.'s (2013) conclusion that children with language problems failed to connect socio-emotionally with the model to imitate body movements.

4.1.1 Aims of the project

The aim of this project was to examine the relationships between motor and verbal imitation and language ability, in typically developing children, and in children with autism and/or language impairments. Specifically, it aimed to examine two questions. The first was whether children with language impairment show the same pattern of imitation deficits as children with autism, including particularly poor body movement imitation, and difficulties copying the stylistic quality of the actions (Hobson & Hobson, 2008; Williams, Whiten, Suddendorf, & Perrett, 2001). Secondly, the project considered whether imitation problems could account for the nonword repetition difficulties in some children with ASD. If imitation problems underlie nonword

repetition difficulties in ALI, one could predict that children with ALI will exhibit poorer imitation skills than children with ANL, and we would predict correlations between motor and verbal imitation. The extent to which correlations between verbal and motor imitation could be explained by joint relationships with broader motor and cognitive factors was also examined.

4.2 Method

4.2.1 Participant characteristics

The characteristics of the sample are summarised in Table 4.1. As before in Chapter 3, the children in the TD group were a sub-sample of a larger number of typically developing children seen, selected in order to provide a group matched on nonverbal cognitive skills to the other groups. In addition to the children in Chapter 3, five additional typically developing participants and one additional autism-normal language participant are included in the group comparisons of motor imitation (they were excluded from verbal imitation analyses in Chapter 2 as they did not complete all of the nonword repetition tasks).

4.2.2 Procedure

Details of recruitment, data collection and general experimental procedures are described in Chapter 3. The children were assessed on a number of language measures, three nonword repetition tasks, nonverbal cognitive ability, fine motor skills, and motor imitation tasks (described below).

Table 4.1 Child sample characteristics

Variables		*Participant group*			
		TD (*N*=19)	**ALI (*N*=8)**	**ANL (*N*=14)**	**LI (*N*=17)**
Age (months)		104.00 (16.58)	112.00 (27.20)	105.79 (20.58)	111.35 (20.51)
Gender		12 F, 7 M	3 F, 5 M	1 F, 13 M	5 F, 12 M
WASI Matrix	*Raw Scores*	13.32 (5.44)	11.00 (6.85)	16.43 (7.91)	11.29 (7.19)
	Standard Scores	44.89 (6.42)	38.38 (19.91)	49.07 (12.44)	39.38 (11.10)
WASI Vocab.	*Raw Scores*	34.26 (9.13)	16.38 (6.05)	32.00 (11.89)	18.82 (7.21)
	Standard Scores	53.58 (9.45)	27.00 (5.55)	52.00 (11.11)	30.69 (8.14)
CELF Concepts and Directions		10.05 (2.46)	1.50 (0.93)	8.93 (3.17)	3.12 (2.03)
NEPSY Sentence Repetition		10.44 (2.68)	3.25 (2.55)	10.21 (3.07)	3.81 (2.01)
NEPSY Finger Tapping		11.74 (1.85)	8.83 (2.56)	11.00 (2.48)	8.81 (2.14)

Table 4.1: *Characteristics of the final sample. With the exception of gender, this table shows the means and standard deviations for each group on the measures shown.*

Table 4.2 3di scores

3di Scale	*Participant group*	
	ANL (*N*=14)	**ALI (*N*=8)**
Social Reciprocity *(Min. with clinical significance = 10)*	15.2 (4.13)	15.7 (4.52)
Communication *(Min. with clinical significance = 8)*	15.9 (2.59)	14.0 (4.34)
Nonverbal Communication *(Min. with clinical significance = 7)*	8.9 (2.13)	7.6 (3.87)
Repetitive Behaviours *(Min. with clinical significance =3)*	7.8 (1.89)	5.6 (2.77)

Table 4.2: *Means and standard deviations for two autism groups on the 3di sub-scales.*

4.2.3 Motor Imitation Tasks

Three tasks of motor imitation were included. Two of these tasks were from the Mature Imitation Task battery (Rogers, Cook, & Greiss-Hess, 2005). These subtests included the imitation of body movements (BM) and imitation of actions with objects (AO). The entire MIT battery includes numerous others imitation sub-tests, but these two tasks were specifically selected because it was felt they best mapped onto the direct and indirect routes in neuropsychological models of imitation, and were comparable to previous investigations examining object-based and body-based imitation in younger samples of autistic and language-impaired children (Dohmen, Chiat, & Roy, 2013; Stone, Ousley, & Littleford, 1997)

In these two tasks, the children watched videos of a man performing either meaningless body movements (for example, tapping one's arm below and then above the elbow), or actions with objects (such as striking a xylophone, or rolling play-dough with a rolling pin). After each clip, the children were asked to immediately imitate what they saw, as closely as they could. The children's responses were filmed and scored offline by two trained scorers who were blind to the children's group membership and the hypotheses of the study. The task was scored according to the unpublished manual for the Mature Imitation Task. The scoring codes the children's accuracy of the imitations on a number of set parameters. For the body movement tasks, the start and end points, the movement (e.g. arm moves across horizontal plane), and posture change (e.g. fingers closed to fingers together) are all marked, and each imitation trial has a maximum score of six. For the actions with objects task, the imitation attempts are marked on children's grip of the tool, whether the tool is used appropriately, the movement, and also whether the children copy the style in which the action is performed, with a maximum score of four. For three of the actions, the style is

performed gently. In the other three, they are performed harshly. Both the body movement and actions with objects tasks had six trials each.

The third imitation task was the standardised hand imitation task from the NEPSY (Schmitt & Wodrich, 2004), a task which has previously been found to be poor in children with language impairment (Brookman, McDonald, McDonald, & Bishop, 2013). In this task, the experimenter demonstrates a hand posture and asks the child to copy them. A correct imitation within 20 seconds is scored as correct. There are twelve postures, tested for each hand separately.The NEPSY hand imitation task differs from the two tasks from the MIT in several key ways; the demonstrator is live; the imitation targets are postures rather than movements; and the target to be imitated remains on show for the children to refer back to, so there are reduced working memory demands in this task. Working memory is considered a deficit in SLI, and could be a limiting factor in these children's imitation abilities. It was reasoned that if this was the case, the language-impaired sample may perform poorly on the MIT tasks, but show comparably intact imitation in the NEPSY hand imitation task.

4.2.4 Fine motor task

The fine motor skills task was the NEPSY finger tapping test. In this task, there are two stages. In stage one, the child is asked to tap their index finger and thumb together, as fast as they can, with their dominant hand. The experimenter times how long the child takes to tap 32 times. The test is then repeated with the non-dominant hand. In stage two, the child is asked to tap their thumb and each of their other fingers (index, middle, ring and then little finger) in a recurring sequence. As before, they are asked to do this as quickly and as accurately as they can. The experimenter times how long the child takes to complete 8 sequences. This is then repeated with the non-

dominant hand. The timings for both stages are summed, and this corresponds to a standardised score.

The data for the MIT tasks is missing for one typically developing child, due to technical failure at recording. Furthermore, one child in the ANL group found the NEPSY hand imitation and finger tapping tasks very frustrating, and these tasks were not completed. One child in the ALI group similarly could not complete the finger tapping task, instead tapping all fingers together rather than specific fingertips, and their data are not included.

4.3 Results

4.3.1 Inter-rater reliability

The two blind coders scored equal numbers of typically developing, language-impaired and autistic children. Twelve of the overall sample of children (before exclusion due to language impairments not being strong enough to be included in the LI groups, and selection of the TD group to match the nonverbal skills of the other disorder groups) were coded by both scorers to check for agreement between the scorers. This subset included two TD, five LI, and five ASD children. Intra-class correlations (ICCs) were calculated for scores on the body movement and actions with objects tasks. Inter-rater reliability is considered fair when the ICC value is between .40 and .59, good when between .60 and .74, and excellent when between .75 and 1.0. (Cicchetti, 1994; Hallgren, 2012). For the body movement imitation task, single-measure ICC was .919, and for actions with objects, this was .656 – thus, there was excellent agreement on the body movement task, and good agreement on the actions with objects task.[5]

[5] While the agreement for both tasks was adequate, it is interesting to consider why agreement was so different for the two tasks. One possibility was that scoring the actions with

4.3.2 Groups' performance on motor imitation tasks

The means and standard deviations for the groups' performances on the three imitation tasks are shown in Table 4.3. Analysing the groups' performances together in one mixed factorial ANOVA was not deemed appropriate as group sizes were unequal. Instead, group effects on these tasks are analysed separately in three one-way ANOVAs. For the hand imitation task, a non-parametric ANOVA is used (the Kruskal-Wallis test), as this variable and its residuals were not normally distributed.

Performance on the body movement imitation task is considered first. Levene's statistic was significant for the imitation of body movement scores, indicating that variance within the groups was not equal. Because of this, Welch and Brown-Forsythe tests were used to assess whether the groups were significantly different, as these tests are robust to the effects of unequal variance. No significant group effects were identified (Welch's adjusted F (3, 21.91) = 2.14, p =.124).

For performance on the actions with objects task there was a significant effect of group (F (3, 59) = 3.456, p =.023). Bonferroni post-hoc tests indicated that ALI group were significantly worse that TD group (p =.048). No other significant differences were seen, although there a marginal trend for the ANL group to perform worse that the TD group (p = .08).

objects tasks included a very subjective judgement on whether the child imitated the style of the model appropriately, while other judgements (grip, tool use and movement) were perhaps easier to score objectively. To investigate this, ICCs were also calculated for the different parameters in the actions with object task (grip, movement, tool use and style). Scoring on all these different parameters had good to excellent agreement, but agreement on the style aspects of the actions was much lower than the other three criteria. This would suggest that for this task, judging the style of a child's imitation was harder to score objectively than other aspects of the movement, and led to less excellent overall agreement than for the body movement imitation task.

For raw scores on the hand imitation task, there was a significant group effect on the performance of this task: $\chi^2(3) = 21.13$, $p<.001$ ($N = 57$). Pairwise comparisons using Dunn-Bonferroni tests to correct for multiple comparisons indicated that while the language impaired groups were significantly different from the non-language impaired groups, children in the ALI and LI groups did not differ from each other, and children in the ANL and TD groups did not differ from each other: ALI versus TD – Dunn's z = 3.52, p =.003; ANL versus ALI - Dunn's z = 3.18, p = .009; LI versus ANL - Dunn's z = 2.84, p = .027; LI versus TD - Dunn's z = 3.31, p =.006

Table 4.3 Groups' performances on imitation tasks

Imitation task	*Participant group*			
	TD ($N = 19$[1])	**ALI** ($N = 8$)	**ANL** ($N = 14$[1])	**LI** ($N = 17$)
Body movement (BM)	64.20 (13.03)	45.83 (22.37)	61.90 (22.89)	55.88 (12.88)
Actions with objects (AO)	71.30 (7.94)	58.85 (9.03)	61.61 (12.68)	65.44 (11.86)
Hand imitation (HI)	22.74 (2.21)	18.13 (3.64)	22.31 (3.77)	18.76 (4.37)

Table 4.3: *Means and SD of the groups' performance on the three imitation tasks. The means for hand imitation are based on raw scores (the maximum score for this task is 24). The MIT has not been normed – these values represent percentage.* [1]*Data missing for one TD participant on BM and AO, and one ANL participant on the hand imitation task.*

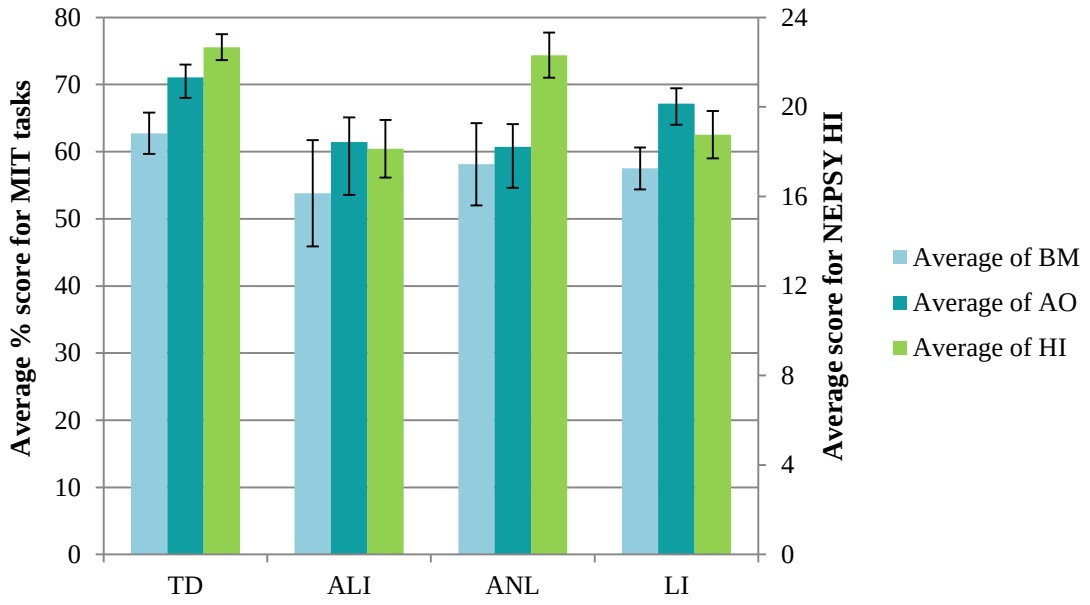


Figure 4.1: The groups' mean performances on the three imitation tasks. *Error bars represent standard error. Note that hand imitation and the two subtests from the MIT are on different y axes. MIT scores are percentage scores, and NEPSY HI score are raw scores.*

4.3.3. Imitating the style of the model

Given previous accounts of an autistic deficit in copying the style of actions, the groups were compared on this specific parameter of the actions with objects task. As there were six actions with objects trials, children could score between 0 and 6 on the style parameter, making it arguably an ordinal variable. Median style score was 5 for the TD group, 4 for the ANL group, 3 for the ALI group and 4 for the LI group. A Kruskal-Wallis test was used to examine group effects on the style score. No significant group effects were found: $\chi^2(3) = 4.82$, $p = .185$, $N=57$.

4.3.4 Correlations between motor imitation and verbal imitation

As one of the aims of the study was to examine the links between motor and verbal imitation, the correlations between performance on the motor imitation tasks and nonword repetition tasks was calculated. For this test, the sum of the LD score across all conditions and syllables was used as a measure indicating overall nonword repetition

ability. Nonword repetition tasks were skewed, thus Spearman's Rank was used. As the correlation between three imitation tasks and nonword repetition was being assessed, the corrected threshold for a significant association is .05/3=.017. These correlations were assessed across the whole sample, with two-tailed tests. Scatterplots of the data are in Figures 4.2 to 4.4. As can be seen from Figure 4.4, there were clear ceiling effects for the NEPSY hand imitation task.

As the two MIT tasks are not normed, raw scores of the NEPSY hand imitation task were used for parity. However, using only raw scores could potentially mean that correlations between verbal and motor imitation would be driven by age effects. To account for this, age is partialled out from the correlations.

Nonword repetition correlated significantly, and moderately to strongly, with the three motor imitation tasks. The correlations are negative, indicating higher LD scores correlate with lower imitation scores – because high LD scores reflect increased errors in nonword repetition, this suggests that poor nonword repetition is associated with poor motor imitation. Partialled for age, for hand imitation, r_s= -.654, p <.001. For body movement imitation, r_s = -.407, p =.004. For actions with objects, r_s = -.431, p =.002.

Two other variables were considered as potential mediators of the relationship between nonword repetition and motor imitation. These included nonverbal cognitive skills and fine motor skills, as assessed by the WASI Matrix and NEPSY finger tapping tasks respectively. As before, because the MIT tasks are not standardised, partialling the relationship between these imitation tasks using standardised scores for the matrix and finger tapping tasks would mean that only age effects on the partialled variables would have been controlled. Instead, raw matrix and raw finger tapping scores were partialled, in addition to age.

The zero order, first and second-order partialled correlation coefficients are shown in Table 4.4. Partialling fine motor skills did reduce the correlations between verbal and motor imitation, but both correlations with hand imitation and actions with objects remained significant after correcting for multiple comparisons. Partialling out nonverbal cognitive skills reduced the association between body movement imitation to non-significance, however the correlations between hand imitation and verbal imitation, remained significant.

Table 4.4 Correlations between verbal and motor imitation tasks

	Zero order correlations	**First order partials: age**	**Second order partials: age, fine motor skills**	**Second order partials: age, and nonverbal skills**
NEPSY Hand imitation	-0.604***	-0.654***	-0.499***	-0.556***
MIT Body Movements	-0.425**	-0.407**	-0.304*	-0.160
MIT Actions with Objects	-0.449***	-0.431**	-0.401**	-0.242

Table 4.4: *Correlation coefficients between verbal and motor imitation tasks. The verbal imitation task was nonword repetition, and the measure is the sum of LD (a measure of phonemic distance between correct and incorrect responses), pooled across three nonword repetition tasks (described in Chapter 3). Coefficients are Spearman's Rank. Significance is indicated thus: *p< .05, **p< .017, ***p< .001. Zero order correlations are between imitation scores (percentage score for MIT tasks, raw scores for NEPSY HI) and the sum of LD. First order partials are partialled for age in months. Second order partials are partialled for age and raw scores on the NEPSY finger tapping task, or the raw scores on the WASI Matrix task. Note that there were two children who did the WASI-2 – their raw scores are not included in the correlation considering nonverbal skills, as the raw scores of the WASI-2 are not comparable to the older WASI.*

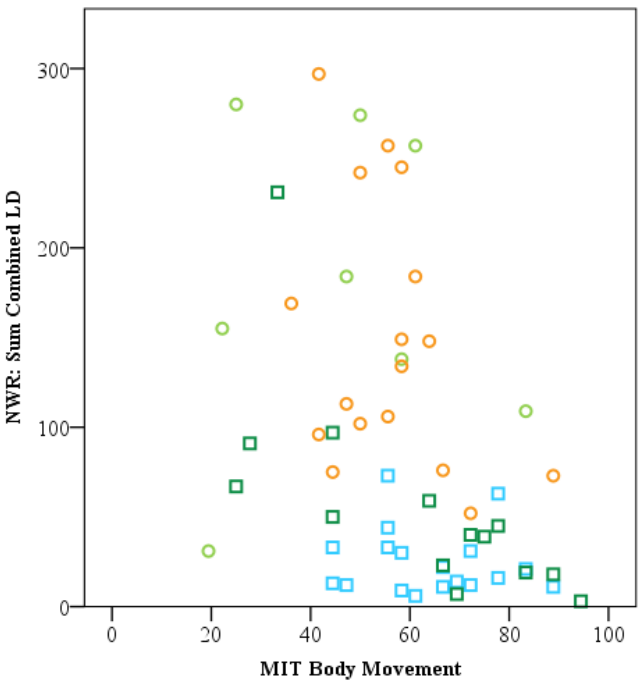


Figure 4.2: Scatter plots showing the association between the imitation of body movements and nonword repetition. *The scale of body movement imitation is a percentage score. Squares represent children without language impairment, circles those with language impairment.*

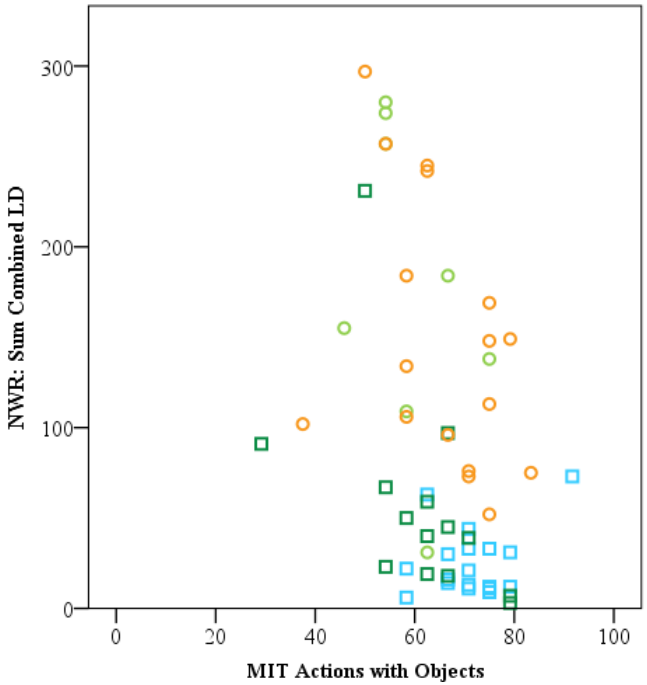


Figure 4.3: Scatter plots showing the association between the imitation of actions with objects and nonword repetition. *The scale of actions with objects imitation is a percentage score. Squares represent children without language impairment, circles those with language impairment.*

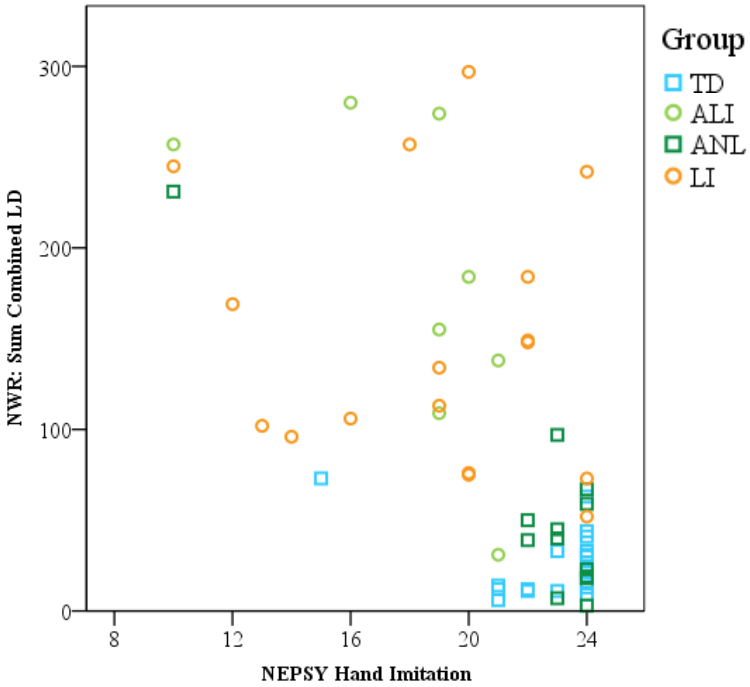


Figure 4.4: Scatter plots showing the association between the imitation of hand postures and nonword repetition. *The scale of hand posture imitation is a raw score (the maximum score is 24). Squares represent children without language impairment, circles those with language impairment.*

4.4 Discussion

Considering the groups' performances on the three tasks of imitation, there were group effects for two of them – imitation of actions with objects, and imitation of hand postures. Only children who were autistic and had language impairment showed significant deficits on the action with objects task, compared to the typically developing group. For the hand imitation task, both language impaired groups showed significant deficits compared to the typical and autistic group that did not have language difficulties. No group effects were found on imitating the style of actions with objects. Across the whole sample, motor imitation and nonword repetition correlated significantly and strongly. Partialling out nonverbal cognitive skills or fine motor skills suggested that these factors could account for the correlations between the body movement and actions with objects imitation tasks and nonword repetition performance. However, even after partialling out these skills, hand imitation performance remained significantly correlated with nonword repetition performance.

The failure to find significant group effects for meaningless body movements is surprising. Imitation of meaningless actions has been suggested to be particularly impaired in ASD, both by proponents of broken mirroring systems playing a role in ASD, and others who have disputed such theories (Hamilton, 2013; Williams et al., 2001). Nonetheless, a recent meta-analysis of imitation studies in autism did not find any significant effects of imitation task domain (including oral-facial, hand-body, and object-directed tasks) (Edwards, 2014). Indeed, previously reported differences between body movement and object-based imitation may have been at least in part driven by differences in task difficulty. Actions with objects are arguably easier than meaningless body movements, because the objects themselves provide some limitations on what possible actions can be performed, and children can bring previous knowledge about

familiar objects to the task of imitating them (for example, in the current study, children would hold the baton for the xylophone trial or the rolling pin for the rolling dough trial correctly before seeing the video – demonstrating that they already had a good idea of what the model would do) (see Williams, Whiten, & Singh, 2004).

While there were limited group differences for the imitation tasks from the MIT, hand imitation did pattern with language status, with both LI and ALI groups showing significantly poorer hand imitation than the other two groups. Furthermore, nonword repetition ability was correlated with this task, even after partialled out effects of age, nonverbal cognitive skills and fine motor ability. Previous research on hand imitation tasks with language-impaired children has found conflicting results; impairments on this same task by were also found by Brookman et al. (2013), but Wray et al. (2015) did not find hand imitation deficits in their language-impaired sample.

Imitation difficulties in the non-autistic language-impaired children were found for the hand imitation task only. The LI group's selective difficulty with this task is unexpected on two fronts. Firstly, it could have been anticipated that children with LI would show increased difficulty with the body-movement task compared to the hand imitation task, as the body movement task placed greater reliance on working memory, a domain argued to be weak in some children with language problems (Baddeley, 2003). Secondly, it may have been expected that the body-movement task from the MIT and the NEPSY hand imitation task would pattern together - both are not goal-oriented or object-based, and therefore are presumably both supported by the direct mimicry route. However, some of the hand postures in the NEPSY could be construed as having communicative or symbolic meaning. For example, one posture is holding a hand up to one's ear like a telephone. While not goal-oriented or object-based, such an action is not

meaningless, and could perhaps be parsed in terms of the communicator's intentions, via the indirect route.

These limited imitation problems in the language-impaired sample in this study are largely in conflict with previous investigations. Vukovic, Vukovic, and Stojanovik (2010) and Marton (2009) found impaired body movement imitation in children with SLI, and we also did not replicate the pattern of findings obtained by Dohmen et al. (2013), who found poor performance of their language-delayed toddlers on a body movement task, but intact performance on imitation of instrumental actions.

There are several factors that could contribute to the varied results in the studies of imitation in language impaired children conducted so far. The first is whether the researcher scoring the imitation task was blind to the group membership of the children. In the studies by Stone et al. (1990), Vukovic et al. (2010), Marton et al. (2009), Brookman et al. (2013), and Wray et al. (2015), researchers do not report blinding the scorers. Indeed, while Stone et al. (1990) and Wray et al. (2015) did not find imitation deficits, these studies were not chiefly investigating imitation deficit in language impaired samples *per se* – Stone et al. (1990) were using their LI sample as a control group for studying imitation in ASD, and Wray et al. (2015) were primarily examining children's gesture. The study conducted by Dohmen et al. (2013) is actually the only study thus far to have found impaired imitation skills in a language-impaired sample with blinded scoring. The current project represents only the second investigation to blind code language-impaired children's imitation performance. The children's responses were taped and their performance scored by trained undergraduate students who were not told what groups of children were involved in the study. The hand imitation task however was scored "live" by a researcher who knew the group membership of the children, and the hypotheses of the study. Could the findings of

selective hand imitation difficulties in children with language impairments in this study simply be due to experimenter bias?

Dohmen et al.'s (2013) study contains some additional important differences to other investigations that could also shine light on what factors influence whether or not imitation deficits are found. Firstly, there is the age of the children. Dohmen et al. (2013) investigated imitation performance in toddlers, and is the study of imitation with the youngest sample of language impaired children. Furthermore, the pattern of performance in Dohmen et al.'s (2013) language impaired sample arose largely due to *refusals* rather than reduced accuracy. Returning to the points made in Chapter 2, this suggests reduced *propensity* to imitate, and may reflect early social difficulties in this sample of children.

Considering further the issue of blinding and experimenter bias, the fact that only children with ALI showed deficits in the tasks that were coded blind could potentially be explained by coders suspecting these children had developmental disorders. Scorers were not told that the study included children with ASD, but the children with ALI arguably did appear more developmentally delayed, exhibited more behaviours (such as repetitive actions, or "stimming") that may have led coders to suspect they have a developmental disorder, which may have biased the scoring. Children with language impairment on the other hand, given that the video task did not require any verbal responses, may have appeared no different to the typically developing controls. Indeed, many of the ANL children arguably showed fewer autistic behaviours in the videos than the ALI children, and thus may have been less influenced by bias.

Another possibility is that the pattern of findings in this study is due in part to the children with ALI and LI having poorer motor abilities or poorer general cognitive skills than the other groups. A thorough assessment of the motor skills of these children was not undertaken, but the children in the ALI and LI groups demonstrated poorer performance on a task of fine motor ability. This is in keeping with reports of motor difficulties being common to both children with language difficulties, and children with autism (Green et al., 2009; Hill, 2001). However, while correlations between the MIT tasks and nonword repetition performance were reduced to statistical insignificance by partialling out nonverbal cognitive skills and motor skill, the associations between hand imitation and nonword repetition persisted.

4.4.1 Limitations

Some of the limitations to do with the sample of this study are discussed in Chapter 3, and so are not discussed again here, but rather specific issues to the study of the motor imitation skills are considered. This includes the limited number of imitation types studied; for example, oro-motor and facial imitation abilities were not assessed, and these could perhaps be expected to show stronger associations to verbal imitation than manual tasks. Body movement and actions with objects were specifically selected from the MIT as these tasks seemed to best map onto the neuropsychological models of imitation, in which meaningful and meaningless actions are suggested to be supported by different pathways in the brain. Furthermore, previous reports had suggested that imitating meaningless body actions is a selective deficit in ASD, and that imitation of goal-orientated actions with objects was comparatively spared, and similar patterning of imitation deficits had been reported in SLI (Dohmen et al., 2013; Stone et al., 1997). Thus, these tasks seemed best suited to pulling apart these different pathways, across disorders. As well as the limited number of imitation tasks, all of these tasks were based

on accuracy, and allowed no assessment of the children's spontaneous imitative behaviours (which may be more important, according to the social motivation hypothesis, outlined in Chapter 2).

In addition, given the clear importance of motor skills in the imitation abilities of both children with autism and children with language impairment (as discussed in Chapter 2), the study could have benefitted from having a more thorough motor assessment for the disorder groups, especially an assessment of praxis skills, which have been shown to be important in the imitative deficits in ASD (Dziuk et al., 2007; Mostofsky et al., 2006). Only one task – the NEPSY finger tapping task – allowed for the study of the children's fine motor skills, and the standard format of this task does include some imitation in the explanation and administration of the task (thus, this sub-test represents a motor test that is confounded with imitation). Partialling out this task suggested a fairly limited effect on the relationship between motor and verbal imitation compared to partialling out nonverbal skills, but including other motor tasks would have allowed a better examination of this issue. Indeed, ideas about the imitation deficits in ASD have moved on to theories that autistic individual's own action representations may be abnormal, leading to motor difficulties and preventing the development of representations that can be used for imitation and understanding of others (Cattaneo et al., 2007; Fabbri-Destro, Cattaneo, Boria, & Rizzolatti, 2009 - see Section 2.6). Given reports of motor and social impairments in SLI, it could be informative to consider whether these children also show evidence of abnormal action representations. Furthermore, including more motor tasks would have allowed for the examination of whether any overlapping imitation deficits were specifically shared imitation problems, or if actually the groups shared more broadly impaired praxis or motor skills.

The NEPSY hand imitation task was selected as an additional imitation task, as it represented a standardised task (unlike the MIT) that has previously been found to be impaired in some studies with language-impaired children (Brookman et al., 2013). However, while there were strong group differences on this task, many of the typically developing children and children in the ANL group were at ceiling. Potentially, a harder hand posture imitation task may have allowed differences between the ANL and TD group to emerge, and this no doubt means that the correlations with the hand imitation tasks are largely driven by deficits in the LI and ALI groups, as there is so little variation in the TD and ANL groups.

Finally, two children from the disorder groups were unable to properly do the NEPSY finger tapping task, and one of these children was also non-compliant for the NEPSY hand imitation task. The rest of the sample did do the task, but for a number of the language-impaired children these tasks appeared to elicit some frustration, and seemed to require extra encouragement to complete. I speculate that this may be related to the fact that children with language impairments often exhibit motor difficulties, and that these children may have been at least partly aware of their own motor impairments, and did not want to do a task they knew they would do badly on. Several investigations have noted that non-autistic children with language impairments suffer from reduced self-esteem (Jerome, Fujiki, Brinton, & James, 2002), but no reports appear to have considered the children's understanding of their own motor problems. Arguably the acceptance in the research community that motor impairments are the rule rather than the exception in SLI has come relatively recently. It is interesting to consider that the children themselves may have had some insight into their own motor difficulties, perhaps preceding such insight from experimental psychologists.

4.4.2 Summary

Limited group differences were found on tasks of motor imitation. Strong associations were seen between verbal and motor imitation, though nonverbal cognitive abilities and fine motor skills appeared to play an important role in this relationship, suggesting that verbal and motor imitation may only be linked at their broadest levels. Nonetheless, the relationship between hand posture imitation and nonword repetition persisted, even after partialling out these factors. This task, the only task to showed significant differences between language-impaired and normal-language groups, was a task administered and scored live, and so was potentially subject to greater experimenter bias. This study has highlighted the importance of blind scoring procedures in behavioural tests in developmental disorders research.

PART TWO: Mu suppression as a neurophysiological measure of the MNS

CHAPTER FIVE: Mu suppression in the past, present and future

5.1 Introduction

Mu is a rhythm observed in a typical human EEG, usually defined as the frequency band 8-13Hz, thought to arise from the sensorimotor areas (typically electrode sites C3, C1, Cz, C2, C4). When a person is at rest, the cells in the sensorimotor cortex fire in synchrony. When a person performs, observes or imagines themselves performing an action, the firing of these cells becomes desynchronised. This desynchronisation leads to reduced mu power, compared to when the cells were firing together (Pfurtscheller, Neuper, Andrew, & Edlinger, 1997). Changes in mu power (i.e. the strength of the mu frequency band) have been used in recent years as a means to study the human mirror neuron system (MNS). Typically, investigations examining these changes take the form of mu suppression studies, in which the power of mu is compared between a baseline condition (in which an experimenter does not expect to excite mirror neuron activity) and an experimental condition (see Figure 5.1).

Mu suppression studies have increased in number over the last decade. A recent meta-analysis surveyed 85 studies conducted since 1990 (including data from 1,707 participants). This included only studies that examined mu rhythm activity in typical participants, and used an experimental paradigm that had an action observation condition or an action execution condition (or both) (Fox et al., 2015). Of these 85 studies, 49 were conducted since 2010. Cuevas, Cannon, Yoo, and Fox (2014) also note an increase in the use of mu suppression studies with infants.

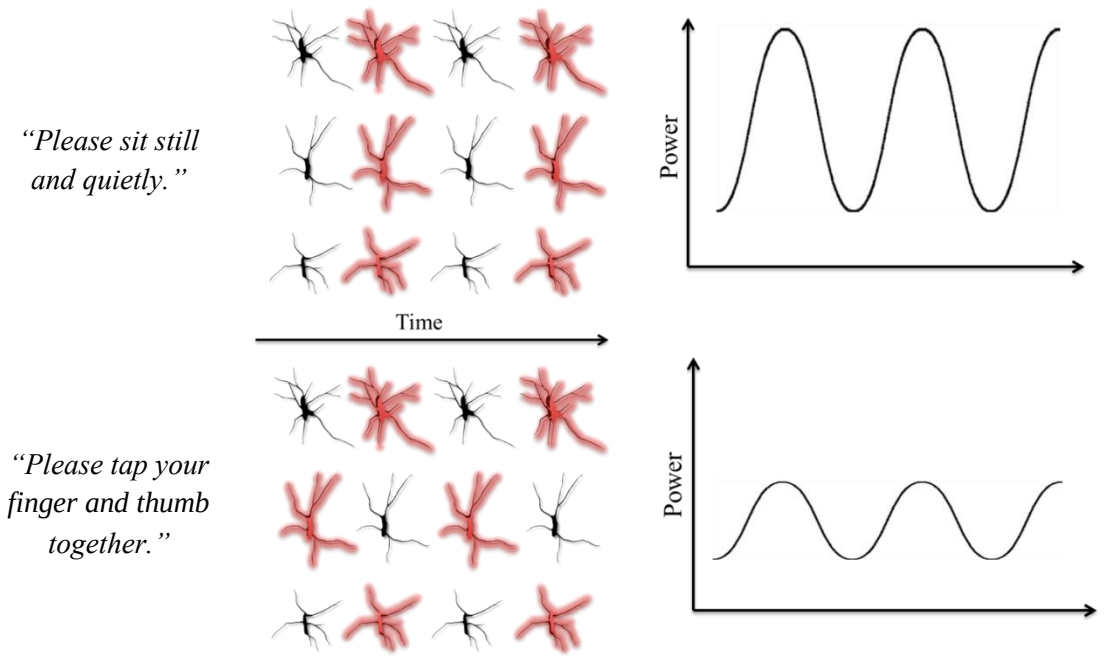


Figure 5.1: Mu suppression experiments. *Two conditions – a baseline condition (top) and an active condition (bottom) – are represented here. Bold cells represent neurons firing. In the baseline condition, the participant sits motionless. When at rest, the cells in the sensorimotor cortex fire together, leading to higher power in the mu frequency band. In the active condition, the participant is asked to move, generating motor cortex activity. This leads the sensorimotor cells to fire out of synchrony, leading to reduced mu power. Change in mu power is indexed by subtracting the baseline period from the active period. A negative value (suppression) indicates motor cortex engagement.*

While mu suppression studies are on the rise, the human mu rhythm has a long history, going back to some of the earliest EEG experiments, long before the discovery of mirror neurons. One aim of this review is to give a brief overview of the history of mu, before its reconceptualization as a proxy for MNS engagement. Furthermore, given the widespread and increasing usage of mu suppression study designs to explore the existence, reactivity and potential function of the human MNS, it is pertinent to consider current issues in mu suppression studies – does this method represent a sufficiently reliable technique to infer MNS involvement? Thus, a second aim of this review was to outline current problems in contemporary mu suppression studies. Finally, while traditionally mirror neuron studies used stimuli showing quite simple action observations, and the primary function of the MNS was argued to be for action understanding (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996;

Pellegrino, Fadiga, & Fogassi, 1992; Rizzolatti et al., 1988) today theories about the role of the human MNS have expanded to include links to language, social processes and autism spectrum conditions (Gallese & Goldman, 1998; Gallese, 2008; Oberman & Ramachandran, 2007; Rizzolatti & Arbib, 1998). A recent meta-analysis reviewed mu suppression during action observation (Fox et al., 2015), but mu suppression studies have moved well beyond this domain now, using mu to investigate an extended list of potential MNS functions. In light of any methodological issues in the field of mu suppression, this review also aimed to consider what such mu suppression studies have revealed about the MNS

5.2 Mu suppression – a brief history

5.2.1 Early studies of mu suppression – "the rhythm en arceau"

The human mu rhythm was first described by the French scientist Henri Gastaut (Gastaut, 1952, Gastaut et al., 1952), and was termed the Rolandic wicket rhythm, or the *rythme rolandique en arceau*, due to the waves' arch-like or wicket-like shape (see Figure 5.2). Early observations were that these waves seemed to arise from the rolandic regions, at a rate of around 7-11 per second (today, the mu frequency band is commonly defined as 8-13 Hz).

Back in the 1950s, the mu rhythm was thought to be a rare phenomenon, and was even considered indicative of psychopathology. Schnell and Klass (1966) identified mu in just 2.9% of their participants. Gastaut et al. (1959 a, b) found mu in 14% of a population of 500 healthy male adults, and its occurrence was suggested to be related to psychopathic personality traits. Gastaut and Bert (1954) proposed that the cause of the rhythms en arceau could be the same cause of psychosomatic traits in neurotic individuals, suggesting "*the rhythm 'en arceau' is the rhythm of subjects ill at ease in their skin*" (p. 441). Somewhat later, a link between mu and epilepsy was also suggested, and there were even positive

findings concerning the effects of mu biofeedback training and epileptic symptoms (Sterman, Macdoland, & Stone, 1974). (This is particularly interesting in light of recent work linking autism spectrum disorders to mu abnormalities, and attempts to use neurofeedback with this population – See Section 3.3 Mu suppression studies of autism spectrum disorders.)

As noted in Niedermeyer and Silva (2005), the arrival of standardised caps made it easier to routinely identify mu rhythms from EEG sites C3 and C4 (central sites situated over the sensorimotor cortex), and new techniques for analysing EEG also established that mu occurred more commonly than previously thought. Kuhlman (1978) noted that mu activity, with its characteristic wicket shape, was rarely visually identified in the EEGs of their participants, but power spectral analysis revealed it in around half of their sample. By calculating coherence, Schoppenhorst and Brauer (1980) were able to identify mu in 60% of their 54 participants, a much higher proportion than discovered previously. They also suggested that effects of varying vigilance may contribute to difficulties in identifying mu in participants.

As it became apparent that mu was not an unusual or particularly pathological phenomenon, new theories emerged about what mu rhythms could be related to. One prominent theory is that mu represents the resting activity in the sensorimotor cortex, and that suppression of this rhythm reflects these regions of the brain becoming active (Kuhlman, 1978; Schoppenhorst & Brauer, 1980). Indeed, in Gastaut's early studies it was recognised that participants' own movements blocked the mu rhythm, and further experiments found that mu could be blocked not only by spontaneous movements but also passive movements, reflex movements, and movements to command (Chatrian et al., 1959).

Event related desynchronization or synchronisation (ERD/ERS) describes the reduction or increase of a given power band. Largely, ERD and ERS are thought to reflect cortical activation and idling, respectively. The basic principles of using

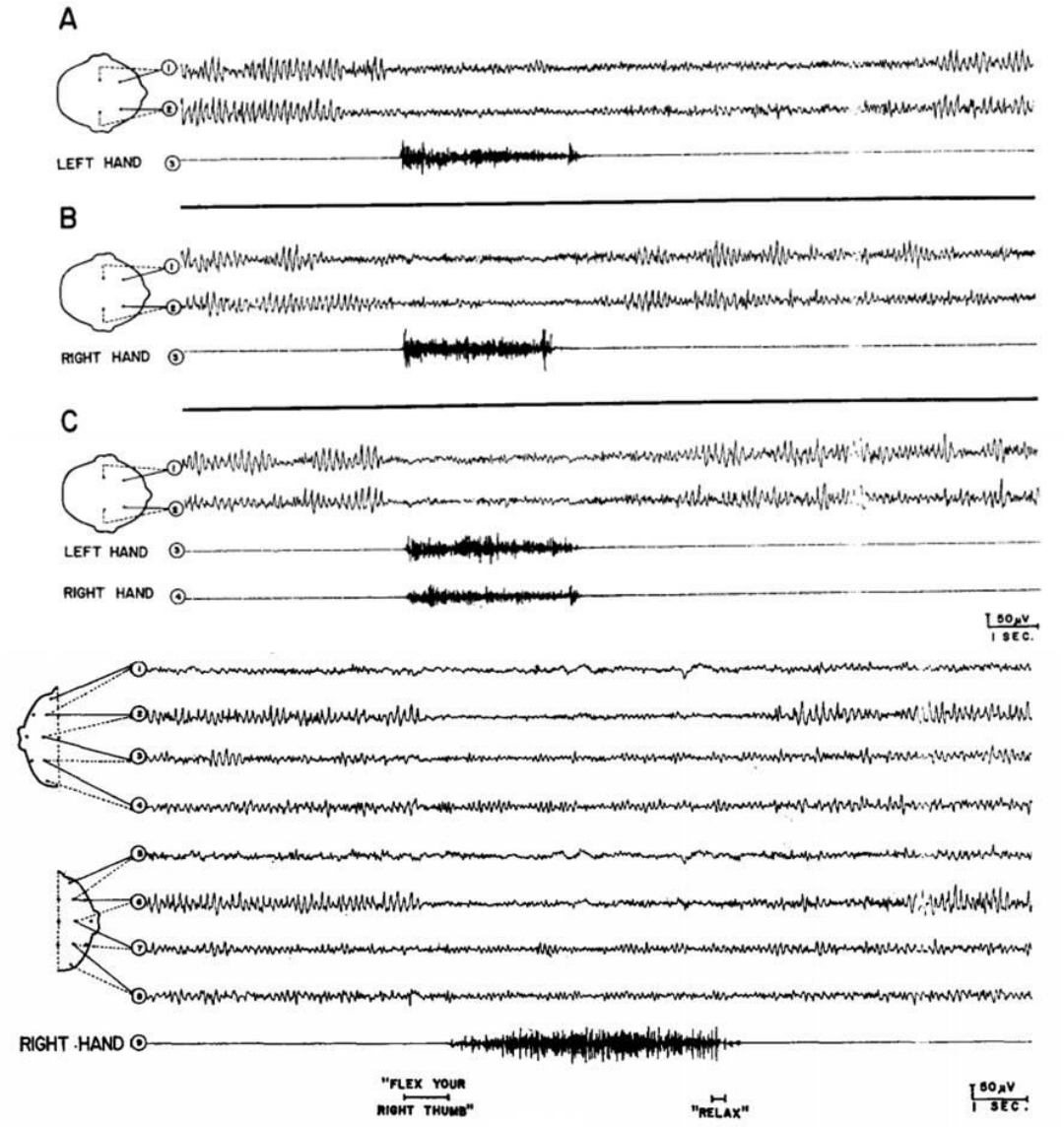


Figure 5.2: Recordings published by Chatrian et al. (1959). *The upper section shows the effects of spontaneous thumb movements on the wicket rhythm. Section A shows the effects of the movements of the left thumb, section B the effects of the movements of the right thumb, and section C the effects of both thumbs. The lower section shows the effects on the rhythm from flexing the thumb to command. Corresponding surface EMG recordings on the thumbs are shown for each recording.*

ERD/ERS in relation to alpha/mu (8-13Hz) and beta 13-35Hz) band activity are documented in Pfurtscheller and Lopes da Silva (1999). Using ERD/S, researchers have demonstrated the reactivity of mu to participants' own movements, and suggested that there may be different types of mu rhythm, perhaps corresponding to different areas of the motor cortex (Pfurtscheller & Neuper, 1994). However, it is mu's reaction to other forms of stimuli that have generated so much research over the last decade.

5.2.2 Mu suppression and the discovery of mirror neurons

As well as reacting to participants' own movements, mu is suppressed by observing the movements of others. Reports of mu suppression during the observation of movements arose as scientists began to examine EEG responses to film projections. In a delightful old paper by Gastaut and Bert (1954), the authors describe their observations of their participants' EEG while watching a film reel of a boxing match: "*It* [mu] *decreases or disappears completely when the subject changes his position on his seat or when he readjusts his tonus. It also disappears when the subject identifies himself with an active person represented on the screen. This phenomenon is particularly interesting to study during a sequence of film showing a boxing match. A few seconds and, at times, less than a second after the appearance of the boxers all type of rolandic activity disappears in spite of the fact that the subject seems completely relaxed and that there is no noticeable change of posture. The relation between the blocking of the "arceau" rhythm and the image of boxers in action is unquestionable. In the middle of this particular film strip, the camera is suddenly turned from the ring to the spectators in the hall for a few seconds. In many subjects the rhythm "en arceau" reappears during this short period and vanishes again as the boxers reappear on the screen.*" (p. 439)

In the 1980s, a team of Italian neuroscientists identified cells in the macaque brain that fired both when the animal performed an action and when it viewed an action being performed by another (Pellegrino et al., 1992; Rizzolatti et al., 1988). These cells were subsequently named mirror neurons, and the observation that the sensorimotor cortex became activated when viewing movement evolved into the mirror neuron theory of action understanding (Gallese et al., 1996). Following the discovery of mirror neurons in the macaque, the phenomenon of mu suppression took on a new interpretation. The responsivity of the mu band arguably shows similar response properties to mirror neurons. Parallels were drawn between mu and mirror neurons, and a reduction in mu activity was suggested to be a signature of mirror neuron activity (Muthukumaraswamy & Johnson, 2004; Muthukumaraswamy, Johnson, & McNair, 2004).

Original mirror neuron experiments in monkeys had suggested that mirror neuron activity was related to goal-directed actions specifically, and classic mirror neuron study stimuli involved a hand interacting with an object. In the animal literature, equivalent movements not directed to an object do not cause mirror neuron activity (Pellegrino et al., 1992). Mu suppression studies with human participants have found that stronger mu suppression occurs when viewing another's hand in a precision grip (i.e. a grip that could be used on an object) rather than in a neutral, non-grip positon, and that object-interaction produced greater mu suppression than conditions without object interaction (Muthukumaraswamy & Johnson, 2004; Muthukumaraswamy et al., 2004). This object effect has been proposed to support the notion that mu suppression is related to mirror neuron activity in humans. Arguably however, a strict interpretation of the animal recording work would suggest that mu suppression should not occur at all when viewing actions that do not relate to an object. Instead, some authors have speculated that human MNS responses to non-object directed actions is a key difference between monkey and human mirror neurons, and

that this difference may represent a departure from our shared ancestors with other primates, which may have played an important role in the evolution of language (Rizzolatti & Craighero, 2004). However, studies also note that mu suppression may be measuring the activity of areas downstream of mirror neurons, rather than mirror neuron areas *per se* (Muthukumaraswamy et al., 2004).

5.3 Is mu suppression a good measure of the MNS?

5.3.1 The scientific quality of mu suppression studies

As mu suppression is already widely used in cognitive neuroscience to imply roles for mirroring systems in higher social processes and clinical disorders, the question of whether mu suppression is good a measure of mirror neuron activity is an important one. A recent meta-analysis of mu suppression studies concluded that while mu suppression offered a valid means to investigate MNS engagement, there were several limitations common in the literature (Fox et al., 2015). These problems included issues relatively specific to the field of mu suppression, including the fact that few studies report changes in alpha-band power at sites other than the central electrodes (making it impossible to be sure that effects were not being driven by changes in power elsewhere), and the related problems of attentional differences between conditions. While echoing these recommendations, several other problems exist in the mu suppression literature, some of which apply also to the wider field of neuroimaging and psychology. These broader points are considered first, before discussing some design issues specific to mu suppression studies.

Firstly, mu suppression studies generally suffer from low statistical power and small sample sizes. In studies that use clinical groups such as autism this is perhaps more understandable, as these groups can be hard to recruit and may also be less willing to undergo neuroimaging than non-clinical participant groups. Where studies have used typical samples

however, sample sizes well below what would be needed for properly powered statistical analyses is less forgivable. It is intuitive to most that small sample sizes reduce the likelihood of detecting a true effect. This may appear to mean that if an effect *is* found, even when the sample size is small, then this effect must be true. However, this assumption is inaccurate – a lack of power also means that significant effects are *less likely to reflect a true effect*. In the field of neuroscience, low power has been argued to increase the risk of false-positives, underlie overestimation of effect sizes, and lead to problems reproducing effects in subsequent studies (Button et al., 2013). How many participants are required for a given study depends on a number of factors including analytical design, number of conditions, the expected effect size, correlations between measures, and more. Thus, there is no set number of participants that a mu suppression study should include. However, as a rough guide, a repeated measures design with 2 factors each containing 2 levels analysing its data with a two-way ANOVA, would need 40 participants to be sufficiently powered to detect a medium sized effect with 90% power.[6] To detect an interaction, 47 participants would be needed.

Secondly, mu suppression is a phenomenon with substantial analytic flexibility, and this is another known risk factor for poor reproducibility (Ioannidis, 2005). For instance, mu suppression studies vary on what frequency band is considered "mu". Frequency bands are not distinctive categories but are flexible ranges that have arisen from the EEG literature, which means that mu suppression papers can employ slightly different frequency bands from each other. Some have argued that the correct frequency band may have to be calculated from individual to individual, akin to functionally defined sites in MRI (Pfurtscheller & Lopes da

[6] This was calculated using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). 90% power is the minimum accepted by most journals offering pre-registration. A conservative estimate of nonsphericity correction was employed – and indeed it is typical for this assumption not to be met. I balanced this conservativism by entering a relatively high correlation between the measures : 0.7, higher than that reported by mu suppression study of Tangwiriyasakul, Verhagen, van Putten, and Rutten (2013) . Lowering this correlation would increase the number of participants needed, and relaxing the nonsphericity correction would reduce the number of participants needed.

Silva, 1999). Indeed, this could be especially important if mu is to be used for neurofeedback, and methods for calculating individual frequency bands have been proposed (Pfurtscheller & Lopes da Silva, 1999). While there may indeed be a theoretical rationale for splitting the mu rhythm, or selecting a higher or lower or narrower or wider band to examine, these decisions are free to be made after EEG data has been collected and initially analysed. This leaves scope for researchers to select a frequency band that provides the best results to fit their hypothesis. A positive step forward taken by some studies is the use of more sophisticated analytical techniques, using independent components analysis (ICA) to extract components that represent mu. This would seem to represent a more objective approach to examining the response properties of the mu rhythm.

Related to the issue of analytic flexibility is that of studies calculating a large number of correlations, or running ANOVAs, without proper correction for multiple testing. These studies are arguably of an exploratory design, and need to be considered as such. While ANOVAs effectively correct for the number of levels within a given factor, they do not automatically correct for the number of factors, or the number of potential interactions between factors. For example, a three-way ANOVA is testing for three main effects, three two-way interactions, and one three-way interaction. In such an ANOVA, the probability of finding no significant effects at all (if alpha is set to 0.05) is 0.95 ^ 7 = 0.70. Thus, the likelihood of finding a significant main effect or interaction is 30%. For a four-way ANOVA this likelihood increases to 54%. ANOVAs are not only problematic in mu suppression literature, but in the wider EEG field and behavioural sciences as well (Cramer et al., 2014; see also the blogpost by Bishop (2013) for a discussion of these issues in relation to EEG).

5.3.2 The problem of alpha

More specific issues in mu suppression work are difficulties ensuring that changes in the 8-13Hz frequency band are really arising from sensorimotor areas, and not other regions in the brain. Activity in this frequency band, commonly called alpha-band activity, occurs all over the head, and changes in it have been implicated in a number of processes. What is alleged to distinguish "mu" from occipital "alpha" is topography and responsivity – while alpha arises from the occipital cortex and reacts to changes in visual stimulation and attention, mu arises from the sensorimotor areas, and responds to participants' own movements.

Perhaps the studies best placed to shine light on this are those that have considered how well mu suppression correlates with other measures purporting to measure the MNS. Such investigations include those that have concurrently taken EEG and fMRI recordings, with the view to investigating whether these two measures were in good agreement, and if mu suppression could serve as a cheaper, more accessible way to study the mirror neuron system (Arnstein, Cui, Keysers, Maurits, & Gazzola, 2011; Braadbaart, Williams, & Waiter, 2013; Mizuhara, 2012; Perry & Bentin, 2009). Broadly, the results have been positive – the BOLD responses in brain areas considered to be part of the human MNS (including the inferior parietal lobe, dorsal premotor and primary somatosensory cortex) correlated with mu suppression. However, despite these correlations, authors have warned that their results also suggest that mu suppression may also be reflecting activity from other networks, including regions involved in visuomotor processes that are not part of the mirror neuron system (Braadbaart et al., 2013; Perry & Bentin, 2009).

Another putative measure of MNS activity is TMS-induced motor evoked potentials (MEPs). Lepage, Saint-Amour, and Théoret (2008) combined EEG and TMS to investigate the relationship between these two measures. As in previous studies, a significant increase in the MEP amplitude and suppression of mu during action observation, imagination and

execution were shown – but there was no correlation between these measures. Other evidence for mu suppression validity comes from MEG studies. MEG is considered superior to EEG in its ability to spatially localise sources. In MEG mu-suppression studies, the areas significantly modulated by observation and execution are the sensorimotor cortices, but effects are also found at occipital areas (Muthukumaraswamy & Singh, 2008; Muthukumaraswamy et al., 2006).

One of the key issues with mu suppression is its overlap with the alpha frequency band. Inadequate separation of mu from alpha could potentially mean that reported changes in mu to certain stimuli are actually changes in alpha, and differences in attention between conditions. For example, Aleksandrov and Tugin (2012) measured mu suppression during a range of control conditions, including periods of mental counting, or watching the movement of a non-biological objects, conditions in which we would not predict mirror neurons to become active. Yet mu suppression during these conditions was not significantly less than in conditions in which participants observed human actions. Furthermore, their most attentionally demanding tasks appeared to produce the strongest mu suppression, suggesting that mental effort could confound mu studies. Similarly, Perry and Bentin (2010) highlight that the significant differences between conditions in their mu suppression study may have been due to differences in attentional demands, rather than differences in the activity of mirror neurons, after finding a comparable pattern of changes in 8-13 Hz power at both occipital and central electrodes. One of the classic markers for alpha is the blocking of eye-opening; in one study, eye opening had the biggest effect on mu suppression, greater than any of the experimental conditions (Cuellar, Bowers, Harkrider, Wilson, & Saltuklaroglu, 2012).

It is well documented that alpha activity is tied to alertness and cognitive effort, and some of the earliest reports noted that this was also the case for mu. Chatrian et al. (1959; 1976) reported that mu waves depended on changes of vigilance, and that mental arithmetic

(Chatrian et al., 1959) or problem solving (Creutzfeldt, Grunewald, Simonova, & Schmitz, 1969) suppressed the mu rhythm. Schoppenhorst and Brauer (1980) state: "*While visually evaluating routine EEGs we noted that mu waves were remarkably unstable. On closer observation we found that this ability could be attributed to slight changes in the degree of vigilance, often hardly discernible from changes in the alpha rhythm"* (p. 25). And later: "*They are very unstable due to their dependence on the degree of vigilance and can be suppressed both by an increase or a decrease in vigilance*" (p. 31).

Perry and Bentin (2010) caution that mu suppression studies should report effects (or lack of effects) at the occipital sites, as well as the central sites, in order to understand the phenomenon being studied. Yet mu suppression studies that have considered the activity at the occipital electrodes have had mixed findings. In their concurrent TMS-EEG study, Lepage et al. (2008) found that 8-13 Hz power at the occipital sites was indeed supressed during conditions when participants observed actions, or imagined themselves performing actions. In contrast, mu suppression studies investigating abnormal mu responses in autistic participants have reported that other than C3, Cz, and C4, no other electrodes showed a consistent pattern of suppression, suggesting their pattern of results could not have arisen due to confounding from changes in occipital alpha (Bernier, Dawson, Webb, & Murias, 2007; Bernier, Aaronson, & McPartland, 2013; Oberman et al., 2005, 2008).

Related to issues around alpha and attention, the choice of baseline technique in mu suppression experiments is very important, and could have an important role in whether mu suppression is observed or not. As pointed out by Pfurtscheller and Lopes da Silva (1999) "*the ERD is measured in the percentage of power relative to the reference interval and therefore it depends on the amount of rhythmic activity in this interval*" (p. 1847). Baselines of inactivity (where participants receive no stimulation and are asked to sit quietly) may inflate alpha power during the baseline period. Then, when the experimental conditions are

subtracted from this period, there is increased likelihood of observing a significant decrease in the 8-13 Hz band, which is interpreted as mu suppression and thus motor cortex or MNS engagement. Potentially, different choices of baseline may contribute to inconsistent findings in mu suppression literature. While the meta-analysis by Fox et al. (2015) suggested that baseline did not affect whether or not mu suppression was observed, other studies have highlighted potential problems with certain baseline techniques (Hobson & Bishop, in press; see Chapter 6).

5.3.3 Mu suppression beyond action understanding

Having laid out the potential shortcomings of mu suppression, we now explore findings from studies that have used this technique to explore the role of the MNS in a number of different functions. While originally the MNS was proposed to function primarily as a substrate for action understanding, the list of potential roles the MNS could play soon expanded. Language, theory of mind, and empathy have all been suggested to have a root in the mirror neuron system. If mu suppression is a sufficient measure of the human mirror neuron system, and if the mirror neuron system is involved in these processes, then one would expect to see mu suppression during tasks that involve language, theory of mind, and empathy. Here, the evidence from mu suppression studies for mirror neuron involvement in each of these processes is explored.

5.4 Mu suppression during language tasks

The suggestion that perhaps language could have a close relationship to the human mirror neuron system has arisen via a number of overlapping and interconnected theories. The first is the motor theory of speech perception (Liberman & Mattingly, 1985). This theory (proposed originally quite independently of and without reference to mirror neurons) argued that producing and perceiving speech relies on the same structures and processes, and that

speech is thus perceived by listeners as a series of gestures or motor commands. While there is superficial parity between the motor theory of speech perception and mirror neurons, language researchers have pointed out that the motor theory of speech perception has long been considered incorrect, and that the revival of this theory following the discovery of mirror neurons has done little to address old problems with the theory (Lotto, Hickok, & Holt, 2009).

The second theory connecting the MNS and language is an evolutionary one, in which a primitive observation-execution matching system is argued to have supported early communication, followed by the development of a more sophisticated MNS which enabled speech. It has been proposed that language may have begun in our ancestors as a mimetic gesture-based communicative system (Bates & Dick, 2002; Corballis, 2003; Donald, 2005; Rizzolatti & Arbib, 1998). Broca's area (a key language area and possible homolgue for F5, the area of the macaque brain in which mirror neurons were originally discovered) has a central role in the evolutionary MNS-language theory; Rizzolatti and Arbib (1998) suggest that Broca's area was originally a region that served action recognition rather than language, and that this was a "neural prerequisite" for the evolution of communication and eventually speech. (In fact, previous authors had speculated that mu suppression was most likely being generated by Broca's area, but the findings of the concurrent fMRI-EEG study by Arnstein et al.'s (2011) do not support this notion.) Such evolutionary theories about mirror neurons would suggest that these cells' mirroring properties are innate – this has been disputed by authors who have argued that mirror neurons could quite well arise due to associative learning (Heyes, 2010, 2013).

A third theory linking the MNS and language is the notion that language is "embodied", and that the existence of the MNS provides evidence for this embodiment. Embodied cognition accounts have been suggested for a number of different domains and

processes, including action understanding (Grafton, 2009) and executive function (Koziol, Budding, & Chidekel, 2012). Broadly, embodied cognition accounts argue that cognition is grounded in perception and action, and true understanding of cognition therefore requires an appreciation for the environment and the resultant perceptual experiences of the organism, and the actions they perform as they move through their world (however, embodied cognition theories are very diverse in their claims – see Wilson, 2002). Language has been considered from an embodied cognition framework (Fischer & Zwaan, 2008; Gallese, 2008), and as described by Gallese (2008), the embodiment process can be considered at multiple levels – the "vehicle level" (this is essentially the motor theory of speech perception), and the "content level" (the semantic content of words). Thus, this theory encapsulates and expands beyond the motor theory of speech perception – motor resonance is not just for articulatory gestures, but for the content of the sentence itself. For example, hearing or reading the word "kick" may lead to the simulation of "kick" in the hearer's motor cortex.

Having laid out the theoretical links between language and the mirror neuron system, what do mu suppression studies suggest about the role of mirror neurons in language and speech processing? Studies examining mu suppression and language have examined both auditory and visual stimuli, and stimuli at the level of phonemes and meaningful sentences. Evidence for mu suppression during speech sounds is considered first, followed by suppression during sentences.

Before examining the findings from studies of auditory speech stimuli, many mu suppression studies have been largely concerned with object-hand interactions. Is mu suppression also observed when participants view mouth and face movements? Findings from MEG studies had previously suggested that mu suppression was limited to oro-facial movements that were *not* linguistic. In agreement with previous findings of an object-effect with hand stimuli, Muthukumaraswamy, Johnson, Gaetz, & Cheyne (2006) found that mouth

movements directed towards an object (a straw) produced the strongest mu suppression, compared to object-less mouth movements, and linguistic mouth movements. Linguistic mouth movements however produced no significant changes in mu power from baseline, leading the authors to conclude that linguistic stimuli are processed differently to other human movements. However, EEG studies have suggested that mu suppression *does* occur when participants view visual speech stimuli. Crawcour, Bowers, Harkrider, and Saltuklaroglum (2009) investigated mu responses to both visual and auditory speech stimuli. Their 14 participants were presented with 9 conditions with different audio-visual pairings, which included visual or auditory speech, visual or auditory noise, and non-biological stimuli (a kaleidoscope pattern or a tone). The only conditions to induce significant mu suppression were those with visual speech stimuli. Conditions in which speech sounds were presented without visual lip movements failed to produce significant mu suppression, suggesting that mu suppression during speech is actually dependent upon visual stimuli being present. (However, it should be noted that this study only examined data from the central electrodes. The extent to which attentional differences between condition could have driven the effects is thus unclear.)

Despite these findings, other studies have found mu suppression to auditory-only stimuli, but task demands may have an important role in determining whether mu suppression is observed. Cuellar et al. (2012) investigated mu suppression during different speech processing tasks, with varying task demands. Their stimuli were auditory only. In the first of their two experiments (using 10 participants), only their most demanding task induced mu suppression, and only in electrode C3. They suggested that this indicated left-lateralised sensorimotor activity when discriminating speech in noise (and only when considering 10-12Hz sub-band). In experiment two (with 13 participants), significant suppression was found only in a task requiring participants to segment the speech stimuli into phonemes, and again

only in electrode C3. Collapsing speech versus tone conditions did produce a significant effect that suggested mu suppression is stronger to linguistic stimuli than general auditory stimuli such as tones. This study only examined changes in power at the central electrodes, however the effects of task demand would seem in keeping with attention and alpha effects. Similar to Cuellar et al. (2012), in their study with 16 participants Bowers, Saltuklaroglu, Harkrider, and Cuellar (2013) found that passive listening to speech sounds did not induce significant mu suppression compared to passive listening to noise, but that active listening (in which participants had to make a judgement about the sounds they heard) did lead to mu suppression. Together these studies suggest that auditory stimuli can lead to mu suppression, but passive listening to speech is ineffective for suppressing mu. Only active listening (requiring discrimination or segmentation) seems to lead to suppression, when the stimuli are auditory only. This would suggest a somewhat limited link between mu suppression and everyday speech processing.

One of the key findings in the mu suppression literature on action understanding is that mu suppression is observed both when a participant *performs* a movement and when they *observe* the movement of another. Indeed, it is this pattern of responses that led to mu being suggested as a proxy for mirror neuron engagement in the first place. It is therefore important for researchers looking to use mu suppression to suggest a role of mirror neurons in the motor theory of speech perception to compare mu activity in the perception and production of speech. This is non-trivial, as asking participants to produce speech during an EEG is likely to create movement artefact. However, two studies do suggest that mu suppression occurs during the production and perception of speech.

Tamura et al. (2012) conducted two experiments with 16 and 11 participants, the first of which examined mu rhythm responses during finger tapping, tongue movements, and articulation of speech sounds (with and without vocalisation). Mu suppression was seen in all

conditions (though it was weakest for articulation with vocalisation, possibly due to the high rate of trials rejected for movement artefact), showing that mu responses were not limited to participants' hand actions but could be demonstrated in speech production as well. Given the difficulties with actual production, the second experiment investigated mu suppression during imagined speech production. Participants initially recorded themselves reading out sentences. Then, while silently reading other sentences during an EEG, the participants heard either their pre-recorded selves normally, with delayed feedback, or with noise added. Finally, in one condition, the participants read the sentences in silence (without the recordings of themselves reading). Mu suppression was found in all these conditions, and Tamura et al. (2012) argued that their results demonstrated mu suppression during imagined speech production. However, a key limitation of this study is that production and perception conditions were in two separate studies – a crucial test of mirror neuron functioning would be demonstrating mu suppression both during perception and production of the same stimuli within the same study.

An interesting approach was taken by Jenson and colleagues (Jenson, Thornton, Saltuklaroglu, & Harkrider, 2014; Jenson, Bowers, et al., 2014) who investigated mu components derived from ICA (independent components analysis) using 20 participants. ICA is a technique used to decompose an EEG signal into independent components (see Onton, Westerfield, Townsend, & Makeig, 2006 for a review of the use of ICA in EEG), and offers quite a different analytical approach to mu suppression than those usually employed. Using ICA, researchers can investigate the presence of a mu component, and whether its response properties match what would be expected of a mirror neuron system involved in language processes. In the study by Jenson et al. (2014), ICA identified a mu component common to conditions requiring imagined speech production, actual production, and the perception of speech sounds (during a discrimination task).

While these studies focused largely on mu's responses to speech, few studies have investigated the role of mirror neuron systems and mu in semantic aspects of language, the "content level" described by Gallese (2008). Van Elk, van Schie, Zwaan, and Bekkering (2010) investigated EEG responses in 24 participants to visually presented sentences, in which either humans or animals performed actions (e.g. "the duck swims in the pond" versus "the woman swims in the pond"). They considered several EEG measures, including mu suppression, beta suppression and the N400. If motor cortex activation is related to motor imagery, human actions (being easier to imagine than animal actions) should show stronger effects. Alternatively, if motor cortex activation is related to the retrieval of lexical-semantic information, they argued that stronger effects should be seen in the animal context. This is because the animal noun limits the range of verbs that could follow, as verbs following animal nouns generally have higher cloze-probability (for example, in the sentence "the duck swims in the pond" the probability of the word "swims" following "duck" is much higher than it following the word "woman" in "the woman swims in the pond"). Considering mu responses alone, stronger mu suppression was observed for the animal sentences. However, this was *not* modulated with the cloze-probability of the independent sentences (although both beta and N400 were). As two out of three of the EEG phenomena showed patterns in-keeping with the lexical-semantic retrieval hypothesis, the authors' conclusion that motor cortex engagement in these tasks represents the retrieval of lexical-semantic information is not unreasonable. However, the mu responses in this study are not easily interpretable, fitting with neither hypothesis laid out by the study's authors.

A study using 30 participants by Moreno, de Vega, and Leon (2013) used mu suppression to examine motor cortex engagement when hearing either concrete-action sentences (e.g. "Now I cut the bread") or abstract sentences ("Now I doubt of the plan"). In a separate condition, participants were also shown actions (not exactly the same as those

described in the sentences) in video clips. Listening to action sentences and observing human actions was found to result in significantly greater mu suppression than listening to abstract sentences. Approaching these results from an MNS framework, this study would suggest that the role of the motor cortex (and mirror neurons) must be limited to verbs or words that have a motor association or a performable action. Thus, the mirror neuron system must only underpin certain sub-sections of language comprehension, and another system must support our ability to comprehend sentences such as "Now I doubt of the plan". Alternatively, these findings are in keeping with an associative account, similar to the associative account of mirror neuron development (see Cook, Bird, Catmur, Press, & Heyes, 2014). Quite possibly action words become associated with performing an action or viewing that action being performed, due to these words often being said or heard when performing or viewing actions. Hearing these verbs therefore activates the motor cortex via these learned associations, but the activation could be epiphenomenal, rather than playing an active role in verb comprehension. (One final issue noted with the analysis of this paper is there seemed to have been no corrections made for the number of factors entered into their ANOVAs.)

5.4.1 Summary of mu suppression in language tasks

Mu suppression studies of language have mainly concentrated on speech perception and production, as opposed to semantic understanding. Such studies have returned rather mixed findings, but many do report suppression during at least some tasks. Some researchers also have advocated investigating mu suppression in broader communication disorders, though no studies of mu suppression in language-impaired populations have yet been conducted (Le Bel, Pineda, & Sharma, 2009).

Considering the evidence for mu suppression in speech perception (as opposed to language comprehension), suppression seems more likely to occur during tasks that require

additional processing, beyond passive listening. Two theories could account for this. Firstly, the recruitment of motor areas only occurs when speech processing is sufficiently demanding. Motor areas are effectively "drafted in" as an additional resource for the task (which would suggest that motor-cortex based speech perception is not the only means to process speech sounds, and acts as an additional support, rather than a core process in speech perception) (for example, see Szenkovits, Peelle, Norris, & Davis, 2012). The second theory is that when speech perception (or perhaps any) tasks become sufficiently difficult suppression in the alpha-band is seen due to attentional effects. Alpha-band activity occurs all over the head, not just in the sensorimotor cortex, and changes in task-engagement and mental activity can lead to suppression of the 8-13 Hz band, potentially quite independent of motor cortex activity.

Finally, there is also a lack of studies that examine mu suppression in both perception and production within the same experiment, a test that seems crucial given the crux of the motor theory of speech perception is the overlap between substrates supporting perception and production. The approach taken by Jenson et al. (2014) represents an interesting way forward, and suggests mu is reactive to both speech perception and production. However, a study that replicates this finding and demonstrates this reactivity is specific to speech stimuli (i.e. not tones or other auditory stimuli) is yet to be conducted.

In sum, it is debatable to what extent speech and language tasks lead to mu suppression independent of other potential confounds, or to what extent mu responses to speech stimuli align with what would be expected if the MNS plays a key role in speech and language processes.

5.5 Mu suppression in social processes

Following on from the original theories purporting that the MNS underpins action understanding, researchers have used mu suppression techniques to investigate the activation

of the MNS in a number of related higher social processes, including empathy and theory of mind. Furthermore, there have been several investigations using mu to assess the role of the MNS in perceiving biological motion. Findings from mu suppression studies examining these skills are reviewed in turn here. Theories that link the MNS to empathy are highly similar to the theories that link the MNS to the separate but related construct of theory of mind. Therefore, these theories will be briefly discussed together, before examining the evidence for mu suppression during these different processes.

5.5.1 Empathy

One theory proposed to explain how we understand and empathise with the thoughts, feelings and beliefs of others is simulation theory. This theory posits that the same mental resources used in our own thinking and emotional responses are also used to understand the thoughts, feelings of beliefs of others (Davies & Stone, 1995). (An alternative to simulation theory is "theory" theory – the notion that understanding others draws on empirical knowledge.) Related to simulation theory is the perception-action model of empathy (Preston & Waal, 2002), which argues that empathy is based on neural simulation – seeing others performing actions, or expressing emotions, engages the same neural networks for the execution of action, or the experience of the same emotions, in ourselves. This engagement leads to associated autonomic and somatic responses. We feel what the other person is feeling. The perception-action model could reasonably be considered to be a biological-level account of the cognitive model of simulation theory, and the human mirror neuron system could be posited to be this site of perception-action linkage. If the MNS has a role in empathy, then tasks considered to require it or cause it should result in greater mu suppression than tasks that do not.

Evidence for MNS engagement during empathy-related tasks was examined by Moore, Gorodnitsky,and Pineda (2012). They considered their 22 participants' mu responses while viewing (happy or disgusted) faces, or buildings. The viewing of faces was administered under two conditions – in one, participants were asked to empathise with the person they viewed, and in another they were asked to rate how attracted they were to the face. In the buildings condition, participants rated how much they liked the building. Greater suppression was obtained when participants viewed faces as opposed to buildings, although there was no effect of the empathy/non-empathy condition, and buildings produced significant suppression too.

Two theories are posited to explain these results – one is that an empathic response (and a mirror neuron response) to faces is automatic, and thus there is no effect of asking participants to empathise or not, as the activation of the MNS occurs anyway. An alternative hypothesis that the authors concede is that as no facial EMG was recorded participants might automatically (and unconsciously) mimic the facial expressions of the faces they view. Participants' own facial movements would lead to motor cortex engagement, and thus mu suppression. The finding of significant mu suppression to buildings is argued to need more research.

Perry, Bentin, Bartal, Lamm, and Decety (2010) examined mu responses to viewing painful or non-painful stimuli. Twenty-eight participants viewed pictures of "painful" (a hand being pricked with a needle) or "non-painful" (a hand being touched by a cotton bud) stimuli. They were also told that some of these pictures were of a patient with a rare disease in which the needle would not cause pain, but the cotton bud would. Significant suppression of alpha-band activity was seen to painful stimuli - including when participants viewed cotton buds touching the purported patient's hands (demonstrating empathy). Strongest suppression for stimuli arose from the occipital rather than the central regions, but the authors argue that the

actual pain effect (stronger suppression to painful versus non-painful stimuli) arises from the central and frontal regions, rather than the occipital (however, they failed however to achieve a significant site by pain interaction).

The studies detailed so far consider average mu responses to different conditions, requiring or not requiring empathy. A different line of inquiry has examined individual differences in empathy and MNS activity, using mu suppression as a measure of MNS engagement. Indeed, some studies have suggested that female participants exhibit stronger mu suppression than males (Cheng, Tzeng, & Decety, 2006; Cheng, Lee, et al., 2008), a difference purported to be related to gender differences in empathy, and also to complement the extreme male brain theory of autism (Baron-Cohen, 2002). Furthermore, several studies have also investigated correlations between individual mu suppression, personality traits and self-reported social and empathy skills. Cheng et al. (2008) noted a positive correlation between mu suppression and the personal distress subscale of the interpersonal reactivity index (IRI), and suggest that "*the EEG mu rhythm can be a potential biomarker of empathic mimicry*"(p. 4). They also noted a negative correlation between mu suppression and the systemising quotient (SQ), a dimension related to the extreme male brain theory of autism.

However, the significance of these correlations is doubtful. Firstly, the IRI has four subscales, and the authors also report using the empathizing quotient, systemizing quotient and emotional contagion scale. Thus, there were seven measures which were investigated for a correlation with mu suppression, but the authors do not report any corrections made for multiple tests. Furthermore, the effect sizes of these statistically significant correlations are small. Across their 40 participants, the correlation between the systemising quotient and mu suppression was -0.124, and for the personal distress scale it was 0.118.

Other studies have found moderate relationships between mu responses and questionnaires. In a later MEG study, this group considered mu suppression to viewing painful versus non-painful stimuli, and found a correlation between mu suppression and the perspective taking subscale of the IRI of 0.36 and 0.37 (the correlations were given separately for the right and left hemispheres respectively) (Cheng, Yang, Lin, Lee, & Decety, 2008). Woodruff, Martin, and Bilyk (2011) investigated the relationship between mu suppression and self-other discrimination, a key component of contemporary theories about empathy and perspective taking. In their sample of 39 participants, they found a moderate correlation between the perspective taking component of the IRI, and the *difference* between mu power between execution and observation conditions. The greater the difference, the higher the score on the questionnaire (r =-0.36).

Yet other studies have reported failing to find correlations between mu suppression and measures of empathy (Moore et al., 2012; Silas, Levy, Nielsen, Slade, & Holmes, 2010; Perry, Troje, & Bentin, 2010. The latter study actually found a significant correlation between mu suppression and empathy is the *opposite* of the predicted direction). When Silas et al. (2010) investigated the associations between socio-emotional scales, mu suppression and gender in their sample of 33 participants, they did find that mu suppression was stronger in females, and that females scored higher on self-report socioemotional questionnaires – but there were no correlations between individual differences and mu suppression. They suggest that while the sex difference in mu suppression may be real, it is unrelated to socio-cognitive abilities.

Leading on from work on empathy, social scientists have also considered how mu suppression may be used to study the neural mechanisms for intergroup relations and prejudice. Drawing on the perception-action model of empathy, Gutsell et al. (2010) hypothesised that individuals with more prejudice would show reduced mu suppression to the

outgroup: "*These* [intergroup] *biases… might be a manifestation of a more basic and general bias: perception–action-coupling for gross motor responses – the physiological process thought to be at the core of interpersonal sensitivity – might be impaired in response to disliked outgroups. Such a fundamental bias, would not only make it difficult to empathize with outgroup members' suffering, but also to understand their actions and intentions, potentially hampering smooth intergroup interactions and communication"* (p. 842). In a sample of 30 Caucasians, Gutsell et al. (2010) found significant differences between the mu suppression towards ingroup versus outgroup members, and significant correlations between prejudice scores and mu suppression towards outgroups. The correlation they report is moderately large ($r = 0.52$). (Arguably, one could link Gutsell et al.'s findings to those of Oberman et al. (2008), described in the later in Section 5.6 *Mu suppression studies of autism spectrum disorders*. Oberman et al. (2008) found that mu suppression in their autistic sample was modulated by familiarity with the model – presumably ingroup members are much more familiar with their own group.) Gutsell and Inzlicht (2013) discuss further research that followed on from these findings, which suggests that mu-suppression biases can be modified by engaging participants in a perspective-taking task, and that mu-suppression biases correlate with beliefs about genetic overlap between different racial groups.

5.5.2 Theory of mind

Despite considerable amount of research on empathy and mu suppression, only one study was found that used mu suppression to investigate MNS involvement in theory of mind. Pineda and Hecht (2008) argue their mu suppression study of 23 participants provides evidence of a dissociation of different theory of mind routes. They appeal to a theory of mind framework by Tager-Flusberg and Sullivan (2000) which suggests that theory of mind could be considered as having socio-cognitive and socio-perceptive components. (One could broadly link the socio-perceptive component to the simulation account of theory of mind

outlined earlier, while the socio-cognitive account may be thought of as similar to the "theory" theory of mind approach.)

Pineda and Hecht (2008) employed tasks argued to measure these different socio-perceptive and socio-cognitive components. To measure socio-perceptive processes, they used a task that required participants to match images of eyes, based on the eyes' emotion, race or gender (the latter two acting as control tasks). For the socio-cognitive processes, they used a cartoon task, in which participants guessed the last panel of a comic strip. The comics require either mental attribution (understanding what the person is intending to do), or an understanding of physical causality. For the physical causality comics, some contained characters, but intention reading was not required (e.g. seeing someone's scarf blown off by the wind), while others contained no characters at all (e.g. seeing a bomb explode). The authors argue that their results support a distinction between socio-cognitive and socio-perceptive tasks, and that the MNS is more involved in socio-perceptual than in socio-cognitive tasks. This would be in keeping with the notion that the MNS underlies a simulation mechanism that allows us to experience and understand others' minds.

However, the results of this study are difficult to interpret. A direct comparison of the strength of mu suppression in the socio-cognitive and socio-perceptive tasks is not reported - so it is not possible to say whether socio-perceptive tasks result in greater mu suppression. Furthermore, the pattern of suppression across the tasks does not clearly demonstrate a difference between socio-cognitive and socio-perceptive tasks. For example, while significant suppression was seen during the emotion matching task, significantly stronger suppression was seen during the race-matching task (though the authors interpret this as showing mirror neuron activity while participants make these race judgements). The authors also argue that the task with comics depicting physical-causality with characters is actually a socio-perceptive task, as there are correlations between this and the emotion eye-matching task. Yet

it is difficult to see why this physical-causality comic task would tap in the MNS, yet the intention-reading comic task would not. The key difference between these two tasks was the need to infer intention. This would seem, at face value, to be exactly the kind of skill the MNS was originally proposed to underlie. Pineda and Hecht (2008) argue that the intention-reading task is more of a "theory" theory task, requiring empirical knowledge and social-cognition, while the physical-causality character and emotion eye-matching task are socio-perceptive tasks, resulting in MNS activation. However, the distinction between these two types of comics (with one being called a socio-cognitive task and the other a socio-perceptive task) feels arbitrary (especially considering the authors conclude that probably both routes are active in both tasks, and mu suppression occurs in all of them).

5.5.3 Biological motion

Several studies have considered mu responses to biological point light displays. These displays are image sequences created by marking the limb movements of moving bodies with lights. These stimuli offer a solution to the problems of presenting well-matched stimuli to investigate mu's responses – social versus non-social stimuli typically differ on a number of basic perceptual factors, while point-light displays allow for a tighter control over such variables.

Mu suppression to these displays has been argued to evidence that mirror neurons are involved in the processing of biological motion. In a study of 20 participants, Ulloa and Pineda (2007) found significant mu suppression to biological point light displays, but not scrambled motion displays. They argued that their effects were not due to attentional differences, as performance on a continuous performance task did not differ between these conditions – however, no results are reported for regions outside the central electrodes.

In fact, other authors examining mu suppression to point light displays have warned about potential confounding effects from occipital alpha and attentional differences in the different conditions. Perry, Troje and Bentin (2010) examined participants' ability to recognise the different dimensions represented in the point-light displays (emotion portrayed, gender of the model, direction of walking, and direction of rolling for the non-biological point light displays). Participants were slower and sometimes less accurate to make decisions about some of the social dimensions represented in the displays (emotion, gender, intention) than direction of rolling in the non-biological motion condition, suggesting that these tasks were not matched for task difficulty. Furthermore, in the analysis of their EEG data of 24 of their participants, they report results from the occipital regions which showed significant alpha suppression across the conditions, and a pattern of suppression similar to that found at the central sites. Possibly, the authors suggest, biological point-light displays may attract more attention, as these factors have higher ecological value (e.g. needing to know whether someone is walking towards or away from you).

5.5.4 Summary of mu suppression in social processes

The findings so far relating mu suppression to social processes are varied. Studies attempting to use mu suppression as an individual differences measure of the mirror neuron system seem particularly problematic. Few strong correlations have been observed (almost none that survive corrections for multiple comparisons), and it is unclear whether other theories not pertaining to mirror neurons could account for some of the findings. Certainly, many studies do not sufficiently consider the potential confounds of alpha or attentional effects in their designs (Perry, Troje, et al., 2010 is a notable exception that clearly demonstrates the importance of considering such issues). For example, it seems quite plausible that viewing ingroup and outgroup members could have differential attentional

effects. Ingroup members may therefore be more likely to engage our attention, suppressing alpha (rather than mu).

Another consideration is whether participants' own movement could confound the effects. Participants generating their own movement would obviously lead to motor cortex engagement and mu suppression. Therefore, if conditions are able to vary in the amount of movements participants performed, these differences could confound the results of mu suppression studies. Instructing participants not to move may not always be enough, as there is evidence that individuals may mimic others without awareness; automatic mimicry is a phenomenon in which individuals unconsciously mimic the actions or postures they perceive in others. Indeed, automatic mimicry could feasibly mediate reported relationships between empathy and mu suppression as there is evidence that high-empathy participants mimic more than participants than low-empathy participants (Sonnby-Borgstrom, 2002). Across the field, many mu suppression studies do take the step to record EMGs from their participants, in order to discard trials in which participants move, but this may be a particular issue for mu suppression and empathy research, as automatic mimicry was little discussed in these papers. Overall, there seems in many of these studies no need to appeal to a mirror neuron theory account –alternative accounts could explain these results just as well.

5.6 Mu suppression studies of autism spectrum disorders

In parallel to work on mu suppression and empathy and language, mu suppression has also been used to examine the functionality of the mirror neuron system in autism spectrum disorders (ASD). Theories that self-other representations may be impaired in autism arose independently of mirror neurons – Rogers and Pennington (1991) suggested that such impaired representations could account for the reported imitation difficulties seen in this group, and broader social and communication problems. Most recent reviews have argued

that the well-documented imitation difficulties in ASD could be due to an abnormality in the mirror neuron system (Williams, Whiten, Suddendorf, & Perrett, 2001), and following suggestions that mu suppression could represent a proxy measure for mirror neuron activity, it was quite logical to use mu suppression set-ups in samples of individuals with autism. Indeed, mu suppression set-ups are likely to be better tolerated by individuals on the spectrum than other imaging methods, such as fMRI.

It should be noted that while autism is certainly the condition most investigated using mu suppression, an outstanding question is to what extent mu suppression deficits are *specific* to autism. Studies so far have also found reduced mu suppression to biological motion in Down Syndrome (Virji-Babul et al., 2008) and first-episode psychosis (Singh, Pineda, & Cadenhead, 2011). Finding mu suppression deficits across a range of disorders could suggest that atypical mu suppression responses are reflective of quite general abnormalities in brain development, rather than specifically an abnormality in the MNS.

In any case, so far there have been nine studies comparing groups of individuals with an ASD to typical control groups. These studies are summarised in Table 5.1. There have also been investigations into the benefits of mu neurofeedback training, which have argued that this may represent a potential treatment for autism in the future (Friedrich et al., 2015; Pineda et al., 2008; Pineda et al., 2011; Pineda, Carrasco, Datko, Pillen, & Schalles, 2014). At present, mu suppression findings with autistic groups have been decidedly varied, with half of the studies concluding that mu suppression during observations of actions is deficient in autism (suggesting abnormal or impaired mirror neuron systems), and half finding mu suppression comparable to controls.

The most recent paper to investigate mu suppression abnormalities in autism points towards abnormalities in the mu frequency band, but suggests these abnormalities arise from

areas not usually associated with mu, but rather with alpha. When only examining the central electrodes, such as is typically done in mu suppression experiments, Dumas, Soussignan, Hugueville, Martinerie, and Nadel (2014) did replicate previous reports of reduced suppression to actions with objects. However, when considering differences across the whole scalp, Dumas et al. (2014) found abnormalities in the alpha frequency band in the frontal and occipital regions in their participants with ASD. Indeed, there is evidence that the broader alpha band, as opposed to mu, is abnormal in ASD; Mathewson et al. (2012) noted in their study that participants in the ASD group had greater alpha power in an eyes-open condition, and that they showed smaller occipital alpha suppression when comparing eyes-open to eyes-closed conditions than typical controls. Reduced suppression in the alpha band is therefore not specific to mu regions or biological stimuli.

Furthermore, it is plausible that attention may be different between ASD and typical participants when viewing biological motion, and that this could be reflected in differences in alpha activity. Attention to social stimuli has been shown to be abnormal in ASD (see Ceponiene et al., 2003 and Fletcher-Watson, Leekam, Benson, Frank, & Findlay, 2009, for examples and discussion of these issues in both auditory and visual domains, respectively). Previous mu suppression reports argued that attentional differences could not account for mu suppression findings in autism, because behaviourally measured attention (usually measured though continuous performance tasks) showed no differences between groups. However these tasks represent rather gross measures of task engagement, and participants groups score 100%, suggesting that the tasks are almost certainly too easy to pick up on any subtle attentional differences between groups.

Potentially even the positive neurofeedback training findings could be, at least in part, explained using attentional theories. In the investigation by Pineda (2008), in both studies contained in this report, the experimental group showed improvements in measures of ADHD

symptoms and sustained attention ability. No differences in improvements in imitation skills were found when comparing placebo and experimental groups, and improvements in communication, perception, and sociability were inconsistent between studies one and two. Could alleged mu suppression neurofeedback training have had more of an effect on general attention and alpha than on mirror neuron systems and the skills it has been suggested to underpin?

If the "broken mirror hypothesis" is true, and mu suppression represents a good measure of mirror neuron activity, then correlations and associations between mu suppression and behavioural measures supposedly supported by the MNS and impaired in autism should be evident. Imitation skills are particularly important in this theory – they are a key link between the original cognitive theories about self-other representation in autism and MNS theories now. One suggestion has been that mu suppression abnormalities pertain more to variation in imitation ability than to autism *per se*. Associations between imitation and mu suppression have been reported (Bernier, Dawson, Webb, & Murias, 2007; Bernier, Aaronson, & McPartland, 2013) but these have not been well replicated. Ruysschaert, Warreyn, Wiersema, Oostra, and Roeyers (2014) found no association between the quality of children's imitation and mu suppression. Fan, Decety, Yang, Liu, and Cheng (2010) noted that their ASD group had poor imitation skills and yet showed intact mu suppression. Following neurofeedback training with mu there were no differences between the improvements in imitation between the placebo or experimental groups (Pineda et al., 2008). (Further discussion and experimental data on the relationship between imitation and mu suppression can be found in Chapter 7.)

One final comment on mu suppression studies of ASD is on the analyses of these studies. Few of these studies seemed to correct for multiple comparisons in their ANOVAs. Furthermore, of the studies that have analysed their data using ANOVA techniques *none* have

found a significant group effect, nor a significant group by condition interaction in their ANOVAs (Bernier et al., 2013, 2007; Fan et al., 2010; Oberman et al., 2008; Raymaekers, Wiersema, & Roeyers, 2009; Ruysschaert et al., 2014). Thus, it is questionable whether a selective abnormality in mu suppression to observing others is really borne out by current data.

5.6.1 Summary on mu suppression studies of autism spectrum disorders

In sum, the evidence for a specific lack of mu suppression to observing human actions in autism is poor. Rather, the collective evidence suggests there may be differences in the alpha frequency band that are not related to MNS dysfunction. As yet, there is no evidence that mu suppression abnormalities are truly specific to ASD, and robust findings linking mu suppression in autism to imitation skills, a behaviour the MNS has been purported to underpin, are lacking.

5.7 Mu suppression in the future - improving mu suppression studies

Mu suppression started out as a measure of motor cortex engagement but has been reinvented as a proxy for mirror neuron activity. Theoretically, mu suppression could be a comparably cheap, tolerable and simple way to examine mirroring systems in the human brain, and these set ups have been used with both individuals with autism and young children, populations for which imaging with other methods may be not practical. Mu suppression has been explored in language tasks, in its association with empathy and social processes, and in autism spectrum conditions. All of these domains have been linked to the human mirror neuron system, but mu suppression studies provide limited support for MNS involvement in these processes. Potentially, this is because mirror neurons are not involved in these processes, or because mu suppression actually represents a poor measure of MNS engagement.

Table 5.1 Summary of mu suppression studies of autism spectrum disorders

Paper	*Sample*	*Stimuli/Conditions*	*Findings*
Oberman et al. (2005)	10 ASD, 10 TD (age and gender matched), aged 6-47 years.	OM; watching video of hand action (opening and closing hand, same as OM condition); watching video of two BB; WN (baseline). Included CPT – 100% accuracy, so inferred that differences between groups are not due to differences in attention.	TDs showed significant mu suppression to OM and observed movements. ASD group showed significant mu suppression during OM only. The lack of suppression in the ASD group was not due to differences in baseline mu power. Neither group showed significant suppression from baseline during the non-biological motion.
Oberman et al. (2008)	13 TD and 13 ASD, aged 8-12 years.	Stranger opening and closing hand; Familiar (sibling/parent) opening and closing hand; video of own hand movement; BB (baseline). Included CPT (100% accuracy).	TD group showed significant suppression to all 3 biological videos. ASD group showed significant mu suppression for videos of familiar hands and their own hands only.
Raymaekers, Wiersema, & Roeyers (2009)	20 HFA, 19 TDs, 8-13 years.	OM (finger tapping); observing moving hand; BB; WN (baseline). Included CPT (100% accuracy).	Same pattern of findings for both TDs and ASD and no significant differences between groups. ASD groups show intact mu suppression to hand videos, and no significant suppression to BB. Significant correlation between intelligence and mu suppression, but not between symptom severity and mu suppression.

Table 5.1 (Continued)

Paper	*Sample*	*Stimuli/Conditions*	*Findings*
Bernier et al. (2007)	14 ASD and 15 TDs, aged 19-43 years.	Rest condition (baseline); observed hand action with object (gripping a manipulandum); execute action (gripping manipulandum to command); imitate action (gripping manipulandum in imitation of videos.	For ASD group, observe condition showed significantly less attenuation then execute and imitate conditions. Imitate and execute did not differ. For TDs, there was no significant differences between execute and imitate, or execute and observe conditions. The groups differed on mu suppression for observe condition – the ASD group showed significantly less mu suppression to TDs. Mu suppression in observe condition was correlated with imitation skills.
Bernier et al. (2013)	19 ASD, 19 TDs, aged around 6 years	Rest condition (baseline); observed hand action with object (gripping a manipulandum); OM (gripping manipulandum to command). Trials in which children did not attend screen were discarded.	Both groups showed mu suppression during OM and observation of hand actions. No correlation with IQ, or communication impairments in ASD. Correlation between imitation of face expression and mu suppression during observation condition. A subset of children who did not show mu suppression to observation condition (5 TD and 2 ASD) had poor face imitation abilities.
Ruysschaert et al. (2014)	18 ASD and 19 TD, aged 24-48 months	Object observation condition (dangling object); Action observation (goal-directed action); action imitation condition (children encouraged to imitate); observing hand movement condition (experimenter did hand actions similar to action observation but with no object).	Imitation score was comparable between groups. Significant mu suppression during hand movement observation, action observation and action imitation task for both groups. No group differences in mu suppression. No correlations with SCQ in ASD group, moderate correlation in TDs. No correlation with quality of behavioural imitation.

Table 5.1 (Continued)			
Paper	*Sample*	*Stimuli/Conditions*	*Findings*
Martineau, Cochin, Magne, & Barthelemy (2008)	14 ASD and 14 TDs, aged 5-7 years.	Rest; static scene viewing (e.g. lake); scenes with motion (e.g. waterfall); video of woman performing movements with her legs.	Showed desynchronisation of the EEG in the motor cortex and the frontal and temporal areas during observation of human actions. No desynchronisation found in autistic children. Note that while commonly cited, this study mainly reports effects for the *theta* band rather than the *alpha* band.
Fan et al. (2010)	20 ASD and 20 TDS, aged 10-26 years.	Manipulating chess piece themselves (OM); observation of hand interacting chess piece; moving white dot; Static cross rest condition (baseline). Included eye tracking.	No visual attention (as measured by fixation) differences found. Participants with ASD failed to imitate the observed actions, but mu suppression was not different to TDs. Mu suppression during observation was associated with the communication competence of individuals with ASD.
Dumas et al. (2014)	10 ASD, 30 TD, aged 20-41 years.	Observation of meaningless hand gestures; Execution of meaningless hand gestures (not to imitation); spontaneous imitation (subjects were told that they could at will either produce hand gestures of their own or imitate the other's hand gestures); video imitation (where participants imitated the videos); rest condition. Whole brain approach, considered different mu bands (large mu band: 8-13, and 2 narrower bands: 8-10, 11-13).	Analysis of the large mu band showed significant mu suppression for TD and ASD for action execution. Only TD showed significant mu suppression during observation of hand gestures. Considering the whole scalp in the observation condition, TD participants showed a significant suppression of the 8-13 Hz band over the whole scalp, more strongly over occipital parietal region. ASD did not show such significant alpha suppression. For upper band only, significant differences for frontal and occipito-parietal region – greater suppression over OP region in TD group, and an increase of alpha in frontal region in ASD group.

As outlined above, there are many reasons to be cautious of mu suppression findings. This review is certainly not the first to point out the potential pitfalls of confounding from alpha and attentional changes, but it seems that investigations using mu suppression as a means to examine mirror neuron function in a variety of language and social processes do not always heeded these issues. Furthermore, mu suppression as a field is tragically underpowered (though this is sadly representative of wider neuroimaging literature -see Button et al., 2013). Future studies will need to be much more robust if important discoveries can be expected to be made about the MNS using this method.

The review has focused largely on data collected from adults in mu suppression studies (with the exception of some of the studies of ASD), however researchers have also used mu suppression studies with infant populations to try and address questions about the development of mirroring systems. Other researchers have reviewed mu suppression with infant populations, so this literature has not been re-examined here (Cuevas et al., 2014). However, it is worth noting that the review by Cuevas et al. (2014) outlines several pertinent problems in many infant mu studies, including the issues of baseline selection and examining changes outside just the sensorimotor regions. They too highlight the need for researchers to consider changes in power at the occipital region, and point out that topographic maps of power distributions across the scalp provided by some infant mu researchers would actually seems to show suppression at the occipital sites. Broadly, the content of infant mu suppression studies has largely been around the processing of others' actions (arguably the traditional remit of mirror neuron theories), rather than broader functions in language and social processes. Work so far has largely concluded that infant mu rhythms show the same patterns of reactivity to participants' own movement and action observation as the adult mu rhythm, and that mu suppression may represent a means to investigate mirror neuron systems in young children (Marshall & Meltzoff, 2011; Marshall, Young, & Meltzoff, 2011; Nyström,

2008). As is apparent from this review, using mu suppression to examine language and social processes in adults has actually produced few robust findings, and so translating these studies for use with infants, where even less is known about the interpretation of EEG, would seem unwise at present.

Furthermore, a comment not so much on the methodology of mu suppression studies but on their interpretation in wider social cognitive neuroscience; the impression one gets when reading the mu suppression literature is that theories about what the human MNS is doing are sufficiently flexible to fit around whatever mu suppression results are obtained. Of course, given that theories about the MNS evolved and developed, we can expect to see mu suppression to stimuli beyond simple hands interacting with objects. But mu suppression has been demonstrated to viewing static buildings, sheet music, and Rorschach ink blots (Behmer & Jantzen, 2011; Moore et al., 2012; Pineda, Giromini, Porcelli, Parolin, & Viglione, 2011). These are a far cry from the original stimuli used to investigate action understanding. Mu suppression as a field seems to be attempting to simultaneously validate mu responsivity as indexing mirror neuron activity, *and* use this responsivity to suggest what stimuli the mirror neuron system is responsive to, but this logic is circular. Ideally, the field need to agree what we expect the human MNS to respond to, examine whether mu suppression meets these expectations, and reject it as a measure of the MNS if it does not meet them.

The results from the study in Chapter 6 help to shed light on the validity of mu suppression as a measure of the human MNS (Hobson & Bishop, in press). While new data will be useful in making progress, this review also sought to reach back for evidence. Mu changes have long been considered to index motor cortex engagement, well before mirror neurons exploded into the field of cognitive neuroscience. Considering mu's history, and how mu studies have changed over the last decade, should lead to reflection on how mu suppression should be conducted in the future.

CHAPTER SIX: Mu suppression – a good measure of the human mirror neuron system?

6.1 Introduction

Since the discovery of "mirror neurons" in the macaque brain, researchers have investigated the presence of such neurons in humans, and considered what the functional role of the human mirror neuron system (MNS) might be. As explored in Chapter 5, the human MNS has been posited to underpin action understanding, imitation, language and empathy, and has even been theorized to be the cause of an evolutionary leap in our ancestral history (Rizzolatti & Arbib, 1998; Rizzolatti & Craighero, 2004. MNS dysfunction has also been proposed to underlie the symptoms of autism spectrum disorders (Dapretto et al., 2006; Rizzolatti, Fabbri-destro, & Cattaneo, 2009; Rizzolatti & Fabbri-Destro, 2010; Williams, Whiten, Suddendorf, & Perrett, 2001).

Mu-suppression has been used to explore the MNS in both typical and autistic individuals. Mu is a range of EEG oscillations from 8-13 Hz, and is recorded from scalp electrodes corresponding to the sensorimotor regions of the brain (typically electrode sites C3, C1, Cz, C2, C4). When a person is at rest, the cells in the sensorimotor cortex fire in synchrony. When a person performs, observes or imagines themselves performing an action, the firing of these cells becomes desynchronised. This desynchronisation leads to reduced mu power, compared to when the cells were firing together (Pfurtscheller, Neuper, Andrew, & Edlinger, 1997). The key design feature of mu suppression studies is the comparison of an experimental condition to a baseline condition in which one would not expect the MNS to become active. If there is a reduction in mu power in the experimental condition compared to the baseline condition, the interpretation is that the experimental condition has activated neurons in sensorimotor cortex. Because mu suppression is seen both when an individual

performs and observes an action, it has been taken as a proxy for the activity of the human mirror neuron system (Muthukumaraswamy & Johnson, 2004; Muthukumaraswamy, Johnson, & McNair, 2004; Oberman, McCleery, Ramachandran, & Pineda, 2007; Pineda, 2005).

Such a relatively inexpensive and noninvasive technique for gauging the activity of the MNS in humans would greatly facilitate research on this system, although not all researchers agree that mu suppression is a valid index of MNS activity(see Chapter 5). Indeed, mu suppression is rapidly becoming an established measure of mirror neuron activity.

As noted in Chapter 5, one of the concerns raised in the literature surrounding mu suppression is whether it is reliably distinct from changes in alpha activity. Alpha activity was among the first EEG phenomena noted by pioneering electroencephalographer, Hans Berger, yet the precise function of alpha is still unknown. Alpha rhythms have been considered to reflect cortical idling (Pfurtscheller, Stancák, & Neuper, 1996), or the active inhibition of task-irrelevant processes (Klimesch, 1999). While the function of alpha activity is unclear, the reactivity of alpha is well documented. Alpha activity is functionally defined as "blocked or attenuated by attention, especially visual, and mental effort" (Niedermeyer & Silva, 2005). Power in the alpha band is highest when a subject is awake with their eyes closed, and suppressed by mental effort, or drowsiness (Niedermeyer & Silva, 2005). More difficult tasks elicit more alpha suppression (Gevins, Smith, McEvoy, & Yu, 1997; Stipacek, Grabner, Neuper, Fink, & Neubauer, 2003).

Mu is in the same frequency band as alpha (8-13 Hz), but alpha and mu are said to be distinguishable on the basis of topography and reactivity. Alpha activity arises in the posterior and occipital regions, while mu arises from the sensorimotor area. While changes in mu power are typically interpreted as being due to activity in the sensorimotor cortex, alpha

power is thought to reflect attentional engagement (Klimesch, 1999; Pfurtscheller, 1992). Nonetheless, because of the overlap between mu and alpha activity, tight controls of attentional engagement should be a key feature of mu suppression experiments.

6.1.1 Evidence for alpha effects in mu suppression studies

Some authors have warned that mu suppression may be sensitive to activity from areas involved in visuomotor processes that are not considered to be part of the mirror neuron system, and also that changes in mu power may be being driven largely by attentional processes rather than mirror neuron activity (Aleksandrov & Tugin, 2012; Braadbaart, Williams, & Waiter, 2013; Perry & Bentin, 2009). Indeed, Perry and Bentin (2010) cautioned that "*mu suppression reports should always include not only experimental effects at the central sites, but also the occipital regions to help fully understand the phenomenon being studied."* (p.1054). As noted in Chapter 5, where previous mu suppression studies have considered activity at occipital electrodes, findings have been mixed. Ruysschaert and colleagues (Ruysschaert, Warreyn, Wiersema, Oostra, & Roeyers, 2014) investigated changes in the alpha frequency band at occipital sites, but only during their imitation condition (when participants actively copied the movement they saw), not during their observation condition. Thus, while mu suppression during actual movement was specific to the central electrodes, it is unclear whether this was also true for this study's observation condition. Tangwiriyasakul et al. (2013) argued that the correlation between central and occipital electrodes was weak, suggesting that their results had not been affected by alpha. However, the correlation between C4 and O2 was 0.49 ($p<0.001$), a not insignificant correlation. Lepage et al. (2008) entered activity from electrode Oz in their analysis, and found that 8-13 Hz power at this site was significantly reduced during observe and imagine conditions. Other papers have reported that other than C3, Cz, and C4, no other electrodes showed a consistent pattern of suppression (Bernier, Dawson, Webb, & Murias, 2007; Bernier, Aaronson, & McPartland, 2013;

Oberman et al., 2005, 2008). Yet the recent study by Dumas and colleagues suggests that apparent mu suppression deficits in autism are not related to the MNS, but rather to alpha (Dumas, Soussignan, Hugueville, Martinerie, & Nadel, 2014). Thus, it seems unclear to what extent changes in mu power at the central electrodes are reliably distinct from changes in power at the occipital electrodes, regions more strongly associated with alpha.

Overall, while concurrent fMRI studies suggest that mu suppression may represent activity in areas considered part of the mirror neuron system (Arnstein, Cui, Keysers, Maurits, & Gazzola, 2011; Braadbaart, Williams, & Waiter, 2013; Mizuhara, 2012; Perry & Bentin, 2009), other processes that are not observation-execution matching also influence changes in mu power. This casts doubt on previous conclusions reached using mu suppression as an index of mirror neuron activity, particularly on higher level sociocognitive processes where the potential effects of attention may not be immediately obvious.

One way of controlling for potential attentional effects is to compare mu suppression to stimuli that are matched in their postulated engagement of the MNS. For example, Muthukumaraswamy et al. (2004) had participants view a hand interacting with an object vs. a hand interacting with itself. Consistent with the predictions from non-human animal work on mirror neurons, they showed greater mu suppression in the former case (Muthukumaraswamy et al., 2004). We cannot, of course, rule out the possibility that a hand interacting with an object is more attentionally engaging, though participant ratings could be used to test this idea. In sum, as mu suppression is becoming a more mainstream method to measure the activity of the MNS, researchers must control for the possibility that attention (and thus alpha) may influence their results.

6.1.2 Choice of baseline in mu suppression experiments

The potential confound of attentional engagement assumes particular importance when considering the range of methods of calculating mu suppression that have been used in previous work. Mu suppression involves comparing power in the mu frequency band during an experimental condition to a baseline. Some researchers have opted to compare the power in their experimental conditions to a single baseline period, some have used an equivalent number of trials of a control condition, while others have baselined each individual trial separately.

Previous whole baseline conditions have included sitting quietly without stimulation, or visual white noise, or a motion control (e.g. Oberman et al., 2005). Clearly, in studies where participants have been asked sit and watch either no videos or videos that are not very engaging for long periods of time, it is feasible that the level of alpha activity would increase, due to attentional disengagement. For example, one study presented videos of visual white noise and bouncing balls that were 80 seconds long (Oberman et al., 2005). Because alpha and mu waves are in the same frequency band (8-13 Hz), this could lead to an inflated ratio between the baseline and experimental conditions, leading to greater mu suppression. Some of these papers attempted to control for alpha by not including the first and last 10 seconds of a stimulus in their analysis, the assumption being that any confounds caused by alpha will take place in these periods, due to the attentional effects of a stimulus initiating or ending (Oberman et al., 2005, 2008)

Other groups have used a baseline of 1 second prior to the onset of each trial as their comparison, either using a fixation cross, or presenting the first frame of the video as a static frame (e.g. Kumar, Riddoch, & Humphreys, 2013; Muthukumaraswamy, Johnson, Gaetz, & Cheyne, 2006). This design is good for removing effects of long-term shifts in the EEG, for instance due to sweating over the time course of the experiment. By baselining each trial individually, such shift is accounted for, and the attentional effects induced by long baseline

conditions are likely to be reduced, and inflated apparent mu suppression is less likely. Nevertheless, it could be argued that the onset of a moving stimulus would immediately engage attention more than a static image.

The issue of what baseline to use in mu suppression experiments was examined by Tangwiriyasakul, Verhagen, van Putten, and Rutten (2013). They recorded EEG data from 18 subjects, investigating what baselines may be ideal for obtaining maximal mu suppression. Their baselines included active and static stimuli, including bouncing balls, slowly moving flowers, static hand images and white stripes on a black screen. No optimal baseline for the whole group emerged – rather, different participants seemed to show bigger mu suppression effects for different baselines. The authors conclude that these findings suggest that calibration may be necessary for motor imagery experiments, in order to identify which baseline is optimal for the individual participant. However, these findings also suggest that mu suppression is not a reliable phenomenon. Indeed, the paper also reports that four of their participants did not show any mu rhythms in any of the five baseline conditions, and two showed mu, but showed no suppression. Thus, mu suppression, with any baseline, was only found for two thirds of their sample. Furthermore, reports from their participants suggest that attentional engagement could have played a role in these results:

"…*many reported that during the BW* [white stripes on a black screen*] baseline it was difficult to maintain attention. Some of them started counting the white stripes on the screen... During the FL* [flower] *baseline, most subjects felt most comfortable and most relaxed; sometimes they lost their attention… During the dynamic baselines (BB and 2B)* [bouncing ball conditions]*, some subjects said that they usually kept their attention to the ball(s)*." (Tangwiriyasakul et al., 2013, p.7)

6.1.3 Beta activity and the mirror neuron system

The convention of many mu suppression studies, particularly those focused on autistic individuals, is to define "mu" as activity in the alpha range (8-13 Hz). However, the rolandic mu rhythm consists of two spectral peaks, and gets its arch-like appearance from the dual contribution of alpha and beta range activity (Niedermeyer & Silva, 2005). Thus, it is important to acknowledge not only the contributions of alpha but also beta activity in the previous findings in the mu suppression literature.

The beta frequency band is usually defined as 13-35 Hz, with a typical peak frequency of ~20Hz (Niedermeyer & Silva, 2005). Beta activity is historically associated with sensorimotor behaviour (although recently it has been suggested that the role of beta in cognitive and attentional processes has been overlooked; see Engel & Fries, 2010 and Gola, Magnuski, Szumska, & Wróbel, 2013). Studies that have looked at both frequency bands suggest that while "rolandic alpha" (mu rhythm) is linked predominantly to the somatosensory system and somatosensory cortex, beta suppression is more related to motor processing and the primary motor cortex (Hari & Salmelin, 1997; Ritter, Moosmann, & Villringer, 2009). Like mu, beta activity is suppressed by voluntary movements, motor imagery and the observation of movements (Babiloni et al., 2002; Hari & Forss, 1998; Mcfarland, Miner, Vaughan, & Wolpaw, 2000), and changes in beta activity have also been suggested to index mirror neuron activity (Muthukumaraswamy & Singh, 2008; Rossi et al., 2002).

One evident risk in this field is that by focusing on one frequency band, we might miss key phenomena of interest. This study focused on alpha and beta frequency ranges, but studies vary in terms of the precise frequency ranges used to define these, and indeed some argue for finer subdivision of these frequency bands (e.g. Pfurtscheller, Neuper, & Krausz, 2000). This, however, carries the complementary risk that if the choice of frequency band is open-ended, this provides 'researcher degrees of freedom' in post hoc analysis (Simmons,

Nelson, & Simonsohn, 2011).To justify distinguishing different frequency ranges, we need studies that distinguish these a priori and consider whether there are reproducible differences in pattern of results between these.

6.1.4 Aims of this study

The aim of this study was to examine the validity of mu suppression as a measure of the human mirror neuron system, particularly in relation to whether conventional mu suppression designs are confounded with changes in alpha activity and attentional engagement, and also to explore whether the reactivity of beta follows the same pattern as mu. This study was conducted as a registered report, and the methodology and analyses plans were pre-registered prior to data collection.

Consistent with previous studies, videos of hand movements were used to elicit mu suppression. A control stimulus that would not elicit mirror neuron activity, but which would be as engaging as the biological movement condition, was also included. For this, kaleidoscope stimuli were chosen. This study examined whether changes in 8-13 Hz power at the central electrodes are distinct from changes in power at this frequency in the occipital regions, and whether high occipital alpha during baseline tasks could be a confounding factor in previous mu suppression designs.

Finally, this study considered three different baselines that previous researchers have used to analyse their mu suppression experiments, and investigate how they might influence the results. These three baselines included long and short rest periods, and a static period at the start of each stimulus. It was hypothesised that a long baseline condition as opposed to brief or trial-by-trial baselines inflates apparent mu suppression.

This study considered how far the results from each of the three baselining methods showed the pattern that is predicted to be a signature of mu suppression, namely an interaction between condition and electrode site, such that the difference in mu suppression (8-13 Hz) between hand vs. kaleidoscope stimuli is greater at the central than the occipital sites. Subsidiary predictions were that mu suppression would be greater for the hand-with-object versus hand-no-object condition, and that the same overall pattern of activity would be seen for the beta frequency (13-35 Hz) as for the mu frequency.

6.2 Method

6.2.1 Participants

The final sample was 61 typical adult participants (see Appendix F for power analyses containing sample size justification, submitted at the time of preregistration). Participants were recruited largely through the University's research participation scheme, and through poster and email advertisements. The final sample included 19 males and 42 females, with a mean age of 22 years (18-33 years). Our sample included 51 right-handed participants, 9 left-handed participants and 1 ambidextrous participant. The participants had no known neurological disorders, nor any diagnoses of autism spectrum conditions. Participants were required not to consume alcohol, or take any psychotropic medication, or any drugs likely to cause drowsiness, for the 8 hours prior to the experiment.

6.2.2 Stimuli

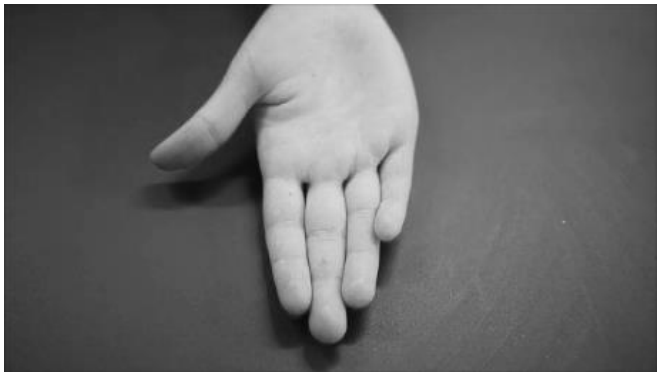
Previous researchers have used a variety of stimuli to test mu suppression to human movement, including: hands grasping a manipulandum (Bernier, Dawson, Webb, & Murias, 2007; Muthukumaraswamy, Johnson, & McNair, 2004), hands manipulating chess pieces (Cheng et al., 2008; Fan, Decety, Yang, Liu, & Cheng, 2010a), a hand opening and closing

with no object (Oberman et al., 2005; Raymaekers, Wiersema, & Roeyers, 2009), mouths sucking or biting, with or without a straw object (Muthukumaraswamy et al., 2006), a hand rotating a coin, or a coin being passed back and forth between two hands (Aleksandrov & Tugin, 2012). The stimuli used in this study were of a hand manipulating a pencil, or performing the exact same manipulative movements but without the pencil. While these stimuli are novel, their features map closely to those previously used in other mu suppression studies. Both object-based and non-object based stimuli were included in the experiment, as previous literature has argued that the presence of an object yields stronger mu suppression (Muthukumaraswamy et al., 2004). It was reasoned that reproducing the object effect would help ensure that the methods and findings are in keeping with and generalizable to other work.

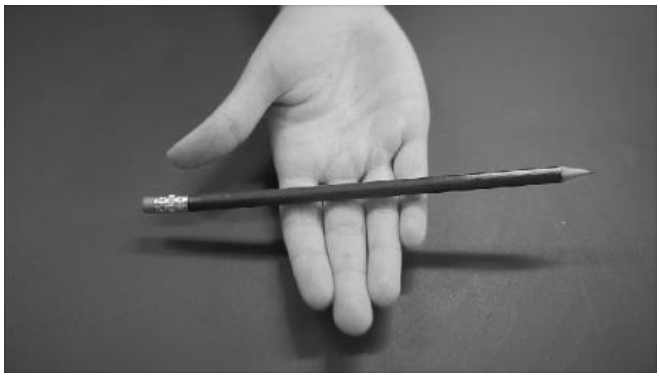
A control stimulus which should not activate the mirror neuron system was also included. Selecting such a stimulus is far from straightforward, as movements of robotic hands have been found to activate mirror neuron areas (Gazzola et al., 2007, but see Tai, Scherfler, Brooks, Sawamoto, & Castiello, 2004), and it has been argued that musical notation can produce significant mu suppression in musicians because of the associations between the sheet music and the movements required to play them (Behmer & Jantzen, 2011). Even stimuli of flowers opening, as used by Tangwiriyasakul et al. (2013), might be argued to be imitable (i.e. you could imagine opening a closed hand to produce a movement that was superficially similar). For these reasons, this study used black and white kaleidoscope videos as control stimuli. If significant mu suppression is seen during the observation of these stimuli it casts serious doubt on the validity of mu suppression as a pure measure of the MNS. Nonetheless, participants were also asked in a post-EEG questionnaire whether they could imagine themselves performing the actions in the videos (see Section 6.2.2.2 *Subjective rating of engagement with stimuli*, and Appendix G).

Equal numbers of trials of videos using the right and the left hand were shown to the participants. The videos can be viewed on the Open Science Framework, under the project title "Mu suppression – a good measure of the human mirror neuron system?" (https://osf.io/yajkz/). Screenshots of the stimuli can be seen in Figure 6.1.

a)



b)



c)

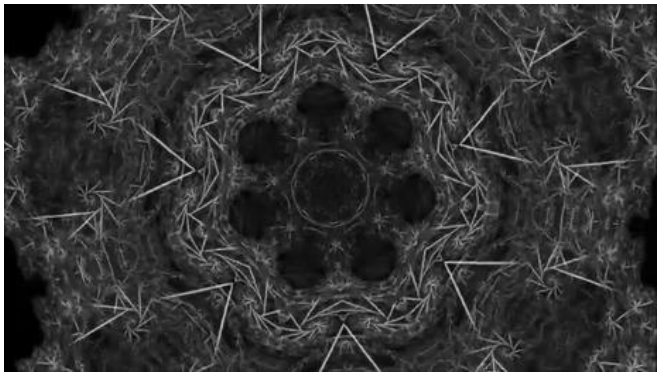


Figure 6.1 Screenshots of the video stimuli used during the EEG recordings. *Participants saw three types of video: a) hands moving with no object; b) hand moving with an object; c) kaleidoscope patterns. All stimuli were black and white. Hand videos were filmed with the use of a metronome to ensure consistency in speed and timing, but were presented to the participants without sound. Videos were filmed with the model using her right hand, and the videos were then digitally flipped to create videos that appear to use the left hand. The hand videos were matched for the movement they showed – they only differed in the presence of a pencil.*

6.2.2.1 Positive control

One of the key characteristics of mu suppression is that it occurs both when a participant observes and performs actions. Not all previous mu suppression investigations have included a movement condition. However, given it is this feature – activation during both observation *and* execution of movements – that has led researchers to propose it as a signature of mirror neuron activity, this investigation included a movement condition, based on a condition used in previous research that successfully elicited mu suppression (Woodruff, Martin, & Bilyk, 2011). This own movement condition acted as an outcome-neutral positive control condition.

6.2.2.2 Subjective rating of engagement with stimuli

To test the hypothesis that previous differences between control and experimental conditions in mu suppression studies are driven in part by differing levels of engagement, participants were asked to rate their subjective levels of engagement in the different conditions. A copy of the post-EEG questionnaire can be found in Appendix G. It was reasoned that if the analysis suggested that apparent mu suppression was being driven by changes in alpha and attentional effects, it would be expected that the pattern of mu suppression seen in the various conditions will follow the same pattern of subjective rating of attentional engagement. This questionnaire also allowed us to check whether participants could imagine themselves performing the "non-imitable" videos, the kaleidoscope patterns, and that they attended to the stimuli sufficiently (see Section 6.2.3.2 *Measure of attention*).

6.2.3 Procedure

The study received approval from the ethics committee at the University of Oxford (Medical Sciences Interdisciplinary Research Ethics Committee Code: C1-2013-190). After

reading the information sheet and signing the informed consent form, participants underwent the EEG. Participants were sat in a quiet room, and watched the stimuli presented to them via a laptop screen. There were three types of EEG condition: a) observing, b) resting and c) moving, based on the conditions used in previous research. In the observing conditions, participants watched the videos of the hand movements and kaleidoscope patterns. During the resting condition, participants were asked to sit quietly but to keep looking at the laptop screen, and not to close their eyes. The EEG conditions and trial types are summarised in Table 6.1.

6.2.3.1. Timings

For each of the observing conditions (hand manipulating pencil, hand with no pencil, kaleidoscope patterns), there were 40 trials. In each video, the first 4 seconds was a static picture of the hand/ kaleidoscope patterns, which served as a baseline (see Section 6.2.5 *Analysis Plan*). These 4 seconds were followed by 2 seconds of movement, and then 2 seconds of a static final frame. The 2 seconds of movement per trial means that each video condition had up to 80 seconds of recording while participants observed the moving videos. The observing conditions were closely modelled on previous work (e.g. Muthukumaraswamy et al., 2006).

The resting condition was modelled on Bernier et al. (2007); participants were asked to sit quietly in front of a blank screen. A single long rest interval of 80 seconds was included in each session, as well as short rest intervals of 8 seconds each, interspersed within blocks of other stimuli.

Table 6.1: Conditions during EEG recording

Condition	*Trial Type*	*Description*
Observation condition	**Hand action with object (HO)**	8 second videos, in which a hand interacts with an object (a pencil). 40 trials in total.
	Hand action without object (HNO)	8 second videos, in which a hand performs actions. There is no object in this video. 40 trials in total.
	Kaleidoscope pattern (KAL)	8 second videos of a kaleidoscope pattern. 40 trials in total.
Rest condition	**Short rest baseline condition**	8 second period of a blank screen. Participants instructed not to move, just like in the video conditions. 40 trials in total.
	Long rest baseline condition	80 second period of a blank screen. Participants instructed not to move, just like in the video conditions. This condition is presented as one continuous trial and later epoched into 2 second periods.
Own movement condition		40 second period in which participants are instructed to tap their finger and thumb. Four 40 second periods for the right hand, and four for the left hand (8 in total).

Table 6.1: *The conditions during the EEG recording.*

For the own movement condition, participants were asked to tap their index finger and thumb together at a steady pace for 40 seconds. This was done four times with each hand. Previous mu suppression research has used this movement to elicit mu suppression (Woodruff et al., 2011). The experimenter was able to watch the participant through a tinted window to ensure that they performed the finger tapping action.

Video stimuli were shown in 8 blocks of 15, with videos playing back to back, except for 5 short rest trials (blank screen) included within each video block. Trials within each block were presented in a semi-random order. The order was constrained, such that a rest trial could not follow another rest trial (to ensure all short rest periods are 8 seconds long, not 16 seconds). Aside from this constraint on the rest trials, videos were presented in a mixed, random order, and each block of 15 videos included hand and kaleidoscope videos (i.e. videos were not blocked by type). The video/rest blocks were interleaved with the movement trials, such that participants watched 2 minutes of videos (with 5 short rest trials), then performed 40 seconds of the finger tapping movement, then watched 2 minutes of videos, and so on. This interleaving of trials was intended to keep participants alert during the EEG. The position of the long resting condition was counterbalanced across participants to occur at one of four places in the experiment – at the beginning, after 2 blocks of videos, after 4 blocks of videos, or after all 8 blocks of videos.

6.2.3.2 Measure of attention

In order to confirm that all participants included in the final analysis viewed and attended the stimuli properly, a coarse behavioural measure of attention was included. Previous studies into mu suppression have used continuous performance tasks as a means of ensuring their participants attended the stimuli (Oberman, Pineda, & Ramachandran, 2007; Oberman et al., 2008). These tasks have typically taken the form of counting a particular event. However, as noted above, alpha activity is known to be affected by mental activity (indeed, previous investigations of alpha have utilised counting targets as a task - see Klimesch, Doppelmayr, Russegger, Pachinger, & Schwaiger, 1998). Therefore, a sufficiently "light" cognitive task is required, so as not to influence the EEG. In the current study, participants were told prior to the EEG recording that they would be asked questions about what they saw during the experiment at the end. During the EEG recording, three grey stars

and three grey arrows appeared on the screen, following or preceding videos or rest periods, but never interrupting them. The stimuli were presented for 1 second each time. Following the recording, participants were asked if they noticed anything during the experiment that was not a video of hands or kaleidoscope patterns. Participants who failed to report any of these extra stimuli, or inaccurately reported how many times these stimuli appeared were considered not to have attended to the stimuli properly, and were excluded from the analyses. While this is arguably a coarse measure of attention, it was reasoned that this task would motivate participants to attend to the stimuli properly, and identify any participants who were unable to do so.

6.2.4 Electrophysiological recording

EEG data were collected from 36 electrodes embedded in a cap using the 10-20 method of electrode placement, including 4 electro-oculograms (above and below the right eye, and to the sides of outer corners of each), and 2 electrodes on the mastoids. Electrolytic gel was applied at each electrode site to reduce the impedance of the electrode–skin contact. The impedance on all electrodes was measured and confirmed to be less than 40 KΩ both before and after testing. Recording was made at a sampling rate of 1000 Hz. The EEG data was recorded using a Neuroscan Nuamps system, and analysed using EEGLAB (Delorme & Makeig, 2004). All recordings were continuous, with no filters applied at the recording stage. Markers identifying the trial type were recorded at the start of the trial for each video and short rest trial, every 8 seconds in the own movement condition, and every 2 seconds during the long rest period. This allowed for the extraction of a similar number of 2 second intervals from the long rest period as for each of the movement portions of the observing conditions.

6.2.4.1 Electromyography

Viewing hand movements could lead to some automatic imitation, even if participants are instructed to remain as still as possible. In order to identify and exclude rest or observation trials in which participants generated muscle activity, an electromyogram (EMG) was recorded from the extensor digitorum communis (the arm muscle that extends the fingers). EMGs were recorded from these muscles on both the left and the right arm, using disposable ECG electrodes, at a sampling rate of 1000 Hz. The EMG data was recorded as additional channels in our EEG dataset and made bipolar in our analysis script. Details on how the EMG data was used to exclude movement trials can be found in Section 6.2.5 *Analysis Plan*.

6.2.4.2 Current Source Density

EEG data was transformed to a "reference-free" format using current source density (CSD) transformations. CSD estimates are second spatial derivatives of recorded field potentials (see Tenke & Kayser, 2005 for more details on CSD). CSD is essentially a spatial filter that minimises the problem of volume conduction, providing more accurate topographical results.

6.2.5 Analysis Plan

Analysis was conducted using the following steps, using EEGLAB version 6.1 run in MATLAB. The script for analysing the data is available on Open Science Framework (https://osf.io/yajkz/).

Using this script, the continuous file was first epoched into segments starting at onset of the trial marker (0 seconds) and lasting for 7 seconds. All trials were baselined to be centred on an average of zero. Trials containing extreme values (greater than 350 μv) other than eye channels or frontopolar channels were removed. This is a much more extreme cut-

off than is usually used because the goal at this point was just to remove trials with excessive movement artefact, but not to remove blinks.

Any observation or rest trials in which the EMG activity recorded from the electrodes was above an individualised threshold then removed. A non-active EMG was subtracted from the EMG recorded from the extensor digitorum communis to create a bipolar channel. The EMG activity in the own movement conditions was converted to root mean square values across all own movement trials, separately for the left and right arm. A threshold of 1.5 standard deviations below this average was used to remove trials in the rest or observation conditions that show muscle activity greater than this value.

The bipolar eye channels were subtracted to give one channel for vertical eye movements and another for horizontal eye movements. Data were then subjected to independent component analysis using single-order blind identification (see Bishop, Hardiman, & Barry, 2011). This was achieved by transforming the weight matrix for components into z-scores across all electrodes, and identifying those that have a z-score greater than 4.0. This is an arbitrary large value which has been determined in previous studies to identify signals due to blinks or to other artefact. Components whose activity is heavily focused on a single electrode were then subtracted from the signal.

To be included in the final analysis, a minimum of 16 trials per condition were required, after bad trials were rejected. Following the rejection of bad epochs, the remaining data were re-referenced offline to a CSD derivation using a CSD MATLAB Toolbox (Kayser & Tenke, 2006a, 2006b). The functions in the Toolbox were utilised by our analyses scripts. The Toolbox is freely available here: http://psychophysiology.cpmc.columbia.edu/software/CSDtoolbox/index.html

The analysis was restricted to the sensorimotor and occipital electrodes C3, Cz, C4, O1, Oz and O3. Three methods for estimating mu suppression were compared, where the period from 2 seconds to 4 seconds post-trial onset is described as the early interval, and the period from 4 to 6 seconds post-trial onset as the late interval. Note that these terms correspond to the static and active portions of the trials where hand stimuli are used. A frequency decomposition was conducted using the EEGLAB "spectopo" function, separately for early and late intervals for each of the six conditions: (a) Hand No Object, (b) Hand with Object, (c) Kaleidoscope patterns, (d) Short fixed stimulus; (e) Long fixed stimulus (f) Own Movement. Mean log power in the frequency range 8 to 13 Hz is defined as $10*\log_{10}(\mu v^2/f)$, where f is frequency in Hz. The three methods are as follows:

Method 1. Within-trial baseline: Mean log power in the early interval was subtracted from mean log power in the late interval for all three observe conditions.

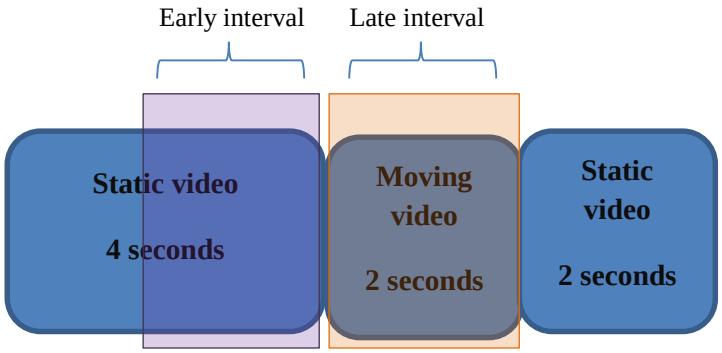
Method 2. Between-trial baseline: Mean log power in the late interval for the short rest trials was subtracted from that in the late interval for trials with hand or kaleidoscope stimuli, and own movement condition.

Method 3. Single long baseline: Mean log power in the long rest period was subtracted from that in the late interval for trials with hand or kaleidoscope stimuli, and own movement condition. In addition, as a further control, log power in the long rest period was subtracted from mean log power in the late interval for the short fixed stimuli trials: a contrast where no mu suppression should be observed. Figure 6.2 is a diagram depicting the three baselining methods. For the main analysis, three 2-way repeated measure ANOVAs were conducted, for the three different baselining methods (short rest trials, long rest trials, and trial-by-trial baselines). In each analysis, the first factor is condition and the second factor is site (central and occipital). For the comparisons with rest trial baselines, all four conditions (hand no

object, hand with object, kaleidoscope patterns and own movement) were compared. For the trial-by-trial baseline, the own movement condition is excluded, since the same movement is executed continuously. Electrodes C3, Cz and C4 are averaged together, as are electrodes O1, Oz and O3.

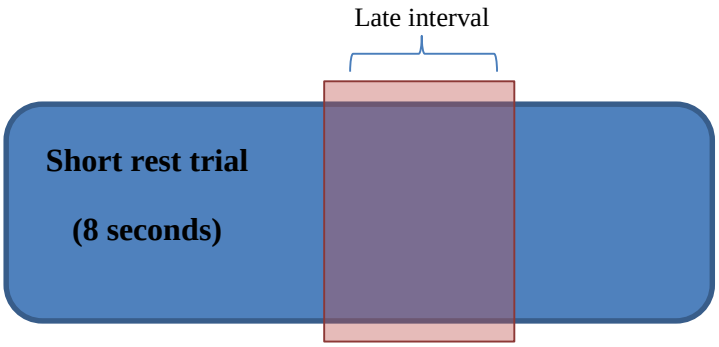
1. **Within-trial baseline**

Late interval compared to early interval for each trial.



2. **Between-trial baseline**

Late interval in short rest trial compared to late interval in video/own movement



3. **Single long baseline**

Long rest condition (split into 40 x 2 second epochs) for comparison to late interval in video/own movement trial

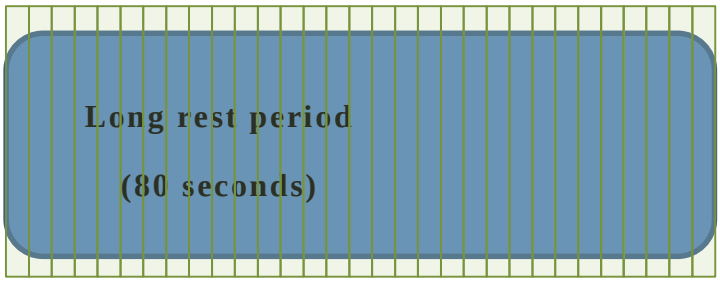


Figure 6.2: A diagram depicting the analysis using the three different baselining techniques*. The period of the video in which the hand/kaleidoscope pattern moves (the orange section) is compared against one of three baselines:*

1. *The 2 seconds early interval period immediately preceding the video when there is a static picture presented (purple);*
2. *The average power of the late interval period during the short (8 second) rest trials (red);*
3. *The average power of the late interval periods in the long rest condition (green). The long rest period is composed of 40 x 2 second epochs.*

In each 40 second own movement trial, there are 5 triggers every 8 seconds used to divide the movement trials up into 5 epochs. The own movement trials are analysed the same way as the video trials, comparing the power in the late interval the late interval in the short rest trials and the average power in the long rest condition.

Results were analysed using repeated-measures ANOVAs rather than paired comparisons, so that specific interactions between condition and electrode site could be tested. Because the three baselining methods are not independent, no direct comparisons were made between them. Rather, I considered how far any of the three methods showed the pattern of results that is predicted to be a signature of mu suppression, namely: on ANOVA, an interaction between condition and electrode site should be seen, such that the difference in suppression between hand-with-object vs. kaleidoscope stimuli is greater at the central than the occipital sites.

In addition, it was predicted that in the positive control condition (own hand movements) significant suppression of 8 – 13 Hz power would be seen at central sites (tested using one-sample t-test to compare observed power change to zero). Predictions about the hand-no-object observation condition are less clear-cut. The early mirror neuron theory focused on grasping of objects, and would not necessarily predict any MNS activity for these stimuli, but subsequent studies of mu suppression suggest it can occur with no object (e.g. Cochin & Barthelemy, 1999; Muthukumaraswamy et al., 2004). Following Muthkurawasamy et al. (2004) an interaction, such that mu suppression would be greater for the hand-with-object versus hand-no-object condition, was predicted. Finally, it was predicted that the same overall pattern of activity across baselines and stimuli would be seen for the beta frequency (13-35 Hz) as for the mu frequency. In order to limit the chance of Type I error, electrode sites (C3, Cz, C4, O1, Oz and O3) were pre-selected.

If suppression of 8-13 Hz activity is seen to hand stimuli, but with a similar pattern of results for the central and occipital sites, this would suggest that differences between conditions could be accounted for by changes in alpha activity associated with attentional changes. In this case, it was planned that the results from the engagement questionnaire would be used as a covariate to see if this could account for these results.

6.3 Results

6.3.1 Excluded participants

In total, 109 participants were recruited during the course of the study. 27 participants were excluded for failing the attention check task. A further 13 participants were excluded as reliable EMG signals could not be obtained; as the EMG recordings were used to retain or exclude trials we could only include participants with data from these channels. A further three were excluded due to very poor EEG recordings, and five recorded datasets were found to have had too many trials rejected by our analysis script, and were therefore replaced with new participants. In total, 48 participants were excluded. In the final sample of 61 participants, a high number of trials were retained for each condition, following automated rejection in our analysis script (see Table 6.2).

Table 6.2: Trials retained for each condition

Condition	HNO	HO	Kaleidoscope	Own movement	Short rest periods	Long rest periods
Mean no. trials retained	34.08	32.87	33.59	39.07	33.98	33.52

Table 6.2 *Average number of trials retained per condition in final sample of 61 participants. Note that from the participants' perspective, there is only one long rest trial – the average presented here represents number of epochs retained. HNO = hand no object trials; HO = hand with object trials.*

6.3.2 Post-recording questionnaire responses

Table 6.3 shows the responses to the questionnaire, given to participants after the EEG recording session. The results of the engagement questionnaire were to be used as a covariate, if results from the occipital and central sites were found to be the same. However, given that mu suppression in all baseline techniques was weakest for kaleidoscope videos, but this stimulus was rated the most engaging by participants, and given the dissociation between mu and alpha in this condition, this was not deemed appropriate. Analyses on the

questionnaire responses can be found in Appendix H *Supplementary unregistered analyses for mu suppression registered report*. These show that the kaleidoscope videos were rated as the most interesting stimulus (although the actual ratings of engagement were not dissimilar across the conditions).

Table 6.3: Post-recording questionnaire responses

	Kaleidoscope	*HNO*	*HO*	*Rest period*
% Rated most interesting	65.6	6.6	26.2	1.6
% Rated least interesting	4.9	14.8	1.6	78.7
Average engagement	3.57 (1.02)	2.97 (.95)	3.39 (.95)	2.05 (1.04)
Average difficulty to perform	4.39 (1.05)	1.98 (1.11)	1.78 (1.02)	*N/A*
% Judged could imitate	11.5	96.7	95.1	*N/A*

Table 6.3: *Responses to the post-recording questionnaire. Numbers in parentheses represent standard deviation. When rating average engagement, participants were asked to rate on a scale of 1-5 with 5 being very engaged. When rating difficulty to perform, participants were asked to rate on a scale of 1-5 with 5 being very difficult to perform. The questionnaire can be found in Appendix G.*

6.3.3 Results for the single long baseline

For each baseline technique, a two-way ANOVA was run, followed by the six planned comparisons (hand-object versus kaleidoscope, hand-no object versus kaleidoscope, and hand-object versus hand-no object, both at the central and the occipital sites). Correction for multiple comparisons was not performed since comparisons were planned before the data was collected. The mean changes in mu/beta power and standard errors are shown in Tables 6.4 and 6.5.

The results when the single long baseline condition was used to calculate mu/alpha (8-13 Hz) and beta (13-35 Hz) suppression is considered first. Figures 6.3A and B show the

results using this baseline. For the mu band, there was a significant effect of site: F (1,60) = 5.36, p = .024. Condition did not have a significant effect. There was also a significant interaction: F (1.36, 81.74) = 79.83, p <.001. Contrasts comparing suppression across the video conditions revealed that changes in the 8-13 Hz band were significantly different between the kaleidoscope and hand-object conditions at the occipital sites (F (1, 60) = 14.18, p < .001), but not at the central sites. Similarly, suppression during hand-no-object videos was significantly different from the kaleidoscope videos, in the occipital regions only (F (1,60) = 15.17, p <.001). Central mu suppression for hand-object and hand-no-object videos did not significantly differ. One-sample t-tests found that none of the video conditions had average suppression that was significantly lower than 0 at the central sites, however the own movement condition produced average mu suppression significantly below 0: t (60) = -6.25, p <.001.

For the beta band, there were no significant main effects of site or condition, but there was a significant interaction effect: F (2.02, 121.37) = 50.72, p <.001. Contrasts comparing suppression in the 13-35 Hz band across the video conditions revealed that suppression for kaleidoscope and hand-object videos at the occipital sites was significantly different (F (1, 60) = 7.05, p = .010), as was suppression for hand-no object and kaleidoscope videos at the occipital sites (F (1, 60) = 9.05, p = .004). Hand-object and hand-no object videos did not significantly differ at either site. One-sample t-tests found that none of the video conditions had average suppression that was significantly different from 0 at the central sites, though

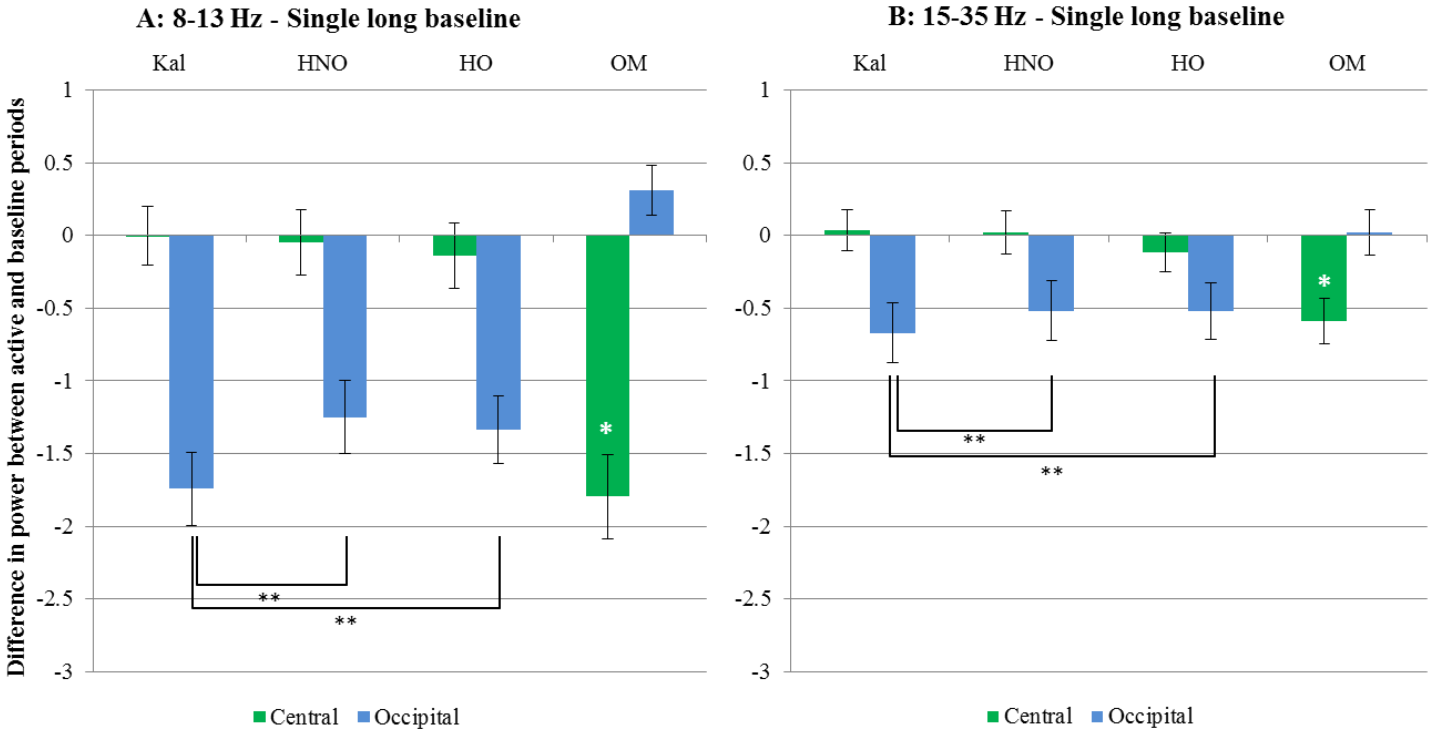


Figure 6.3: Results from the single long baseline. *Graph A shows changes in the 8-13 Hz band (alpha/mu). Graph B shows changes in the 13-35Hz (beta) band. Kal = kaleidoscope, HNO = Hand (no object), HO = Hand (with object), OM = own movement. Error bars are standard error. Planned comparisons between the video conditions that were significant are highlighted and asterisked: * indicates p<.05, ** indicates p<.01. Where one –sample t-tests found that suppression at central sites was significantly below 0, this is marked with a white *.*

suppression to own movement was: $t(60) = -3.84$, $p < .001$.[7]

Overall, with the long baseline, neither mu nor beta showed the pattern corresponding to the mirror neuron hypothesis. The only case where there was a selective suppression at central electrodes was when the participant engaged in hand movement. When observing hand movements, no suppression was seen. The occipital electrodes showed evidence of alpha suppression, which was greatest when observing the kaleidoscope patterns.

6.3.4 Results for the between-trial baseline

The between-trial baseline was calculated by subtracting the average mu or beta power across the short rest trials from the active periods of the video conditions and own movement condition. Figures 6.4A and B show the results. Note that the pattern of differences between conditions will be the same as for the single long baseline analysis – this is because the same averages across the four conditions (the three video types and the own movement condition) are subtracted from a common average, this time based on the average power across the short rest periods. For mu, there was a significant main effect of site: $F(1.60) = 8.34$, $p = .005$. The effect of condition was not significant. There was a significant interaction: $F(1.36, 81.72) = 79.83$, $p < .001$. Contrasts comparing suppression across the video conditions showed that kaleidoscope and hand-object videos differed at the occipital sites only ($F(1, 60) = 14.18$, $p < .001$, the same as for the single long baseline). Similarly,

[7] These planned comparisons are *not* corrected, as they were pre-registered. However, for the 6 comparisons being made for each ANOVA, the Bonferroni corrected alpha cut-off would be 0.05/6 = 0.008. However, there are arguably not just 6 comparisons being made but rather 12 (6 for each frequency band for each baseline). Correcting for this number of comparisons, alpha would be 0.004.

Employing these corrected thresholds would mean that a number of the planned comparisons would no longer be significant. For clarity, the comparisons affected are listed here. For both the single-long baseline and the between-trials baseline, the difference in beta suppression between kaleidoscope and hand object videos would no longer be significant. For the within-trial baseline, for the mu band, the difference between kaleidoscope and hand-object videos would no longer be significant.

suppression during hand-no object videos was significantly different from the kaleidoscope videos, at the occipital regions only (F (1,60) = 15.17, p <.001). Suppression to hand-object and hand-no object videos did not significantly differ at either site. On one-sample t-tests, mu suppression to the video stimuli was significantly below 0 for all three video conditions (For HO: t (60) = -2.85, p = .006; for HNO: t (60) = -2.36, p=.021; for kaleidoscope: t (60) = -2.51, p = .015.) The own movement condition also produced average mu suppression significantly below 0: t (60) = -7.52, p <.001.

For beta, there was no main effect of site, nor condition, but there was a significant interaction between site and condition: F (2.02, 121.37) = 50.72, , p <.001. Contrasts comparing suppression across the video conditions for the 13-35 Hz band found that kaleidoscope and hand-object videos differed significantly at the occipital sites (F (1, 60) = 7.05, p = .010), as did hand-no object and kaleidoscope (F (1, 60) = 9.05, p = .004). Hand-object and hand-no object videos did not significantly differ at either site. One-sample t-tests showed that suppression to all three video conditions were significantly below 0 at the central sites. (For HO: t (60) = -4.99, p <.001; for HNO: t (60) = -3.60, p = .001; for kaleidoscope: t (60) = -3.81, p <.001). Own movement also resulted in beta suppression significantly below 0: t (60) = -6.85, p <.001.

In sum, with the between-trial baseline, the overall pattern of results was similar to that for the long baseline, except that there was evidence of suppression of both mu and beta at central sites. However, this suppression was no different for conditions observing hand movement than for the kaleidoscope condition, indicating it was not a reflection of mirror neuron activity.

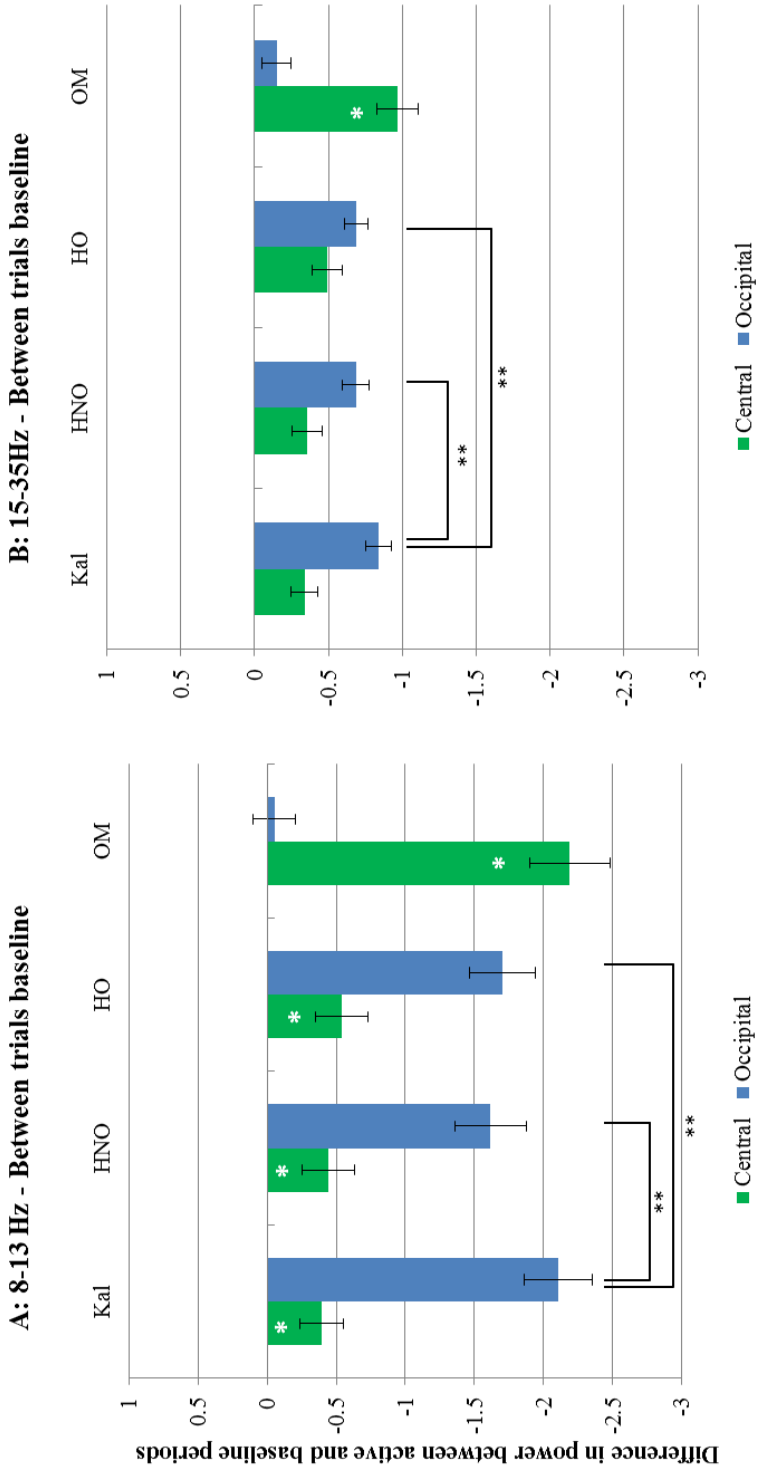


Figure 6.4: Results from the between trials baseline. *Graph A shows changes in the 8-13 Hz band (alpha/mu). Graph B shows changes in the 13-35Hz (beta) band. Kal = kaleidoscope, HNO = Hand (no object), HO = Hand (with object), OM = own movement. Error bars are standard error. Planned comparisons between the video conditions that were significant are highlighted and asterisked: * indicates p<.05, ** indicates p<.01. Where one –sample t-tests found that suppression at central sites was significantly below 0, this is marked with a white *.*

6.3.5 Results for the within-trial baseline

This baseline was calculated by subtracting mu or beta power during the static image component of the videos from the active portion of the videos, on a trial-by-trial basis. Figures 6.5A and B show the results. (Note that standard error is smaller with this baseline, as the active periods of the videos are baselined with the static portions of the videos from the same condition, hence reducing effects of between condition variation.) For mu, there was a main effect of site: F (1, 60) = 36.23, p < .001. There was no main effect of condition. There was also a significant interaction between condition and site: F (2, 120) = 12.93, p < .001. The planned contrasts revealed that the kaleidoscope and hand-object videos were significantly different both at the occipital (F (1, 60) = 11.54, p = .001) and central sites (F (1, 60) = 4.82, p = .032). Hand-no object and kaleidoscope videos were significantly different at the occipital sites only: F (1, 60) = 16.27, p <.001. Hand-object and hand-no object videos did not significantly differ at either site. One-sample t-tests found that only the hand-object videos produced mu suppression that was significantly below 0: t (60) = -2.76, p =.008 (although there was trend for near significance for the hand-no object videos: t (60) = -1.97, p = .054).

For beta, there was a main effect of site: F (1, 60) = 8.154, p = .006. There was no main effect of condition, nor an interaction. None of the planned contrasts were significant. One-sample t-tests found that suppression for all three video conditions were significantly below 0. (For HO: t (60) = -3.74, p <.001; for HNO: t (60) = -2.50, p = .015; for kaleidoscope: t (60) = -2.24, p = .029).

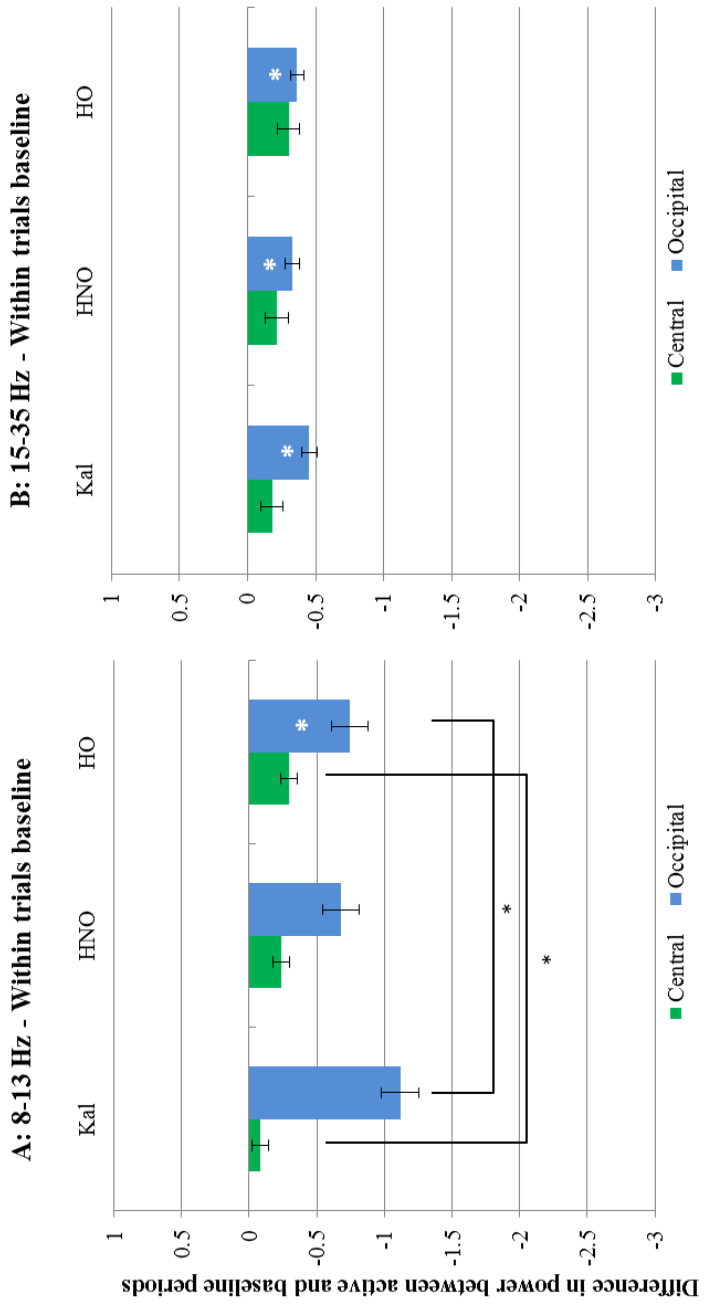


Figure 6.5: Results for the within trial baseline. *Graph A shows changes in the 8-13 Hz band (alpha/mu). Graph B shows changes in the 13-35Hz (beta) band. Kal = kaleidoscope, HNO = Hand (no object), HO = Hand (with object). Error bars are standard error. Planned comparisons between the video conditions that were significant are highlighted and asterisked: * indicates p<.05, ** indicates p<.01. Where one –sample t-tests found that suppression at central sites was significantly below 0, this is marked with a white *.*

To summarise, the within-trial condition was the only baseline to show the predicted pattern of mu suppression that would be consistent with mirror neuron activity. When observing a hand manipulating an object, there was significant mu suppression, whereas this was not seen when observing the kaleidoscope patterns: mu suppression differed significantly between these two conditions. As indicated in Figure 6.4, observing a hand moving without an object showed a trend in the same direction as the hand with object, but this fell short of statistical significance. This pattern was not seen for the beta frequency band, where suppression at central sites was seen to all three types of stimuli, without any difference between conditions.

Table 6.4: Mean changes in 8-13 Hz power

Single-long Baseline				
	Kaleidoscope	HNO	HO	Own Movement
Central	-0.001 (0.20)	-0.050 (0.22)	-0.143 (0.22)	-1.796 (0.29)
Occipital	-1.742 (0.25)	-1.250 (0.25)	-1.337 (0.23)	0.313 (0.17)
Between-trials Baseline				
	Kaleidoscope	HNO	HO	Own Movement
Central	-0.397 (0.16)	-0.445 (0.19)	-0.538 (0.19)	-2.191 (0.29)
Occipital	-2.109 (0.25)	-1.617 (0.26)	-1.704 (0.24)	-0.054 (0.15)
Within-trials Baseline				
	Kaleidoscope	HNO	HO	
Central	-0.085 (0.11)	-0.240 (0.12)	-0.295 (0.11)	
Occipital	-1.118 (0.14)	-0.681 (0.12)	-0.745 (0.11)	

Table 6.4: *Mean changes in alpha-band (8-13 Hz) power for each condition, and baseline technique. Numbers in parentheses represent standard error.*

Table 6.5: Mean changes in 13-35 Hz power

	Kaleidoscope	HNO	HO	Own Movement
Single-long Baseline				
Central	0.039 (0.14)	0.020 (0.15)	-0.116 (0.14)	-0.590 (0.15)
Occipital	-0.672 (0.21)	-0.519 (0.20)	-0.521 (0.20)	0.019 (0.16)
Between-trials Baseline				
Central	-0.338 (0.09)	-0.356 (0.10)	-0.493 (0.10)	-0.968 (0.14)
Occipital	-0.839 (0.08)	-0.686 (0.09)	-0.689 (0.08)	-0.149 (0.10)
Within-trials Baseline				
Central	-0.181 (0.08)	-0.213 (0.09)	-0.303 (0.08)	
Occipital	-0.453 (0.05)	-0.331 (0.05)	-0.362 (0.05)	

Table 6.5: *Mean changes in beta-band (13-35 Hz) power for each condition, and baseline technique. Numbers in parentheses represent standard error.*

6.3.6 Short rest periods versus long rest period

Average mu power from the long rest condition was subtracted from the average power across the short rest periods, and the significance of this difference assessed using a one-sample t-test. This showed that both central mu and occipital alpha power were higher in the short rest periods than the long rest condition, and that the difference was significantly different from 0 (for central sites: t (60) = 3.01 p = .004; for occipital sites: t (60) = 5.42, p <.001).

6.3.7 Unregistered analyses - percentage of participants showing expected mu suppression effects

In addition to the analyses above, I considered how many of the participants showed expected mu suppression effects – that is, mu suppression significantly below 0 when performing and observing actions. The report by Tangwiriyasakul et al. (2013) suggested

around a third of participants do not show predicted effects. This section of the analysis was modelled on their paper, in which they used t-tests to assess for each participant whether or not they demonstrated significant changes in mu for their different baseline techniques.

Thus, for each participant, a 95% confidence interval (CI) was calculated for the change in mu power at electrodes C3, Cz and C4 (the paper by Tangwiriyasakul et al. generally considered channels separately, so we did the same for parity), for the observation and own movement conditions. In order to be considered to have shown the expected mu suppression effect in a given condition, a participant was required to show a CI that did not cross zero (demonstrating mu suppression significantly below 0) for *at least one* electrode site. How many participants showed significant mu suppression during their own movement was also examined; Tangwiriyasakul et al. (2013) did not include an own movement condition, but as mu suppression is considered an index of motor cortex activation this provided a positive control. For this, right hand and left hand movement conditions were considered separately.

Table 6.6 shows the percentage of participants who *failed* to show the expected mu suppression effect at any of the three electrodes, for each hand video condition and own movement condition, by baselining technique. While not all participants showed significant mu suppression to their own movement, between a sixth and a third of participants failed to show the expected suppression effect when observing the hand videos. Consistent with the prior analyses, the greatest proportion showing mu suppression was for the hand-with-object condition using the within-trial baseline (note that it was not possible to baseline the own movement condition with this technique).

Table 6.6: Percentage of participants not showing mu suppression effects

Baseline technique	**Video condition**		**Own movement condition**	
	HNO	*HO*	*R*	*L*
Within-trial baseline	21.3	16.4	N/A	N/A
Between-trial baseline	21.3	21.3	8.2	4.9
Single long baseline	24.6	29.5	4.9	3.3

Table 6.6 *The percentage of participants who do not show expected mu suppression effects at any of the central electrode sites (C3, Cz and C4), for the video conditions and the own movement conditions (shown with the right and left hand separately). A given participant was considered to have shown mu suppression if the 95% CI around the average difference in mu between the active period and baseline period did not cross zero, for any of the three electrode sites.*

6.3.8 Summary of results

The introduction of this study outlined that a key condition for mu suppression to be considered a valid indicator of MNS activity would be observing an interaction between condition and electrode site, and that the difference in suppression between hand and kaleidoscope stimuli would be greatest at the central sites. Although significant site by condition interaction effects were seen for the 8-13 Hz band, these effects were not due to significantly stronger central suppression to biological videos – instead, these statistical interactions were due to stronger occipital suppression to kaleidoscope videos and strong central suppression to participants' own movements. For the hand videos, suppression was always stronger at the occipital sites. The only analysis providing evidence of specific central mu suppression to hand videos was that using the within-trial baseline. It would appear that the static-period (within trial) baseline represents a better baselining technique –this was the only baseline in which the planned comparisons found specific suppression for hand-object videos. Furthermore, a higher proportion of individual participants showed mu suppression effects when considering this baseline.

For the beta band, the only main effect of condition was for the beta band for the between-trial baseline (and even for this effect, hand-object and hand-no object videos did not differ from kaleidoscope videos at the central sites, only at the occipital). Similar to the mu-band, this study failed to find evidence of a specific reaction of the beta band to hand videos.

6.4 Discussion

Rest periods are commonly used in mu suppression investigations as baseline conditions. However, using two different rest-baselining methods and examining changes in power at both the central and occipital sites, this study failed to find evidence for specific mu suppression to videos of human movement. The final method, using a baseline measure from a static stimulus at the start of each trial, gave much less variability in measures of mu suppression (as indicated by the narrower confidence intervals around the mean values for this baseline), and did give a pattern of results that was consistent with mirror neuron activity, although as found in previous research, this was much reduced compared to the mu suppression when performing movements (Woodruff & Maaske, 2010).

The control conditions (watching kaleidoscope patterns, and performing finger tapping movements) show that it is possible to dissociate mu from occipital alpha. The positive control condition, in which participants performed movements themselves, confirms that desynchronization of mu at the central sites captures the activity of the motor areas. Furthermore, in this condition, where no visual stimulus is observed, there was no alpha suppression at occipital sites, whereas alpha suppression was substantial when watching visual stimuli. Indeed, if mu suppression was simply a reflection of alpha confounding, and mu suppression was inherently tied to changes in attentional engagement, it would be predicted that the kaleidoscope videos, rated the most engaging by participants, would show

both the strongest occipital *and* central suppression. Instead, while kaleidoscope videos yielded significantly stronger occipital alpha suppression than the biological videos, the difference between the hand and kaleidoscope videos at the central sites was non-significant, or in the opposite direction. This is an encouraging finding, as it suggests engagement and attentional issues are factors that can be separated from mu suppression, and should be considered and controlled in future mu suppression work.

Similar to Tangwiriyasakul et al. (2013) it was found that a significant minority of the participants failed to show the expected suppression effect to hand videos, even in theoptimal within-trial baseline condition. These participants were typical adults with no reported history of any neurological disorders, nor any diagnoses of autism spectrum conditions. These observations highlight that mu suppression to observing human action is not a universal finding, limiting its power as an experimental tool.

In some of the baseline techniques, significant suppression at the central sites to videos of kaleidoscope patterns was observed, stimuli we would not predict to activate the mirror neuron system. The question then arises as to whether participants might somehow have imagined themselves performing the movements they observed. This seems implausible. These stimuli are highly abstract, and were selected as stimuli that could not be easily embodied. Furthermore, participants were asked at the end of the recording if they felt they could perform or imitate the patterns, and almost all of the participants reported that they could not. Limited differences in central mu suppression between hand and kaleidoscope videos call in to question the specificity of mu suppression, and again weaken arguments that this is a valid measure of the MNS.

Broadly, these results are consistent with a recent meta-analysis of mu suppression studies (Fox et al., 2015), which was published during the data collection phase of this

registered report. The current report included more participants than any of the studies included in their meta-analysis, and (unlike many of the investigations reviewed in the meta-analysis) is sufficiently powered. Fox et al. (2015) determined from the studies they reviewed that there is strong, central-specific suppression during action execution, no significant effect of biological (hand) versus non-biological (kaleidoscope) conditions on suppression during action observation, and a lack of central-specific effects during action observation – results similar to our findings for the first two of the baseline conditions. Interestingly however, despite these similarities, we have arrived at different conclusions. Fox et al. (2015) argue that mu suppression can indeed be used to index mirror neuron system activity. I, by contrast, argue that evidence for mu suppression is only apparent when a specific kind of within-trial baseline is adopted that controls for some extent for variability across a session. When other baselines are used – as was the case for many of the studies in the meta-analysis – the impression is that mu suppression is typically confounded with alpha suppression, which occurs in response to the presentation of a new visual stimulus.

Fox et al. (2015) did consider a number of moderating factors in their analysis, including type of baseline used, and found no moderation effects of baseline on the effect size of mu suppression. However, this is not in conflict with our argument that baseline is an important factor. Technically, the strongest mu suppression in the current study was observed when using a short rest baseline, but it is apparent that these results are confounded by alpha. Considering baseline's effect on the *strength* of mu suppression alone will not prove that this is indeed an important factor – I argued that *specificity* is important for determining mu suppression's validity. This study opted to consider the pattern of significant and non-significant mu suppression across our conditions, and only the within-trial baseline showed a pattern of suppression that was specific to biological stimuli.

Given these results, treating mu suppression as a proxy for mirror neuron engagement, and using it as a basis for neurofeedback therapy, requires serious caution. While the original function of the MNS was purported to be action-understanding, theories about the human MNS have evolved radically to encompass potential roles in a number of social and communicative functions, including empathy and language . Indeed, as described in Chapter 5, several investigations have used mu suppression in an individual-differences approach, as a gauge of the quality or responsiveness of an individual's mirroring system. This is then correlated with personal characteristics, such as empathy or prejudice (e.g. Cheng et al., 2008; Gutsell & Inzlicht, 2010). This study cannot speak to whether the MNS is involved in such processes or not, but it is worrisome that such studies may be taken as evidence that mu suppression is a valid and reliable measure of the human MNS, especially as such studies would seem generally quite underpowered to examine these correlational questions, and corrections for multiple comparisons have not always been adhered to. In fact, a study by Silas, Levy, Nielsen, Slade, and Holme (2010) which *did* use appropriate corrections concluded empathy measures were unrelated to individual differences in mu suppression.

This study found that mu suppression is not consistently demonstrated from individual to individual (even in typical participants). One possibility is that individual variation in mu suppression is meaningful and related to some characteristic that this study failed to measure. Nonetheless, this study suggests that mu suppression is not specific to viewing biological stimuli (we argue a key characteristic of the MNS), and thus its use as a measure of the quality of an individual's mirror neuron system seems dubious. Furthermore, I am not aware of any data on reliability of mu suppression– that is, how variable is an individual's mu suppression within and between testing sessions? Correlating mu suppression with individual differences in empathy or prejudice would seem to imply it has some relatively fixed or stable quality to it.

Chapter 7 explores how mu suppression in this sample of participants relates to individual differences in imitation skill, self-reported communication ability and autistic traits. Another individual differnce that this study's data was able to examine at was gender. Although the meta-analysis by Fox et al. noted that studies with predominantly male samples reported stronger effects, previous investigations studying gender differences in mu responses reported that females exhibit stronger mu desynchronisation to observation of biological movement (Cheng et al., 2008; Cheng, Tzeng, & Decety, 2006; Silas, Levy, Nielsen, Slade, & Holmes, 2010). This study's sample had a high proportion of females, so any sex difference in mu suppression could influence our results. Accordingly, a further unregistered analysis was conducted to explore this issue. No gender effects on activity at the central sites during action observation were found (See Appendix H: *Supplementary unregistered analyses for mu suppression registered report*).

Inconsistent findings in relation to gender raise questions about correlations found with behavioural measures of individual differences. As Vul, Harris, Winkielman and Pashler (2009) noted, correlations between behavioural and neurofunctional measures often overestimate effects: "*Such an analysis will inflate observed across-subject correlations and can even produce significant measures out of pure noise*." (Vul et al., 2009, p.279). Their article was concerned with fMRI studies, but it raises warnings about the dangers of studying individual differences using neurofunctional measures of unknown reliability. It is recommended that any researchers investigating correlates of mu suppression should first establish the reliability of their measures.

As well as considering the validity of mu suppression as an index of mirror neuron system activity, a second frequency band was also investigated, the beta band. Researchers have suggested that mu maybe be more related to sensory processing rather than motor activity, and changes in beta power, not mu, are indicative of motor cortex activity (Coll et

al., 2015; Ritter et al., 2009). The recent meta-analysis of mu suppression studies called for further investigation of beta-band responses (Fox et al., 2015). Overall, as predicted, the pattern of results obtained for beta was similar to those obtained with the alpha/mu band. This study's results suggest beta suppression is no better an index of mirror neuron activity than mu. However, other investigations have used post-movement beta rebound effects (rather than suppression during stimulus presentation) to examine beta's responses. Following medial-nerve stimulation, when beta typically "rebounds" to higher than pre-stimulation levels, showing participants videos of actions has been found to suppress this rebound effect (Muthukumaraswamy & Johnson, 2004). Reduced rebound suppression has also been noted in participants with autism (Honaga et al., 2010). Further work will need to be done to ascertain whether post-movement rebound effects offer a better measure of MNS engagement than simple suppression during stimulus presentation.

6.4.1 Controlling for attention and alpha effects

In their recent meta-analysis, Fox et al. (2015) discussed the problems of attentional effects and alpha on mu suppression investigations. They argued that mu suppression studies should include a condition in which no action is observed or executed, but in which participants experience the same attentional demands as the other experimental conditions. This attention condition could then be subtracted from experimental conditions to control for attentional confounds. Although this recommendation is well-justified, in practice it is hard to implement because we do not have a way of matching attentional demands across tasks. Some previous investigations have used continuous performance tasks to ensure participants maintained attention to the screen, but this may be problematic when long resting baselines are used (when there are naturally no stimuli for participants to continuously count or monitor).

The current investigation used an attention check that was less demanding than a continuous performance task, which was selected to motivate participants to attend to the screen, and to provide a broad filter for those who failed to do so. It is noteworthy that a large number of participants (27 of 109) failed this attention check task and had to be replaced. The final sample included only participants who passed this test, but the high rate of attrition does suggest that attentional engagement does need to be considered and sufficiently monitored or controlled for in mu suppression studies. Future work will need to strike a balance between demanding attention tasks (which could increase alpha suppression and lead to confounding), and ensuring that participants are paying sufficient attention to the stimuli they are observing.

It was predicted that the long rest period would inflate alpha levels in the baseline, and thus inflate apparent mu suppression. However, the results do not support this – significant mu suppression was not seen for any of the video conditions using this baseline. Indeed, stronger occipital alpha and central mu power was seen in between-trial baseline, which used short rest periods. This result is unexpected, as the stimulus the participants are seeing in the short and long rest periods is exactly the same. What this suggests is that the time-course of alpha and mu responses is also important. Sampling alpha/mu levels in the middle of the long rest condition is not the same as sampling them at the beginning or end of this period, and similarly sampling these levels when participants are viewing a blank screen but when they have just been viewing dynamic videos is not the same. It may be that over the length of the long resting baseline, alpha levels change, or that going from viewing a video to a blank screen may induce greater alpha enhancement than sitting without stimulation for a long period of time.

Kaleidoscope videos produced significantly more suppression in the alpha and beta bands at the occipital regions, regardless of baseline. These stimuli were also rated by participants as the most engaging. These stimuli did differ from the hand videos in a number

of ways, and potentially very slight differences in overall level of motion, or contrast, could have had an impact on the differences in alpha suppression between these video types. However, this finding does not explain why this study failed to find an effect of video condition on central mu suppression, in two of the three baseline techniques.

Finally, one point to note is the suggestion made by Fox et al. (2015) that the tight association between alpha and mu might be a reflection of "a close coordination of action and attention". (Fox et al., 2015, p6). While Fox et al. (2015) themselves do not elaborate much on this point, what their idea entails is that mu and alpha reflect separable but highly related processes, and to an extent seeing changes in alpha should not alarm us, as they may be an inherent part of action processes, alongside motor activity. It is an interesting notion. For this study however, if one accepts that during action attentional processes are highly probable and perhaps and natural part of action processes, why are the alpha and mu responses to participants' *own* movements so clearly distinct? The current data would have to suggest this close coordination only occurs for observing others' actions.

6.4.2 Object effects in mu suppression

Greater mu suppression to videos in which participants interact with an object has been found previously (Muthukumaraswamy & Johnson, 2004). I did not replicate this finding and the recent meta-analysis by Fox et al. (2015) also failed to find a significant moderator effect of object versus non-object-directed stimuli. Potentially, one reason previous reports may have found stronger mu suppression to transitive versus intransitive actions may have been more related to the presence of goals or discernible actions, as opposed to the mere presence of an object. In Muthukumaraswamy and Johnson's (2004) investigation, their stimuli involved precision grips made on an object, and precision-gripping movements made without contacting an object. A precision grip may be described a goal-

oriented action, whereas in our stimuli the videos in which the hand interacts with a pencil are less clearly goal-based actions.

Another possible explanation for why in both the current study and in the recent meta-analysis mu suppression object effects have not replicated could be that mu suppression is more related to sensory rather than motor stimuli, and that the tactility of stimuli affect the strength of mu suppression observed. Using a cross-modal repetition suppression design, a recent paper by Coll et al. (Coll, Bird, Catmur, & Press, 2015) showed that repetition effects were only found in conditions where the tactile components of the stimuli were repeated, not when the motor components were repeated. They thus concluded that mu suppression is more related to sensory rather than motor mirroring. Arguably, the stimuli in the current study differed from Coll et al.'s in that there was quite minimal contact between the hand and object in the hand-object videos. Potentially this could mean that there was not enough of a tactile element to the videos to cause sufficient suppression and obtain a significant object effect.

6.4.3 Suggestions for going forward

I would not want to suggest that the procedure used in the current study is "bullet-proof" or a "gold-standard" way of doing mu suppression studies. Instead, I hope that this study will serve as a platform for discussion around how best to conduct investigations going forward, so that researchers can converge upon a reliable setup that is most likely to provide solid ground for robust breakthroughs in understanding. Mu suppression studies are already widely used and cited. What do mu suppression studies in the future need to consider? Several important suggestions were outlined in the recent meta-analysis by Fox et al. (2015). These include the need for execution *and* observation conditions to be included in future

studies, and ways to deal with potential alpha confounding. Here, I re-iterate salient points and add further novel suggestions.

These findings highlight the importance of considering and presenting the results from regions associated with alpha, outside of the central sensorimotor strip. As described earlier, a recent paper utilised a whole-brain approach to analysis to re-examine the issue of mu suppression deficits in autism (Dumas et al., 2014). When only examining the central electrodes, the previous reports of mu suppression abnormalities in autism were replicated. However when their analysis widened to include other regions it was clear that the key sites of difference between control and autistic participants were not at the central sites, but rather in the frontal and occipital regions. Indeed, Fox et al. (2015) found in their meta-analysis that for action-observation conditions, effects were not specific to the central regions, and they noted that many studies failed to report findings from other sites. To be confident that mu suppression is indexing changes in activity in motor areas, it must be ruled out that these changes could be coming from elsewhere.

As discussed in Chapter 5, another concern when reading mu suppression literature is that there seems to be much room for analytic flexibility, a factor known to be associated with poorer reproducibility (Ioannidis, 2005). For example, the parameters of the mu band are not fixed, and different studies use different definitions of the mu band, with some suggesting the mu-band needs to be further divided up (Pfurtscheller et al., 2000). Having agreement, discussion and transparency around how data collected from mu suppression studies is analysed will be important. The analyses used in the current study were based on what seemed to be the prevailing approach in the field, and the analysis scripts have been made open to the scientific community to download, use and adapt.

Finally, considering data at the individual level will be useful in ascertaining to what extent mu suppression to action observation is a reliable phenomenon, dependable enough for experimental or proposed clinical use. In common with some previous experiments, it was noted that mu suppression is not observed in a significant minority of typical participants. If mu suppression is to be continued to be used for inferring the processes mirroring systems are involved in, or comparing groups (such as autistic and typical participants), understanding why so many participants do not show expected mu suppression effects will be important. Studies of mu suppression in autism usually present data at the group level, comparing average changes in mu, but it would be intriguing to know whether the proportion of participants showing expected mu suppression effects differ between the groups - do more participants with autism show no mu suppression, or an indeed an increase in mu when observing actions?

6.4.4. Conclusions

This study is one of if not *the* largest mu suppression experiment to date, investigating mu suppression's validity as a measure of human mirror neuron system activity, and the importance of baseline methodology. The results suggest that mu suppression calculated using resting baselines is not specific to biological stimuli, nor the central motor regions. Similar results were found for beta-band suppression. Using a baseline of a static image improved the specificity of mu responses, but even when this baseline technique was used, a significant minority of typical participants did not show the expected mu suppression effects. This has implications for the future use of mu suppression in experimental settings and for clinical applications.

CHAPTER SEVEN: Mu suppression and individual differences in autism traits, imitation abilities and self-reported communication skills

7.1 Introduction

As explored in Chapter 5, relationships between a number of domains and the mirror neuron system have been proposed, including action understanding, imitation, language and communication, social processes, and autism. These accounts would predict behavioural skills proposed to be supported by the MNS and neurophysiological measures of MNS activity would show associations. As discussed in Chapters 5 and 6, several researchers have used mu suppression in an individual differences approach as a measure of a given person's mirror neuron system's quality or responsiveness. If mu suppression can be considered a valid means to assess mirror neuron functioning, then it is reasonable to predict that mu suppression and these domains should show correlations. Chapter 5 discussed studies examining language (namely speech perception and language comprehension), social processes and autism, using mu suppression as means to explore these domains in relation to the MNS. Here, I consider the evidence for a relationship between mu suppression and imitation, autistic traits and communication skills.

7.1.1 Does imitation mediate findings of abnormal mu suppression in autism?

Mu suppression has been suggested to be related to individual differences in imitation skill, and some have even suggested that the link between autism spectrum conditions and abnormalities in mu responsivity may be mediated by imitation (Bernier, Aaronson, & McPartland, 2013; Bernier, Dawson, Webb, & Murias, 2007). Two studies have found mu suppression and imitation to correlate. Bernier et al. (2007) compared 14 adults with ASD and 15 typical controls on their changes in mu power during imitation, observation and execution tasks, and also assessed their participants' imitation skills. In keeping with previous

investigations, they found their autistic group showed impaired imitation skills (on all types of imitation tested), and reduced attenuation of the mu wave when observing others' actions.[8] Bernier et al. (2007) noted that behaviourally assessed imitation skills were correlated with mu attenuation during the observation and imitation EEG conditions. This correlation was particularly strong for facial imitation. The authors suggest this may be because this kind of imitation has an even greater reliance on the observation/execution matching system, the activation of which is captured by mu suppression.

In a later study with autistic and typically developing children (with 19 participants in each group), Bernier et al. (2013) did not find significant differences between groups in terms of their mu responses to observing others, but again noted an association between mu responsivity and behaviourally assessed imitation abilities. Correlational analyses found significant associations between mu suppression in the observation condition and facial imitation ability (this study did not have an imitation EEG condition). Correlations were not found between mu suppression and hand imitation, nor autism related communication impairments (drawn from the ADI-R communication subscale score – previously an association between mu suppression and this subscale has been shown by Fan, Decety, Yang, Liu, and Cheng, 2010). Considering the mu responses of the children individually, five typical children and two autistic children did not show mu suppression during the observation task. Comparing this group of seven children to the rest of the sample on their imitation performance, they were significantly poorer than the other children on their facial imitation abilities. The authors speculate that perhaps mu suppression may not be related to autism *per*

[8] However, no main effects of group or interaction effects actually emerged from the analysis. This conclusion was interestingly driven in part by a non-significant difference in mu suppression between execution and observation conditions in their *typical controls*, while a significant difference was observed in ASD participants. Arguably, the non-significant effect in the typical group is what is unusual, rather than the effect in the ASD group, as execution of action has been shown to lead to greater mu - see Woodruff & Maaske, 2010.

se, but rather to imitation abilities – which often are poor in autism, but can also vary greatly. This could perhaps explain the heterogeneous findings of previous mu suppression studies.

These two studies hint at an association between imitation and mu suppression, but two other studies have argued against it. Ruysschaert, Warreyn, Wiersema, Oostra, and Roeyers (2014) also examined the relationship between mu suppression, imitation ability and autism symptomatology in sample of young children (with a mean age of around 40 months), including 18 with ASD and 19 typical controls. They found no correlations between mu suppression and performance on their imitation tasks, and autism symptomatology (as measured by the SCQ) correlated with mu suppression in the typical group only (and even then, the correlational was only marginal – $p = .08$). While Fan et al. (2010) did not examine the actual correlations between imitation and mu suppression, they point out that while their autism group showed poor imitation performance, these participants' mu suppression was intact, suggesting a dissociation between mu responsivity and imitation skills.

So far, the evidence for a relationship between imitation and mu suppression is equivocal. However, these studies lack the power to properly ascertain whether or not correlations between behavioural characteristics and EEG phenomena are present - correlational questions such as these require larger samples, but even the largest correlational study conducted into imitation and mu suppression had 38 participants. Exploration of these relationships with larger sample sizes is needed.

7.1.2 Do imitation skills pattern with autistic traits?

Considering the hypothesis put forward by Bernier et al. (2013), that imitation may be more strongly related to mu than autism diagnosis, it would be useful to know to what extent autistic traits and imitation skills are related to one another. Intriguingly, while imitation is considered to be a predictor of language and social development, and poor imitation is

considered a key deficit in autism, no studies have examined the relationship between imitation skill and autistic traits in the general population. Examining such associations in typically developed populations may seem unwarranted, but autism has been argued not to represent a qualitatively distinct group of individuals, but rather an extreme end of a distribution of traits (e.g. Constantino & Todd, 2003). Exploring the patterning of skills considered poor in ASD with autistic traits in the general population could help to inform autism research about what deficits in ASD can be thought of as being at the extreme end of a distribution, rather than qualitatively disordered.

Indeed, as autism is considered to be a highly heritable condition, family members, including the siblings and parents of autistic individuals, are assumed to carry some of the genes associated with autism. Studying their performance on a variety of tasks has been used to investigate the broader autism phenotype – that is, what sub-clinical characteristics are manifested in relatives who do not have clinical autism. Understanding this can help inform researchers what difficulties in autism are genetically transmitted. Considering the importance that has been given to imitation difficulties in both old and new theories about autism (Rogers & Pennington, 1991; Williams, Whiten, Suddendorf, & Perrett, 2001), it is interesting to ask - are imitation deficits part of the broader autism phenotype? In fact, little evidence exists to answer this question, though two studies have investigated imitation performance in "at-risk" younger siblings (children whose older siblings have autism, and who are therefore at a higher risk of being diagnosed with the condition themselves). The findings of these studies conflict; imitation difficulties were found in one sample of at-risk siblings, but not in the other (Toth, Dawson, Meltzoff, Greenson, & Fein, 2007; Young et al., 2011).

7.1.3 Do EEG measures of the MNS correlate with autistic traits?

As well as studies investigating differences in mu suppression between clinically autistic and non-autistic groups, other studies have suggested that within the general population MNS responsivity measured with EEG is associated with level of autistic traits in non-clinical samples. Several papers by one group of researchers have suggested links between scores on the Autism Spectrum Quotient (AQ) – a questionnaire considered to capture an individual's level of autistic traits– and beta responsivity (Cooper, Simpson, Till, Simmons, & Puzzo, 2013; Puzzo, Cooper, Cantarella, Fitzgerald, & Russo, 2013; Puzzo, Cooper, Vetter, & Russo, 2010). As discussed in Chapter 6, changes in the beta band have also been suggested to capture the activation of the MNS (although the study in Chapter 6 did not find beta to fit the pattern of reactivity expected of the MNS). While perhaps in keeping with the theories positing that the MNS has a role in ASD, these studies are limited by not correcting in their analyses for multiple comparisons, inflating Type 1 error.

Puzzo et al. (2010) examined the EEG responses of groups of typical participants who scored high or low on the AQ to observations of hand actions and static hand images. This study looked at both mu band and beta band activity, at sites over both the motor and premotor areas. They found a group difference for beta, at premotor sites only, and this effect was driven by enhanced ERD to *static* hands in the high AQ group. Thus, while there was a difference between static and moving hands for the low AQ group, this difference was not significant in the high AQ group, but not because the sensorimotor cortex was not active during the action observation condition, but because of activity during the static hand period.

In another study, Puzzo et al. (2013) considered the effects of repetitive TMS to the inferior parietal lobule (IPL – an area considered part of the human MNS) on beta and mu responsivity to similar stimuli. They included participants with average or high AQ scores. The results are complex, and are largely concerned with beta rather than alpha. In the sham TMS condition, they replicated their earlier finding that participants with high AQ scores

show limited differences in beta ERD at premotor areas, when comparing the conditions in which participants view static hands and moving hands. As shown previously, participants with average AQ scores did show a difference in beta ERD in these conditions. They predicted that stimulation to the IPL would enhance alpha and beta ERD during action observation (the moving hand condition). This was not found – there was no difference between active and sham TMS condition on either frequency band during action observation. However, TMS did increase beta ERD during the *static* hand period, in the *average AQ group*. Thus, in the average AQ group, stimulation to the IPL appears to increase the reactivity of the sensorimotor cortex to static stimuli, but not to action observation. For the high AQ group, there were no differences between static and moving, in either sham or active TMS. Taking these findings all together, the authors suggest that perhaps as the MNS is already maximally active in the average AQ group, stimulation does not activate it any further. However, MNS activity is able to be increased in the static period. In the high AQ group, they suggest the system may be *hyperactive* during static hand observation, leading to no differences between static and moving conditions, and limited effects of TMS.

One concern with this paper is the large number of factors in the analysis, without any apparent correction applied to account for the increased likelihood of finding a significant effect due to multiple comparisons. The factors in the analysis included AQ group (high versus low), TMS condition (sham versus active stimulation), stimulus condition (static hand versus video) and electrode site. This analysis was run twice for two different regions (frontal and central), and the dependent variables included multiple frequency bands. For each frequency band and region, the analysis is testing for four main effects, six two-way interactions, three three-way interactions and one four-way interaction. The likelihood of finding a significant main or interaction effect – for each analysis and dependent variable – is over 50%.

Similar limitations apply to a third paper by this group. Cooper et al. (2013) also investigated EEG responses to actions performed by happy, neutral or angry actors in a sample of participants with high or low AQ scores. They suggest that the beta responses to these facial expressions show a different pattern in the two AQ groups. However, the high number of factors in their analysis (group, expression, site, and frequency band) mean that significant effects are hard to interpret, given the problem of multiple comparisons (no mention of corrections for multiple comparisons is made).

7.1.4 Aims of this project

This project aimed to test the relationships between the proposed EEG measure of the human mirror neuron system – mu suppression – and behavioural characteristics related to autism. Specifically, one aim was to examine if the correlation between imitation skill and mu suppression during action observation found in previous studies could be replicated. In addition, the relationship between mu responsivity and level of autism traits and communication competence was considered. Given previous researchers have suggested that mu suppression in other communication disorders should be examined, and given theories suggesting that the MNS underpins language and communication, a self-report measure of communication ability (the Communication Checklist Self Report – CC-SR) was also included.

Finally, aside from relationships with mu, there are theoretical reasons to predict that these behavioural measures could be related in their own right. As there is currently limited evidence about imitation ability in the broader autism phenotype, it seemed also interesting to examine how level of autism traits relates to imitation ability. Also, the CC-SR and AQ could arguably be said to assessing highly related constructs – communication competence (including pragmatic skills and social interest) and autism traits. However, no previous

reports seem to have examined whether these two measures correlate. The relationships between participants' imitation abilities, autism traits and self-reported communication skills were thus also explored.

7.2 Method

In addition to the electrophysiological data collected during the study (described in Chapter 6), participants were also asked to perform behavioural tasks of imitation, and fill-in two self-report questionnaires of autism traits and communication skill. The EEG data was collected and analysed as described in Chapter 6. In the present study, only suppression to hand videos was examined. As described in Section 6.2.3, the hand video stimuli included transitive and intransitive videos of hand actions, which were presented in mixed blocks throughout the EEG session. Behavioural imitation skills were assessed using the Mature Imitation Task (MIT), autistic traits using the Autism Spectrum Quotient (AQ), and communication skill using the Communication Checklist – Self Report (CC-SR), all described below. All the behavioural tasks were administered before EEG set-up and recording, starting with the CC-SR, followed by the AQ, and then the MIT.

There were no language criteria for the mu suppression registered report - because of this, not all of the participants were native English speakers. This is problematic for the CC-SR, which was developed and normed on native English speakers. Scores are therefore open to bias where individuals who have good communicative competence in their own native languages score lower when completing the questionnaire about themselves while in the UK, due to the increased demands of communicating in a second language. Participants who spoke English as a second language were therefore asked to complete the questionnaire reflecting on how they are when communicating in their mother tongue. This is still imperfect, as the items in the CC-SR have been developed and selected for English-speakers,

but it was felt it would provide a better picture of individuals' communicative competence where they were non-native English speakers.

7.2.1 Mature Imitation Task

The Mature Imitation Task is an unpublished imitation battery (Rogers, Cook, & Greiss-Hess, 2005), and has been used previously in investigations of mu suppression and imitation (Bernier et al., 2013, 2007). The MIT includes a number of different imitation tasks including a) single facial postures b) sequences of three face postures c) single meaningless uni-manual hand postures d) sequences of three meaningless uni-manual hand postures e) single hand postures requiring both hands f) meaningless body movements and g) actions on objects (administered always in this order, using the same trial order in each subtest). Participants were shown all the actions with pre-recorded stimuli, with a male model, to ensure the consistency of presented actions. Participants were instructed to copy the model as closely as they could, and their responses were filmed and coded offline according the unpublished manual for the MIT.

The MIT does not currently have published norms. In the current study, percentage scores were calculated for hand, face, body movement and actions with objects (shown in Table 7.1). Bernier et al. (2007) summed their scores, rather than calculating percentages. However, total imitation score, face summary score and hand summary scores were calculated to compare the imitation performance of this sample with that of Bernier et al (2007). The imitation performance of the typical adult group in Bernier et al. (2007) (N=15) is comparable to that of the sample in this larger study. In Bernier et al. (2007), total imitation, face summary and hand summary scores had averages (and standard deviations) of 185.2 (13.1), 57.7 (4.3) and 45.9 (7.8), respectively. In the current study, the average scores were: 180.36 (12.33), 55.02 (7.43), and 43.42 (8.62). The performance of the current sample

of typical adults therefore seems quite comparable to that of Bernier et al. (2007), and suggests good fidelity of the MIT.

Table 7. 1 Mean imitation performance of sample

Imitation measure	*% score*
Actions with objects	74.94 (7.37)
Body movements	85.96 (10.06)
Facial imitation (*includes single and sequenced facial tasks*)	85.97 (11.60)
Hand imitation (*includes single and sequenced hand tasks, and complex hand tasks*)	67.84 (13.47)
Total imitation	79.45 (5.43)

Table 7.1: *Means and standard deviations for the imitation performance in the MIT.*

7.2.2 Communication Checklist – Self Report

The CC-SR is a questionnaire measure for adults to rate their own communication skills (Bishop, Whitehouse, & Sharp, 2009). It contains 70 items, and the answers load onto three sub-scales; structural language, pragmatic language, and social engagement. The structural language sub-scale assesses a participant's mastery of the linguistic aspects of language, including speech, syntax and semantics. Items for this scale include statements such as, 'I mix up similar words' and 'I leave out sounds in words'. The other two subscales are concerned with social communication. The pragmatic language skills subscale assesses participants' level of appropriate communicative behaviours (for example, 'I am told I keep going on about things no one else is interested in' and 'I copy what other people say'). The social engagement scale assesses poor social communication relating to a passive

communicative style (for example, ‘I feel anxious when I am with other people’ and ‘People say that my face looks blank’). For each item, participants indicate the frequency with which the statement applies to them, selecting between: less than once a week (or never), about once a week, once or twice a day, or several times a day (or all the time). Lower scaled scores (though higher raw scores) indicate an increased level of communication difficulties. The CC-SR has been standardised and normed. Standard scores for each subscale have a mean of 10 and an SD of 3. The current sample of 50 participants, with means of approximately 11 and standard deviations of approximately 3 for each of the subscales, thus showed quite a typical range of scores (see Table 7.2 for exact means and standard deviations).

7.2.3 Autism Spectrum Quotient

The AQ is a widely used self-report questionnaire measure considered to capture the level of autism traits shown by an individual (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Ruzich et al., 2015). It was developed based on the idea that autism lies on a continuum of traits, and thus autistic traits can be measured with a quantitative approach. The AQ comprises of a series of statements (e.g. ‘I would prefer to go to a library than a party’), for which participants are asked to mark whether they agree or disagree with the statements describing themselves. It contains 50 items, assessing social skill, attention switching, attention to detail, communication, and imagination. A score of 32 or higher is usually considered to suggest a level of autism traits usually found in clinically diagnosed cases (Baron-Cohen et al., 2001). In a recent systematic review of the AQ, the mean score, drawn from a pooled sample of 4931 typical participants was 17 (15, when studies explicitly excluded participants who are part of the broader autism phenotype), and the mean for individuals with ASD was 35 (based on a pooled sample of 1374). The mean AQ score of the current sample thus appears to be quite typical for a non-clinical sample (M = 14.8). Three

participants in the current sample did score 32 or higher on the AQ, indicating a high level of autism traits.

7.2.4 Participants

The final sample of participants for the registered report examining mu suppression was 61. To be included in this sample, participants had to pass an attention check (they had to report seeing three grey stars and three grey arrows during the EEG recording – see Section 6.2.3 *Procedure*), and yield enough trials for the EEG analysis. Of these 61, 50 completed the MIT imitation assessment – the other 11 participants did not undertake the imitation tasks due to time restraints. For parity, these 11 participants were not considered for the correlations between the questionnaires and the EEG measures, so all correlations are based on the same set of 50 participants. This sample included 32 females and 18 males.

Table 7.2 Participant characteristics

Variables		*Mean/Frequency*
Gender		32 F, 18 M.
Age (years)		22 (3.53)
Non-native English speakers		14
AQ		14.8 (8.29)
CC-SR Standardised Scores	Structural-language	11.4 (3.49)
	Pragmatic-language	11.1 (3.07)
	Social interest	11.2 (3.04)

Table 7.2: *Participant characteristics (N=50). Frequencies for participants' gender and language-status are shown. For variables age and AQ, the mean is shown with the SD.*

7.2.5 Correlational analysis plan

The actions observed by participants during the EEG recording included hand actions with objects (HO), and hand actions without objects (HNO). Previous investigations have examined the correlations between imitation and mu suppression to transitive stimuli, and none have considered the potential for differential relationships between imitation abilities and mu suppression to transitive versus intransitive stimuli. However, by examining the behavioural correlates of two separate mu suppression measures to an already moderately large number of behavioural measures, it was clear this would inflate Type 1 error, or require such a stringent statistical significance threshold that it was highly unlikely that any associations would survive corrections, thereby inflating Type 2 error. It was therefore decided to average the mu suppression of HO and HNO together, and to consider the relationship between just this measure and behavioural measures. As there were directional predictions about the correlations that could be derived from both previous empirical investigations and theories, the correlations were assessed with one-tailed tests. Specifically, reduced mu suppression (i.e. suppression values closer to or above zero) was predicted to be associated with increased autism traits, poorer self-reported communication skills, and poorer imitation skills. (Note that because higher AQ scores relate to higher AQ traits, and higher CC-SR scores relate to better communication skills, we predicted positive correlations between mu suppression and AQ, and negative correlations between mu suppression and the CC-SR.)

As described in Chapter 6, there were several methods of baselining mu suppression measures explored in the current dataset. Given the results of the registered report, the suppression scores used in these analyses represented those baselined with the static image, as this baseline appeared to yield the best results in terms of specificity to biological motion. Beta-band activity (15-35Hz) was not examined, and neither was suppression at the occipital sites. Thus, this study considered individual differences in the average suppression to hand

videos (including both HO and HNO videos) at the central electrodes, in the 8-13 Hz range, using the within-trial baseline described in Chapter 6.

7.3 Results

7.3.1 Correlations between mu suppression, autism traits and communication skills

Several of the imitation subscales (including actions with objects, body movements, face movements) and the questionnaire measures (AQ and the social-engagement subscale of the CC-SR) were not normally distributed according to the Shapiro-Wilk's test. Because of this, Spearman's Rank was used to calculate the correlations between the measures. One-tailed correlations were calculated for the association between averaged mu suppression to hand videos and scores on the AQ, CC-SR and imitation tasks. There were nine comparisons being made at this stage, and thus correcting for multiple comparisons would mean the adjusted significance threshold would be 0.006.

Table 7.4 shows the correlations between the measures. Only one association was found that would survive this threshold and it was with imitation total score r_s=.384, p=.003. Figure 7.1 shows the scatter plot for this correlation. Mu suppression also correlated with hand imitation (r_s = .265, p=.032) but this would not survive correction for multiple comparisons. No other associations were significant, even to the uncorrected p = .05 level.

7.3.2 Correlations between behavioural measures

Correlations between imitation scores, self-reported autistic traits and communication abilities were also considered. Considering all relationships between the imitation measures and questionnaire measures meant there were 20 comparisons being made, thus a stringent correction for multiple comparisons would be p =.0025. In addition, correlations between the

AQ and the CC-SR subscales were also examined. Adding these further three comparisons would mean the corrected threshold would be p =.0022.

One-tailed tests were again used, as there were directional predictions for the relationships between imitation skills and questionnaire measures (it was predicted that participants with poorer imitation would show higher levels of autistic traits and communication difficulties).

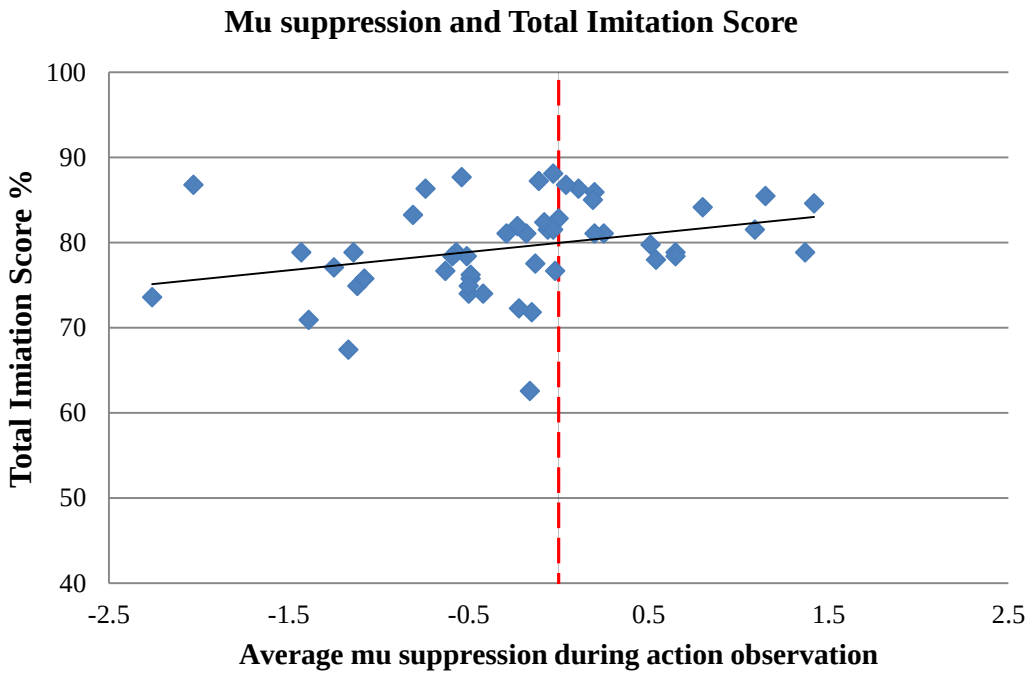


Figure 7.1: Scatter plot showing the correlation between total imitation score and mu suppression during the hand videos. *The red dashed line emphasises mu suppression values of zero, which would indicate no differences between mu power during the baseline period and during the movement period of the hand videos.*

Total imitation score correlated with AQ score (r_s = -.236, p =.049) and the structural-language subscale of the CC-SR (r_s =.299, p =.018), but neither of these associations would survive corrections for multiple comparisons. Of all the imitation sub-scales, only the imitation of hand actions was found to show any significant correlation with the

questionnaire measures; with AQ, r_s = -.341, p=.008; with the social engagement subscale of the CC-SR, r_s = .257, p=.036; with the structural language subscale of the CC-SR, r_s = .375, p =.004. Again, these associations would not survive stringent correction for multiple comparisons.

Scores on the AQ correlated moderately to strongly with all three CC-SR subscales. With the pragmatic skills subscale, r_s = -.401, p =.002. With the social engagement subscale, r_s =-.656, p <.001. For the structural language subscale, r_s =-.377; p =.004, but this correlation would not survive correction for multiple comparisons. Figures 7.2 and 7.3 show the correlations between AQ and the pragmatic language and social engagement subscales.

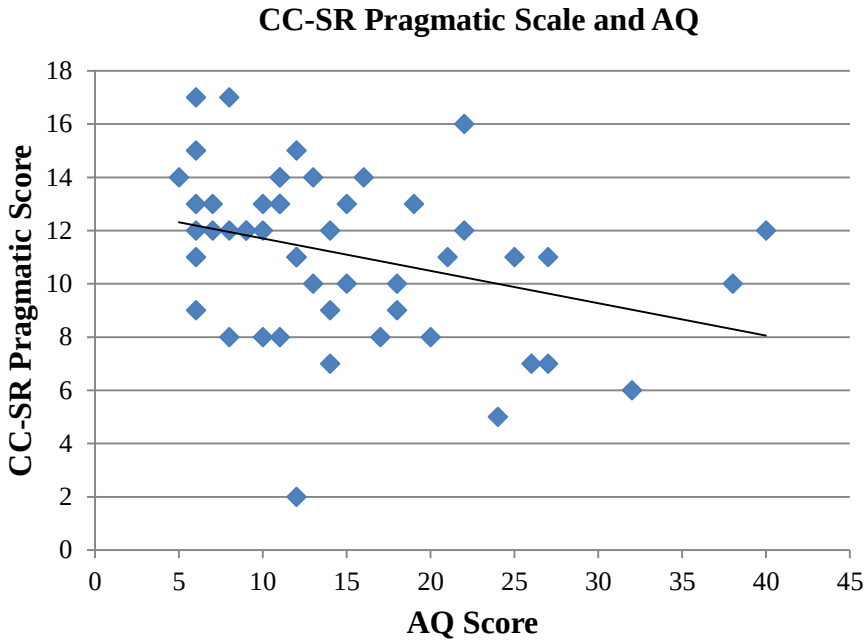


Figure 7.2: Scatterplot showing the correlation between the CC-SR pragmatic subscale score and AQ score.

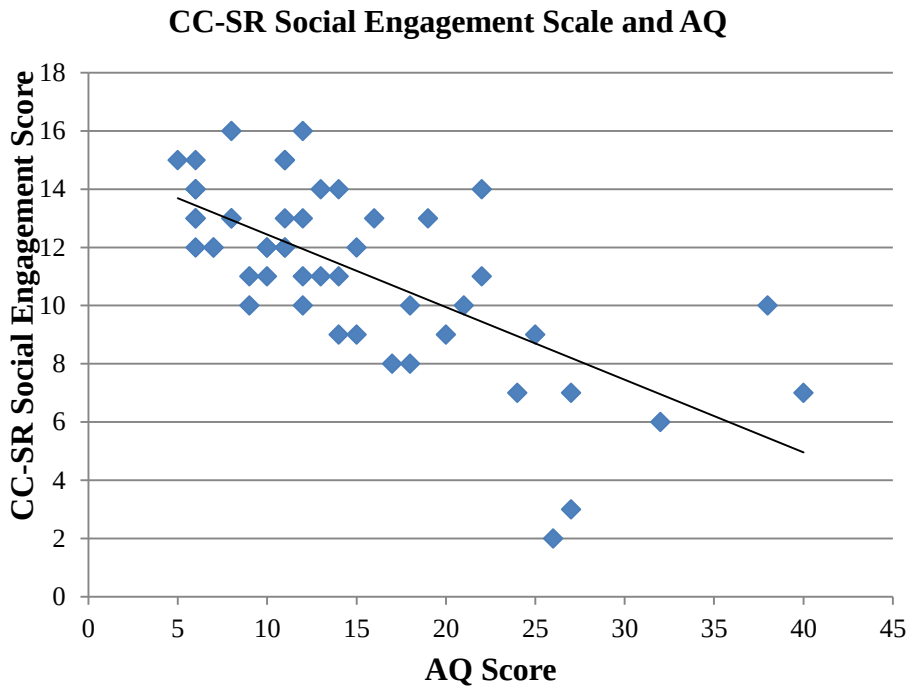


Figure 7.3: Scatterplot showing the correlation between the CC-SR social engagement subscale score and AQ score.

Table 7.3 Correlation Matrix

	Mu suppression	*AQ*	*CC-SR LS*	*CC-SR PS*	*CC-SR SE*	*Imitation Total*	*Imitation: Actions with Objects*	*Imitation: Body Movements*	*Imitation: Face actions*	*Imitation: Hand actions*
Mu suppression		-.012	.081	.122	.021	.384**	.206	.015	.117	.265*
AQ	-.012		-.377**	-.401***	-.656****	-.236*	-.042	-.183	-.044	-.341*
CC-SR LS	.081	-.377**		.692***	.620***	.299*	.087	.003	.076	.375**
CC-SR PS	.122	-.401***	.692***		.712***	.202	.170	.050	.015	.201
CC-SR SE	.021	-.656***	.620***	.712***		.103	.079	-.054	.100	.257*
Imitation Total	.384**	-.236*	.299*	.202	.103		.320*	.206	.378**	.795***
Imitation: Actions with Objects	.206	-.042	.087	.170	.079	.320*		.128	-.080	.196
Imitation: Body Movements	.015	-.183	.003	.050	-.054	.206	.128		-.185	.090
Imitation: Face actions	.117	-.044	.076	.015	.100	.378**	-.080	-.185		.105
Imitation: Hand actions	.265*	-.341*	.375**	.201	.257*	.795***	.196	.090	.105	

Table 7.3: *Correlation matrix, of one-tailed correlations between mu suppression, and behavioural measures. Significance values are indicated: *<.05, ** <.006 (the corrected threshold for 9 comparisons), *** <.0022 (the corrected threshold for 20 comparisons), ****<.001.*

7.4 Discussion

While unquestionably still underpowered for correlational analyses with this number of factors, this project is one of, if not *the*, largest correlational study performed with mu suppression, and certainly the largest yet that has examined relationships with imitation ability. Mu suppression during the observation of hand action videos was found to correlate significantly (even after correcting for multiple comparisons) with total imitation score, but *not* in the direction predicted by previous accounts. Weaker mu suppression was found to correlate with better imitation performance. Neither hand nor face imitation abilities were found to correlate significantly with mu suppression after correction for multiple comparisons. Our results do not support the account that reduced mu suppression is related to poorer imitation abilities, or that heterogeneity in imitation skills could account for the disagreement in previous mu suppression findings in autistic samples.

While the direction of this correlation is contrary to what would be predicted by contemporary mirror neuron accounts, two potential explanations are proposed to account for this significant relationship. The first is that increased mu suppression in participants with poorer imitation skills could reflect the fact these participants have mirror neuron systems that are inefficient and exhibit more "effort" during processing. Thus, activity was increased in some participants, but not related to improved performance. This account would manage to preserve the theoretical link between imitation ability and the MNS, but seems highly unlikely given that this direction of association was not described by previous reports. An alternative explanation could be that success on the MIT was more reliant on processes other than perception-action matching. Indeed, for some of the harder subtests of the MIT such as the hand or face sequences tasks, strategies such as verbal labelling may have been more likely to

increase imitation accuracy, rather than relying on perception-action matching alone. If some participants were opting to use strategies such as this, then the EEG observation condition and behavioural imitation task would differ in their reliance on perception-action matching processes (processes assumed to be reflected by mu suppression). This would diminish the association between mu suppression and imitation scores, and may even reverse the association if perception-action matching proved a weaker strategy than labelling for remembering and performing the modelled actions.

As well the relationship between imitation and mu suppression, this study examined the associations between these factors and autism traits and communication skills. The suggestion that imitation, communication and autism are related through a shared connection with the MNS has been raised many times. Oberman and Ramachandran (2007) suggested that a broken MNS could underlie the social, imitation and communicative deficits seen in ASD. The MNS has been argued to have been key in the evolution of our ability to communicate with others (and imitation and mimesis is proposed to have played a role in this development), and indeed authors have advocated the use of mu suppression to investigate MNS dysfunction in children with communication difficulties (Arbib, 2012; Le Bel, Pineda, & Sharma, 2009; Rizzolatti & Arbib, 1998). If mu suppression is assumed to be an adequate measure of MNS function (and the results of Chapter 6 call this assumption into question), then our findings provide quite weak support for the idea that these domains are linked, via a common dependence upon the MNS. Autistic traits and communication skills did not correlate strongly with mu suppression, nor with imitation ability. This latter finding is problematic for the MNS account, even if one accepts that the lack of correlations between mu suppression and these behaviours is due to mu suppression representing a poor measure of MNS activation.

We also examined the agreement between two commonly used questionnaire measures in the study of communication and autism traits – the CC-SR and the AQ. Two of the subscales of the CC-SR correlated with AQ score, the social engagement and pragmatic skills subscales – the correlation with the structural language subscale would not survive stringent corrections for multiple comparisons. One hypothesis that could explain this finding is that language and communication difficulties form part of the broader autism phenotype. Initially, such an account may appear to conflict with the work of Bishop et al. (2004), who found limited evidence of language impairment in relatives of individuals with ASD. In actual fact, the findings here are complementary – as discussed by Bishop et al. (2004), findings of language deficits in relatives of individuals with autism appear most likely to arise when investigations use self-report measures of language and literacy problems. Using the AQ in their own study, Bishop et al. (2004) split their sample of parents into those exhibiting high and normal scores of autistic traits. Those with higher AQ scores were more likely to report early language and literacy problems, despite not showing any group differences to the other groups of parents on direct tests of verbal IQ, nonword repetition or reading. Thus, these parents *report* language problems, but do not demonstrate them in objective testing. Bishop et al. (2004) drew parallels between their findings and those of an earlier study (Fombonne, Bolton, Prior, Jordan, & Rutter, 1997), which similarly found parents of autistic children reported language delay more often than control parents, but showed no differences to this group on direct measures of language and literacy.

Bishop et al. (2004) outline two possible explanations for these findings. The first is that having children with autism makes parents more sensitive to their own communicative abilities, and may make them more likely to notice what they believe to be deficits, biasing their self-reports. In the current study, participants were not asked

about whether they had a relative with autism, but given the prevalence rate of autism it seems unlikely that there would be a high number of subjects with autistic relatives that could drive these correlations, if this bias account were true. Instead, one could consider Bishop et al.'s (2004) alternative suggestion – that the communication difficulties are real, but different to the communication skills assessed by verbal or phonological processing. They note that assessing pragmatic skills objectively is much harder to do than for structural language skills. Discrepancies between objective tests of language skill and self-reports of language problems may reflect the fact that individuals with the broader autism phenotype have pragmatic language deficits that do impact their daily communication, but these pragmatic deficits are not problematic for tests of structural language, such as verbal IQ or nonword repetition tasks. Extending this theory to the current sample, such a hypothesis would appear to fit with the finding that the two sub-scales of pragmatic language skill and social engagement exhibited strong correlations with the AQ, but the correlation with the structural language scale did not survive corrections for multiple comparisons.

Alternatively, these associations could arise because the two questionnaire measures ask similar questions. There are certainly some important differences in the questionnaires' content; the CC-SR includes many items addressing language skills, while the AQ includes items on attention switching, attention to detail and imagination. Nonetheless, examining the questionnaires' items, the correlations between the AQ and the social engagement and pragmatic subscales of the CC-SR may well be at least partly driven by some of the items being highly similar. Table 7.4 shows some of the items that are very comparable to one another.

7.4.1 Limitations

It is important to note a number of limitations with the study. Primarily, the sample had an unbalanced number of males and females, and included a high proportion of non-native English speakers. Previous studies have suggested that females

Table 7.4 Similar items from the AQ and CC-SR

AQ items	*CC-SR items*
I find it easy to work out what someone is thinking or feeling just by looking at their face	I find it hard to know when someone is upset or annoyed
People often tell me that I keep going on and on about the same thing	People say I tell them things they already knew
I enjoy social occasions	I feel anxious when I am with other people
I find social situations easy	I find it hard to express myself in a group of people
Other people frequently tell me that what I've said is impolite, even though I think it is polite	I hurt or upset people without meaning to
I prefer to do things the same way over and over again	When I have free time I choose the same favourite activity
When I talk, it isn't always easy for others to get a word in edgewise	People tell me I talk to much
It does not upset me if my daily routine is disturbed	I am okay when unexpected things happen. For example, I would keep calm if I planned to work on a computer but couldn't because it wasn't working
I am often the last to understand the point of a joke	I don't understand other people's jokes

Table 7.4: *Examples of similar items on the CC-SR and the AQ.*

exhibit stronger mu suppression responses, and lower scores on the AQ (Baron-Cohen et al., 2001; Silas, Levy, Nielsen, Slade, & Holmes, 2010). A higher number of males

would have been useful to the study, and could have increased the variance in the reported AQ scores – more were tested, but were excluded due to failing the attention check criteria. Furthermore, the inclusion of non-native English speakers may have been problematic for the CC-SR.

In addition, while our findings may have interesting implications for reported imitation deficits in ASD, the sample was composed of neurotypical participants, and therefore we must be cautious about the application of these findings to the autistic population. It may be that while imitation and autism traits show limited relationships in typical individuals without a genetic risk of autism, they may be more strongly related in autism, or relatives of autistic individuals. Perhaps a MNS that is "good enough" will allow participants to develop sufficient communication and imitation skills, and thus unless the MNS is *disordered* the correlations between scores on the behavioural measures and MNS responsivity will be relatively limited. This account would suggest that while correlations in typical populations will be weak or non-significant, group differences will arise when comparing groups of participants with and without abnormal mirror neuron systems on measures of MNS functioning, and behaviours underpinned by mirror neurons. As reviewed in Chapter 5 however, the evidence for atypical mu responsivity in autism is equivocal.

Indeed, our participants were not questioned about their family history of autism. Knowing whether any of the sample did have an increased genetic risk of autism could have been interesting, although it seems unlikely that many participants would have had parents or siblings on the spectrum, and the group size would probably have been too small to examine the effect of genetic risk for autism on the association between mu responsivity and the behavioural measures.

Furthermore, level of autism traits and communicative skills were assessed using self-report measures rather than direct testing. As noted in the discussion above, there have been discrepancies between self-report and objective language measures in previous investigations. Indeed, it could be argued that individuals with poor pragmatic or social skills may not be as aware of the communicative errors they make compared to others with good communicative abilities.

Differential relationships between the behavioural measures and mu suppression to transitive versus intransitive stimuli were not explored here, and have also not been explored in previous investigations. Here, the decision to analyse the relationships between behavioural measures of mu suppression to both hand-object and hand-no object videos together was based on concerns about inflating either Type 1 or Type 2 error, and mu suppression to hand-object and hand-no object videos was also very highly correlated.

Finally, as was also noted in Chapter 6, this study did not examine the reliability of mu suppression as an individual differences measure. There is yet limited evidence that mu suppression captures a stable or fixed quality, and therefore reported correlations with personal characteristics must be caveated. In fact, total imitation score exhibited greater variance than many of the imitation sub-scales, and thus perhaps had the best chance of correlating with mu suppression.

7.4.2 Summary

Mu suppression during action observation was found to correlate with a measure of imitation performance, but in the opposite direction of what would be predicted by previous studies and theories. This was the only correlation to arise between mu suppression and a number of individual differences, including imitation subtypes,

communications skills and autism traits. Imitation was not found to correlate with either autistic traits nor self-reported communication skills, which challenges theories that suggest these domains share a reliance on the mirror neuron system. Pragmatic language skills and social engagement skills were found to correlate with level of autistic traits, a finding that compliments previous findings of self-reported communication difficulties in relatives of autistic probands.

CHAPTER EIGHT: Discussion

Through two distinct projects, I have explored issues relating to imitation and autism, from quite different perspectives. The main questions that this thesis aimed to address were:

- Are imitation deficits a site of behavioural overlap between ASD and SLI, and do the two conditions show similar patterns of imitation deficit?
- Why do children with ALI show flat nonword repetition performance? Can broad problems with imitation, autism severity, or the social demands of the task explain this pattern?
- Is mu suppression a good index of the human mirror neuron system?

The answers that my findings have suggested for these questions are considered in turn below.

8.1. Is imitation a shared deficit between ASD and SLI?

Based on previous investigations of motor imitation in both disorders (e.g. Dohmen, Chiat, & Roy, 2013; Stone, Lemanek, & Fishel, 1990; Vukovic, Vukovic, & Stojanovik, 2010), the hypothesis that both autistic and non-autistic children with language impairment might show deficits in imitation seemed reasonable. Indeed, the pattern of deficits found by Dohmen et al. (2013) appeared to suggest that language impaired children might even show a similar profile of imitation impairments to children with ASD, including poor imitation of meaningless body movements compared to meaningful object-based imitations.

In the current study, the limited motor imitation deficits exhibited by the disorder groups make it difficult to judge whether imitation performance in these groups

is comparable. Previous reports had suggested that motor imitation deficits are seen throughout the autistic spectrum (see the review by Williams, Whiten, & Singh, 2004), but no imitation deficits were observed in the current ANL sample. Imitation of hand postures was poor in both samples of ALI and LI children, but given that these children also did poorly on a task of fine motor skill, poor performance in this task may be reflective of motor difficulties in both groups of language-impaired children. If so, this would mean that imitation *per se* is not a site of overlap between ASD and SLI , but rather that children with ALI and SLI both exhibit broad motor impairments. This would support the notion that motor and language abilities are related, and perhaps have overlapping or anatomically close neurobiological structures (as discussed in Vukovic et al., 2010).

8.2 Why do children with ALI show NWR difficulties? Are these reasons the same as for SLI?

Previous authors had suggested that autism severity or broad imitation difficulties could underlie the nonword repetition deficits in some children with autism (Whitehouse, Barry, & Bishop, 2008). However, the findings of this experiment found limited evidence to support either hypothesis. The ALI and ANL groups differed only in terms of their hand posture imitation, but not on imitating body movements, actions with objects, or the stylistic quality of the model's actions. Furthermore, the correlational analyses in Chapter 4 suggest that motor and verbal imitation skills are connected only through their links to broad cognitive factors. In addition, the LI group also showed poor hand imitation, so the pattern of results does not support a phenomimicry account where nonword repetition has different origins in the two disorders (poor PSTM in SLI vs poor imitation in ALI).

Indeed, the results showed comparable nonword repetition performance between the ALI and LI groups. A key difference between the current and previous investigations is the game-like context in which the task was set. This suggests that future studies need to consider ways of designing child-suitable games that engage study participants and sustain their attention.

We may also consider these findings from the perspective that poor nonword repetition may *reflect* language impairment, rather than capturing the root causes of it. In particular, one hypothesis is that poor vocabulary and lexical development in language-impaired children limits the development of fine-grained phonological representations, which is in turn reflected in poor nonword repetition (Munson, Kurtz, & Windsor, 2005; Snowling, Chiat, & Hulme, 1991). If this is correct, then we would predict associations between children's vocabulary and nonword repetition performance, and previous investigations of verbal memory in ASD have also suggested an important role for vocabulary (Gabig, 2008). While there was a correlation between vocabulary and nonword repetition length effects in the current study, this association was not stronger than for the other language measures.

8.3 Is mu suppression a valid measure of the mirror neuron system?

Reviewing previous mu suppression studies of language and social processes, it seems straightforward to propose alternative accounts to explain the changes in apparent mu power in many of these experiments. Indeed, the review in Chapter 5 highlighted that few of these studies have been scientifically rigorous. Of course, when developing and examining potential new methods and approaches, studies are likely to be exploratory and often small at this stage, and I do not wish to undermine the good ideas put forward by researchers. The idea that mu reactivity reflects properties of mirror

neurons is by and large quite reasonable, and merits investigation. The issues with the mu suppression literature have arisen because the technique has been employed prematurely, before work has been undertaken to establish that mu suppression is indeed a good, valid and reliable mirror neuron system activity measure, and to resolve or at least document its limitations. Nonetheless mu is already being used in studies to examine the role of mirroring systems in a huge array of processes. The important stage of establishing mu suppression's quality and limitations as an MNS index appears to have been largely skipped, although there have been a few studies using concurrent fMRI and EEG that tried to establish whether mu suppression was reflecting activity in mirroring systems identified with fMRI (Arnstein, Cui, Keysers, Maurits, & Gazzola, 2011; Braadbaart, Williams, & Waiter, 2013; Perry & Bentin, 2009). However, as noted in Chapter 5, even these studies highlighted the potential role for alpha confounding.

In the registered report in Chapter 6, it was shown that choice of baseline is an important factor in whether mu suppression is seen during action observation. In Chapter 7, the correlations between individuals' mu suppression and autistic traits, communicative and imitation skills was investigated, but associations predicted by MNS theories did not emerge. Together, this thesis suggests that mu suppression does not currently represent a good measure for indexing human MNS activity.

8.4 Implications for clinical and educational practice

There are currently limited therapies available for autism, and the evidence-base for them is generally lacking, though new "miracle cures" are often touted online (Howlin, 2010). Neurofeedback therapy has many qualities that may make it highly attractive to some families and clinicians working with individuals with autism. Firstly, neurofeedback may appear to be a treatment with a biological basis. Given that

psychological therapies for autism are largely very intensive and take a long time to make sometimes only quite small gains, biologically-based treatments may seem to offer a better solution to some families, if such solutions could be found. Neurofeedback is based on the biologically-based assumptions that a) children can learn to alter their own mu rhythms, and b) the resulting normalization of mirror neuron system functioning would lead to an improvement in autistic symptoms. Neurofeedback may also have the added benefit of not requiring children to take drugs, and is thus presumed to result in fewer side-effects than pharmacological treatments. The optimism placed in mu wave neurofeedback therapy is partly understandable from this standpoint.

However, the three chapters in Part 2 of the thesis all suggest that there are currently major shortcomings with applying this work to create neurofeedback therapies for autism, and indeed reasons to be sceptical of the neurofeedback experiments that have been done so far. In fact, the evidence for a mu suppression abnormality in autism at all is very mixed. The key methodological issues with mu suppression will need to be addressed before such a technique can be used clinically.

It was hoped that examining behavioural imitation in ASD and SLI might help shine light on sites of overlap between these conditions. For motor imitation, the imitation deficits identified in both the autistic and language impaired groups were limited, and it is unclear what part experimenter bias might have played in the impaired hand imitation findings. Imitation is a target of behavioural interventions for autism; for example, Ingersoll and colleagues have developed interventions for children which aim to increase their spontaneous imitative behaviours, known as “Reciprocal Imitation Training” (RIT) (Ingersoll & Schreibman, 2006; Ingersoll, 2010). The approach taken by RIT fits well with the social motivation deficit hypothesis of autism, and the notion that reduced social motivation could be important in the reported imitation difficulties

(Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012; Etten & Carver, 2015). However, given the limited imitation deficits I found in my study, would my findings suggest that RIT, and other interventions targeting imitation, are not likely to be fruitful (if there is no imitation deficit to rectify)?

There are two important factors to consider. The first is that RIT is perhaps more concerned with the *amount* children *choose* to imitate, not the overall accuracy of their imitations. The second is that the limited work done so far on RIT has been conducted with very young children. Both of these factors hark back to the points raised in Chapter 2, that imitation *propensity* in young children and imitation *accuracy* in older children are probably capturing different constructs. My findings of accurate elicited imitation in older children may therefore hold limited implications for intervention work of this kind.

8.5 Potential application of pre-registration publication for developmental disorders research

One of the most educational experiences during the completion of this thesis has been undertaking a registered report. It is interesting to consider what this model of publishing could offer for developmental disorders research. The experiments in Part 1 and Part 2 are very different in terms of their power and approach, arguably occupying different ends of a spectrum; one is a small behavioural study on developmental disorder groups, the other a large, well-powered adult study using neuroimaging techniques, with pre-registered methodology and analysis plans. Low power, publication bias and the potential role of pre-registration methods in addressing these problems are currently hot topics of discussion in cognitive science (Chambers,

Feredoes, Muthukumaraswamy, Suresh, & Etchells, 2014). However, there has been comparatively little coverage of these issues in developmental psychology.

There are many parallels between developmental psychology (particularly development disorders research) and other fields in biomedical science; studies using fMRI techniques and animal research also often suffer from small sample sizes. Similarly, negative results (i.e. finding limited differences between disorder and control groups) may prove particularly difficult to publish – even though demonstrating spared aspects of ability in disorder groups could be extremely informative. I have been impressed with the robustness and methodological rigour of preregistration, and having a study with an "in principle acceptance" means that researchers can rest assured that time spent collecting a large amount of data will be rewarded with a good publication – even if the results of their experiment are negative.

Could pre-registration be applied to developmental disorders research? Reflecting on my experience conducting these two experiments, I believe there are specific issues that would need to be overcome in order for pre-registration to benefit developmental disorders research. In particular, careful discussion would need to be had over the exclusionary criteria for such studies. In the mu suppression study, I was required to lay out explicitly how my adult participants would be selected, and who should be excluded; and even though the study used typical adults, I still ended up excluding quite a large number of participants. For that project, it was at least relatively straightforward to recruit more typical adults to replace those who had to be excluded; for developmental disorders research, this is obviously a far more difficult and time-consuming endeavour. Indeed, initially I had intended to exclude children from the study if they had comorbid dyspraxia or ADHD, both common comorbidities in SLI and ASD, and conditions that could be expected to impact upon children's verbal and motor

imitation performances. However, doing so would have made recruitment extremely difficult, further shrinking my sample sizes. As well as the practical issues in excluding comorbid cases, there are arguments to suggest that excluding participants on the basis of comorbid disorders has a profound effect on the findings of studies, and important implications for whether conclusions from such experiments can be generalised to wider populations (Newman, Moffitt, Caspi, & Silva, 1998). In fact, it is unclear to what extent excluding these cases leads to the removal of some important facets of a disorder group. Considering the ideas explored in this thesis about SLI and ASD, the autistic symptomatology in SLI might arise quite independent of autism because the children's language impairments lead to secondary social difficulties, or as a genuine additional deficit faced by most children (Marton, Abramoff, & Rosenzweig, 2005). If studies exclude children with SLI presenting with social difficulties, could they be removing an important aspect of SLI, rather than controlling for autism?

There are three approaches that future pre-registered developmental disorders studies could take. Firstly, they could exclude comorbidities, opting for a traditional "pure" approach. As outlined above, this approach has important implications for how studies' conclusions are used, and would realistically make a study's conduct much harder and longer. Comorbidities are common in developmental disorder groups, with most cases exhibiting symptoms of another disorder, if not meeting clinical thresholds for a full additional diagnosis. A researcher opting for this approach could arguably expect to exclude more participants than they include, if they were looking to exclude all common comorbidities. Forcing developmental researchers to exclude high numbers of cases that are arguably more representative of the groups of interest seems inefficient.

A second approach would be to accept this high level of comorbidity, and not use it as a basis for exclusionary criteria. This, however, also seems unsatisfactory – it

fails to acknowledge that children with developmental disorders are vastly heterogeneous, and potentially risks missing important differences between sub-groups of children lumped together within one diagnostic label.

I believe the best approach would be a third one: to carefully document children's comorbid difficulties, and use this information to investigate the impact of these comorbidities on studies' results. Full investigation of the effects of say, comorbid ADHD, LI, and dyspraxia in autism, would require very large sample sizes to compare many groups. An alternative would be to attempt to use quantitative measures of symptoms of inattention, language ability, and motor problems in regression-type analyses, or as covariates. Indeed, separating children into discrete categories with cut-offs is inherently arbitrary, and fails to recognise that children with one condition can show a high level of symptoms of another disorder, even if a formal secondary diagnosis has not been given. If a child with language impairment shows high levels of inattentive behaviour but fails to meet full clinical criteria for ADHD, should they be really be considered to have "just" language impairment? Capturing attentional abilities quantitatively would remove this dilemma.

Pre-registration could involve registering a study's hypotheses online (using tools such as the Open Science Framework), or undertaking a formal registered report with journals. In the case of the latter, this would assure developmental disorders researchers that even if data collection takes a long time (due to needing sufficient sample sizes to meet journals' power requirements), they will be published at the end. However, this does not escape the fact that researchers will need larger samples. One solution could be increase collaboration across research groups, as is being done in other fields of biomedical science (for a discussion on the importance of increasing collaboration to improve power and reliability of results, see the recent report by the

Academy of Medical Sciences: “Reproducibility and reliability of biomedical research: improving research practice,” 2015).

8.6 General limitations

The nonverbal cognitive weaknesses in the language impaired groups do pose limitations on the interpretations of my results. Having all disorder groups show cognitive abilities within the normal range would arguably have given the results tighter theoretical meaning, as verbal and motor imitation deficits in the LI and ALI groups could not have been put down to generally poor cognitive ability. Ideally, having another group of children with poor language *despite* nonverbal abilities within the normal range would have allowed me to explore the role of nonverbal difficulties on verbal and motor imitation. While these problems arose in my study due to small sample sizes (I could not afford to exclude children with low nonverbal IQ), there are also theoretical arguments for retaining such cases in research studies. As pointed out by Boucher et al. (2008), language ability in low functioning autism has received little research attention because it is assumed that the language difficulties in this group can be straightforwardly put down to their generally poor cognitive ability. Yet this notion misses that language ability in these individuals is often still disproportionately bad compared to other their abilities. Similarly in the field of SLI, there is little support for the idea of ‘cognitive referencing’, or that children with normal nonverbal skills have different language profiles to those with poor language *and* nonverbal skills (Tomblin, 2008). These issues have particularly important ramifications when children are diagnosed and given access to certain service provisions on the basis of nonverbal skills. Indeed, a recent report investigated the effect of nonverbal skills criteria on the prevalence, presentation and functional impairment of childhood language disorders (Norbury et al., in press). The authors found that language-impaired children with

nonverbal skills below typical exclusionary criteria levels showed minimal differences in their language profiles from children with normal-range nonverbal IQ skills, and language problems still showed significant functional impact. The study concluded that low nonverbal IQ should not be used as an exclusionary criterion for service provision, and that research into interventions needs to include such cases.

The small and heterogeneous sample in Part 1 of the thesis could have been increased and perhaps made more homogenous by opting for recruitment through the NHS, rather than through schools. Indeed, this would have allowed for recruitment to be largely driven by speech and language therapists and clinicians, who are more familiar with terms such as "Specific Language Impairment", rather than SENCos (Special Educational Needs Co-ordinators), who vary in their exposure to such terms. NHS-based recruitment might also have yielded more cases of language-impaired children with nonverbal skills within the normal range (cases that adhere more to traditional definitions of SLI). On the other hand, children with language difficulties known to the NHS may represent a group of children with more severe or complex language difficulties, so it is unclear whether recruiting through this method would have led to a more homogeneous sample. The decision to opt for schools-based rather than NHS-based recruitment was based on practicalities – obtaining ethical approval for recruitment through the NHS was thought to require much more time, and it was felt it would hold up the progress of the study.

The experiment in Part 1 contained only some types of imitation, and did not include facial or oral imitation tasks (which could be expected to be more strongly associated with language or verbal imitation). I also did not systematically investigate the effects of semantic or communicative meaning on the children's performances. The limited array of imitation tasks was selected on the basis of the neuropsychological

models of imitation, and previous patterns of ability found in other studies (e.g. Dohmen et al., 2013; Stone et al., 1990; Tessari & Rumiati, 2004). It would also have been desirable to have had a more extensive examination of the children's motor skills, especially given that motor problems are known to be so widespread in both SLI and ASD. However, while both more imitation tasks and more motor tasks could have been informative, the battery was deliberately kept to a reasonable length to avoid tiring the children.

As discussed briefly in Chapter 6, beta suppression may also represent a means to index mirror neuron activity. However, I did not examine the relationship between beta and individual differences in autism traits, imitation and communicative skills, nor did I conduct the same level of literature review into beta suppression's usage in social cognitive neuroscience. Investigating beta was proposed by one of the reviewers of the registered report, as we collected the necessary data to examine beta suppression's validity anyway. In that report, I did not find evidence that beta provided a better index of MNS activity than mu, and for that reason, and to avoid having to make statistical adjustments for multiple exploratory tests, the thesis largely restricted its focus to mu suppression.

8.7 Directions for future research

Here I outline the outstanding questions that I believe to be the most pressing, and that could represent fruitful areas for future research.

The conflation of object-based, goal-based, meaningful and familiar actions appears to be a dilemma in both adult and child imitation research, and also in studies examining the responses of the MNS to human action. In the dual-route neuropsychological models of imitation discussed throughout this thesis, imitation via

the indirect route is effectively imitation that is reconstructed, based on meaning and goals. Often in the literature, object-based and non-object based tasks or stimuli are used, but the presence or absence of an object is not consistently related to meaning or goals, and thus there is discontinuity between neuropsychological models of imitation and the tasks and stimuli used in experiments. For instance, in the mu suppression study by Fan et al. (2010) participants viewed a chess piece being rolled in the model's palm – while transitive, there is no obvious goal or meaning to such an action. Conversely, waving a hand conveys clear communicative meaning, but does not involve object interaction. Presumably imitating such a gesture would be supported by the indirect route. And of course, a gesture that may be "meaningful" to the adult examiner may not necessarily be meaningful to the child if the gesture is new to them – would we expect a six-year-old child to know the "OK" hand sign? How then should it be categorised? It remains unclear whether imitation in children has the proposed dual-route structure outlined in adult neuropsychological models, and there is still limited evidence about whether uneven performance of different tasks (i.e. better object-based, goal-oriented imitation than meaningless body-based imitation) can be put down to broad effects of task difficulty or familiarity.

In addition, we know little about how imitation develops, beyond the early years, and yet studies comparing groups of school-age children to children with autism persist. Establishing the normative structure of imitation abilities, and constructing models for imitation in normal development and adulthood, would allow for much better interpretation of the imitation deficits in autism (and SLI). Furthermore, both my own results and those of other researchers cited in my literature reviews, suggest that there may be a distinction to be drawn between projects examining imitation *propensity*

in pre-school (or younger) aged participants, and imitation *accuracy* in school-aged children.

Ideally, future studies should consider how imitation performance is related to children's motor and semantic knowledge (i.e. their familiarity with gestures that convey meaning and actions on objects). Imitation would also need to be examined longitudinally, incorporating the range of ages that researchers of imitation have investigated – from toddlerhood through to late childhood at least - and include both examinations of spontaneous imitation and elicited imitation, to examine whether imitation propensity and imitation accuracy does need to be distinguished. Clearly such thorough investigation would represent quite a large undertaking, and would be likely to be expensive and time-consuming.

The study in Part 1 is the first to find comparable effects of length on nonword repetition ability in children with ALI and non-autistic language impaired children. The hypothesis that this similarity is absent in previous studies due to motivational reasons is in need of further investigation and replication. Furthermore, the theory that nonword repetition difficulties in ALI and LI are reflective of less fine-grained phonological representations would also predict different effects of word-likeness on the nonword repetition performances of the different groups. Only one study has attempted to examine the effects of word-likeness on nonword repetition in SLI and ALI, but no group effects were found, and this finding is not discussed in the paper (Williams, Payne, & Marshall, 2012).

Part 2 clearly highlights shortcomings with mu suppression as an MNS measure, and it is important that if this technique is to continue to be used that the groundwork is done to establish its best practice. The study in Chapter 6 shows that certain baseline

techniques will provide better results than others, but baseline is just one element of mu suppression studies. For example, the field needs to define the appropriate frequency band for study. The mu band is typically defined as 8-13 Hz, but there have been suggestions that splitting the mu frequency band into upper and lower parts, or having individually-determined bands, could all provide more meaningful results (Pfurtscheller & Lopes da Silva, 1999). A particularly intriguing notion is that of individual frequency-bands. On the one hand, this notion has potential problems, in that it could allow for even more variation and noise in mu suppression experiments, and may allow researchers yet more opportunity to over-fit their analyses. On the other hand, in many other neuroimaging techniques regions are functionally defined, and functionally defining a band-width could be justified in a similar way. In addition, as noted in Chapters 5 and 6, there is a real need to examine the reliability (as well as the validity) of mu suppression. Developing objective and robust means to functionally-define mu bands, and assessing how reliable individuals' mu suppression is, will be particularly important if the pursuit of mu neurofeedback therapy continues, although my results suggest real problems with this enterprise.

One broad theoretical difficulty with the field of mu suppression is how best to validate the measure without already knowing the true nature of the human MNS. Ideally we should establish what the MNS responds to and what its characteristics are before developing measures of it. Yet, we cannot truly understand the nature of the MNS without assuming a given measure is sufficient in order to test our hypotheses about the MNS. While there is no easy solution to this predicament, I feel there needs to be wider acceptance and awareness of the fact that the MNS is a construct, able to change to suit the measures we create for it. Ideally, the best way forward is to create and agree a workable definition of the MNS, based on our current understanding, and

establish and consolidate our knowledge about the measures we have. Whether stricter definitions of the MNS arise from further experimental work or from more theoretical discussions based on the current evidence at hand, research on the MNS and mu suppression needs to have some boundaries set, lest it become totally unwieldy and inconsistent with itself.

8.8 Closing summary

I have explored children's imitation abilities, including children with communication disorders, and one of the techniques neuroscientists have used to understand the neural bases of imitation and communicative deficits in ASD. Through these projects I have encountered complex methodological and theoretical issues with the study of imitation and the MNS. My findings lead me to consider the evidence we still lack for proper understanding of imitation at the behavioural and neurophysiological levels, and to suggest reinterpretations of previous behavioural and EEG reports.

References

Aleksandrov, A. A., & Tugin, S. M. (2012). Changes in the mu rhythm in different types of motor activity and on observation of movements. *Neuroscience and Behavioral Physiology*, *42*(3), 302–307. doi:10.1007/s11055-012-9566-2

American Psychiatric Associaiton. (1994). *Diagnostic and statistical manual of mental disorders (4th ed.)*. Washington, D.C: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, D.C: American Psychiatric Association.

Arbib, M. A. (2012). *How the Brain Got Language: The Mirror System Hypothesis*. Oxford & New York: Oxford University Press.

Archibald, L. M. D., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language & Communication Disorders*, *41*(6), 675–93. doi:10.1080/13682820500442602

Arnstein, D., Cui, F., Keysers, C., Maurits, N. M., & Gazzola, V. (2011). μ-suppression during action observation and execution correlates with BOLD in dorsal premotor, inferior parietal, and SI cortices. *The Journal of Neuroscience*, *31*(40), 14243–9. doi:10.1523/JNEUROSCI.0963-11.2011

Babiloni, C., Babiloni, F., Carducci, F., Cincotti, F., Cocozza, G., Del Percio, C., … Rossini, P. M. (2002). Human Cortical Electroencephalography (EEG) Rhythms during the Observation of Simple Aimless Movements: A High-Resolution EEG Study. *NeuroImage*, *17*(2), 559–572. doi:10.1006/nimg.2002.1192

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, *36*(3), 189–208. doi:10.1016/S0021-9924(03)00019-4

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The Phonological Loop as a Language Learning Device. *Psychological Review*, *105*(1), 158–173. doi:10.1037/0033-295X.105.1.158

Baird, A. D., Scheffer, I. E., & Wilson, S. J. (2011). Mirror neuron system involvement in empathy: a critical look at the evidence. *Social Neuroscience*, *6*(4), 327–35. doi:10.1080/17470919.2010.547085

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet*, *368*(9531), 210–215. doi:10.1016/S0140-6736(06)69041-7

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: Cognitive or affective? *Journal of Autism and Developmental Disorders*, *18*(3), 379–402. doi:10.1007/BF02212194

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, *6*(6), 248–254.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (1994). *Understanding other minds: Perspectives from autism*. New York: Oxford University Press.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-

spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17.

Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology*. doi:10.1002/dev.10034

Bates, E., Thal, D., Whitesell, K., Fenson, L., & Oakes, L. (1989). Integrating Language and Gesture in Infancy. *Developmental Psychology*, *25*(6), 1004–1019.

Behmer, L. P., & Jantzen, K. J. (2011). Reading sheet music facilitates sensorimotor mu-desynchronization in musicians. *Clinical Neurophysiology*, *122*(7), 1342–7. doi:10.1016/j.clinph.2010.12.035

Bénony, C., & Bénony, H. (2005). Precocity of the acquisition of language and type II spinal muscular atrophy in 3-4-year-old children: a study of 12 cases. *European Journal of Paediatric Neurology*, *9*(2), 71–6. doi:10.1016/j.ejpn.2005.02.001

Bernier, R., Aaronson, B., & McPartland, J. (2013). The role of imitation in the observed heterogeneity in EEG mu rhythm in autism and typical development. *Brain and Cognition*, *82*(1), 69–75. doi:10.1016/j.bandc.2013.02.008

Bernier, R., Dawson, G., Webb, S., & Murias, M. (2007). EEG mu rhythm and imitation impairments in individuals with autism spectrum disorder. *Brain and Cognition*, *64*(3), 228–37. doi:10.1016/j.bandc.2007.03.004

Bishop, D. (2002). Motor immaturity and specific speech and language impairment: Evidence for a common genetic basis. *American Journal of Medical Genetics*, *63*(October 2000), 56–63. doi:10.1002/ajmg.1630

Bishop, D. V. M. (1993). Language development in children with abnormal structure or function of the speech apparatus. In D. V. M. Bishop & K. Mogford-Bevan (Eds.), *Language Development in Exceptional Circumstances* (p. 313). Psychology Press.

Bishop, D. V. M. (2013, June 7). Interpreting unexpected significant results. Retrieved from http://deevybee.blogspot.co.uk/2013/06/interpreting-unexpected-significant.html

Bishop, D. V. M. (1997). Cognitive Neuropsychology and Developmental Disorders: Uncomfortable Bedfellows. *The Quarterly Journal of Experimental Psychology*, *50A*(4), 899–923. doi:10.1080/713755740

Bishop, D. V. M. (2000). Pragmatic language impairment: A correlate of SLI, a distinct subgroup or part of the autistic continuum? In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome*. Hove: Psychology Press.

Bishop, D. V. M. (2006). What Causes Specific Language Impairment in Children? *Current Directions in Psychological Science*, *15*(5), 217–221. doi:10.1111/j.1467-8721.2006.00439.x

Bishop, D. V. M. (2010a). Overlaps between autism and language impairment: phenomimicry or shared etiology? *Behavior Genetics*, *40*(5), 618–29. doi:10.1007/s10519-010-9381-x

Bishop, D. V. M. (2010b). Which neurodevelopmental disorders get researched and why? *PLoS ONE*, *5*(11). doi:10.1371/journal.pone.0015112

Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language and Communication Disorders*,

49(4), 381–415. doi:10.1111/1460-6984.12101

Bishop, D. V. M., Adams, C. V, & Norbury, C. F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes, Brain, and Behavior*, *5*(2), 158–69. doi:10.1111/j.1601-183X.2005.00148.x

Bishop, D. V. M., Bishop, S. J., Bright, P., James, C., Delaney, T., & Tallal, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: evidence from a twin study. *Journal of Speech Language and Hearing Research*, *42*, 155–68.

Bishop, D. V. M., & Edmundson, A. (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders*, *52*(April), 156–173. doi:10.1044/jshd.5202.156

Bishop, D. V. M., Hardiman, M. J., & Barry, J. G. (2011). Is auditory discrimination mature by middle childhood? A study using time-frequency analysis of mismatch responses from 7 years to adulthood. *Developmental Science*, *14*(2), 402–416. doi:10.1111/j.1467-7687.2010.00990.x

Bishop, D. V. M., & Hayiou-Thomas, M. E. (2008). Heritability of specific language impairment depends on diagnostic criteria. *Genes, Brain and Behavior*, *7*(3), 365–372. doi:10.1111/j.1601-183X.2007.00360.x

Bishop, D. V. M., Maybery, M., Wong, D., Maley, A., Hill, W., & Hallmayer, J. (2004). Are phonological processing deficits part of the broad autism phenotype? *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *128B*(1), 54–60. doi:10.1002/ajmg.b.30039

Bishop, D. V. M., & Norbury, C. F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology. *Autism*, *9*(1), 7–27. doi:10.1177/1362361305049027

Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioral marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *36*(4), 391–403. doi:10.1111/j.1469-7610.1996.tb01420.x

Bishop, D. V. M., & Robson, J. (1989). Unimpaired short-term memory and rhyme judgement in congenitally speechless individuals: Implications for the notion of "articulatory coding." *The Quarterly Journal of Experimental Psychology Section A*, *41*(1), 123–140. doi:10.1080/14640748908402356

Bishop, D. V. M., Whitehouse, A. J. O., Watt, H. J., & Line, E. a. (2008). Autism and diagnostic substitution: evidence from a study of adults with a history of developmental language disorder. *Developmental Medicine and Child Neurology*, *50*(5), 341–5. doi:10.1111/j.1469-8749.2008.02057.x

Bishop, D. V. M., Whitehouse, A., & Sharp, M. (2009). Communication checklist-self report (CC-SR). *San Antonio, TX: The Psychological Corporation.*

Bishop, D. V. M.,(1998). Development of the Children's Communication Checklist (CCC): a method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *39*(6), 879–891. doi:10.1017/S0021963098002832

Bishop, D. V, M.,Chan, J., Adams, C., Hartley, J., & Weir, F. (2000). Conversational

responsiveness in specific language impairment: evidence of disproportionate pragmatic difficulties in a subset of children. *Development and Psychopathology*, *12*(2), 177–99.

Botting, N. (2005). Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *46*(3), 317–26. doi:10.1111/j.1469-7610.2004.00355.x

Botting, N. (2007). Comprehension Difficulties in Children with Specific Language Impairment and Pragmatic Language Impairment. In K. Cain & J. Oakhill (Eds.), *Children's comprehension problems in oral and written language: A cognitive perspective.* (pp. 81–103). New York: Guilford Press.

Boucher, J. (2012). Research review: structural language in autistic spectrum disorder - characteristics and causes. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *53*(3), 219–33. doi:10.1111/j.1469-7610.2011.02508.x

Boucher, J., Mayes, A., & Bigham, S. (2008). Memory, language and intellectual ability in low-functioning autism. *Memory in Autism: Theory and Evidence*, 268–290. doi:http://dx.doi.org/10.1017/CBO9780511490101.016

Bowers, A., Saltuklaroglu, T., Harkrider, A., & Cuellar, M. (2013). Suppression of the Mu Rhythm during Speech and Non-Speech Discrimination Revealed by Independent Component Analysis: Implications for Sensorimotor Integration in Speech Processing. *PLoS ONE*, *8*(8). doi:10.1371/journal.pone.0072024

Boyles, A. L., Harris, S. F., Rooney, A. A., & Kristina A, T. (2011). Forest Plot Viewer: A New Graphing Tool. *Epidemiology*, *22*(5), 746. doi:10.1097/EDE.0b013e318225c2be

Braadbaart, L., Williams, J. H. G., & Waiter, G. D. (2013). Do mirror neuron areas mediate mu rhythm suppression during imitation and action observation? *International Journal of Psychophysiology*, *89*(1), 99–105.

Brookman, A., McDonald, S., McDonald, D., & Bishop, D. V. M. (2013). Fine motor deficits in reading disability and language impairment: same or different? *PeerJ*, *1*, e217. doi:10.7717/peerj.217

Button, K. S., Ioannidis, J. P. a, Mokrysz, C., Nosek, B. a, Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews. Neuroscience*, *14*(5), 365–76. doi:10.1038/nrn3475

Carpenter, M., Pennington, B. E., & Rogers, S. J. (2002). Interrelations among social-cognitive skills in young children with autism. *Journal of Autism and Developmental Disorders*, *32*(2), 91–106.

Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G., & Rizzolatti, G. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(45), 17825–17830. doi:10.1073/pnas.0706273104

Catts, H., Adlof, S., Hogan, T., & Weismer, S. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, …*, *48*(6), 1378–1396. doi:10.1044/1092-4388(2005/096).Are

Ceponiene, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., & Yaguchi, K. (2003). Speech-sound-selective auditory impairment in children with autism: they can perceive but do not attend. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(9), 5567–5572. doi:10.1073/pnas.0835631100

Chambers, C. D., Feredoes, E., Muthukumaraswamy, Suresh, D., & Etchells, P. J. (2014). Instead of "playing the game" it is time to change the rules: Registered Reports at AIMS Neuroscience and beyond. *AIMS Neuroscience*, *1*(1), 4–17. doi:10.3934/Neuroscience.2014.1.4

Charman, T., & Baird, G. (2002). Practitioner review: Diagnosis of autism spectrum disorder in 2- and 3-year-old children. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *43*(3), 289–305.

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.

Chatrian, G. E. (1976). The mu rhythms. In A. Rémond (Ed.), *Handbook of electroencephalography and clinical neurophysiology* (pp. 104–114). Amsterdam: Elsevier Scientific Pub. Co.

Chatrian, G. E., Petersen, M. C., & Lazarte, J. a. (1959). The blocking of the rolandic wicket rhythm and some central changes related to movement. *Electroencephalography and Clinical Neurophysiology*, *11*(3), 497–510. doi:10.1016/0013-4694(59)90048-3

Cheng, Y., Lee, P.-L., Yang, C.-Y., Lin, C.-P., Hung, D., & Decety, J. (2008). Gender differences in the mu rhythm of the human mirror-neuron system. *PloS One*, *3*(5), e2113. doi:10.1371/journal.pone.0002113

Cheng, Y., Tzeng, O., & Decety, J. (2006). Gender differences in the human mirror system: A magnetoencephalography study. *Neuroreport*, *17*(1), 1115–1119.

Cheng, Y., Yang, C.-Y., Lin, C.-P., Lee, P.-L., & Decety, J. (2008). The perception of pain in others suppresses somatosensory oscillations: a magnetoencephalography study. *NeuroImage*, *40*(4), 1833–40. doi:10.1016/j.neuroimage.2008.01.064

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Science*, *16*(4), 231–239. doi:10.1016/j.tics.2012.02.007

Chomsky, N. (1959). Review of Verbal Behavior. *Language*, *35*(1), 26–58. doi:10.1371/journal.pone.0005772

Cicchetti, D. (1994). Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardised Assessment Instruments in Psychology. *Psychological Assessment*, *6*(4), 282–290.

Cochin, S., & Barthelemy, C. (1999). Observation and execution of movement: similarities demonstrated by quantified electroencephalography. *European Journal of Neuroscience*, *11*(June 1998), 1839–1842.

Coll, M.-P., Bird, G., Catmur, C., & Press, C. (2015). Cross-modal repetition effects in the mu rhythm indicate tactile mirroring during action observation. *Cortex*, *63*, 121–131. doi:10.1016/j.cortex.2014.08.024

Constantino, J. N., & Todd, R. D. (2003). Autistic Traits in the General Population. *Archives of General Psychiatry*, *60*, 524–530. doi:10.1001/archpsyc.60.5.524.Text

Conti-Ramsden, G., Simkin, Z., & Botting, N. (2006). The prevalence of autistic spectrum disorders in adolescents with a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*(6), 621–8. doi:10.1111/j.1469-7610.2005.01584.x

Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. (2014). Mirror neurons: From origin to function. *Behav Brain Sci*, *37*(2), 177–192. doi:10.1017/s0140525x13000903

Cooper, N. R., Simpson, A., Till, A., Simmons, K., & Puzzo, I. (2013). Beta event-related desynchronization as an index of individual differences in processing human facial expression: Further investigations of autistic traits in typically developing adults. *Frontiers in Human Neuroscience*, *7*.

Corballis, M. C. (2003). From mouth to hand: gesture, speech, and the evolution of right-handedness. *The Behavioral and Brain Sciences*, *26*(2), 199–208;

Cramer, A. O. J., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P. P. P., … Wagenmakers, E.-J. (2014). Hidden Multiplicity in Multiway ANOVA: Prevalence, Consequences, and Remedies. *Psychonomic Bulletin & Review*, 1–20. doi:10.3758/s13423-015-0913-5

Crawcour, S., Bowers, A., Harkrider, A., & Saltuklaroglu, T. (2009). Mu wave suppression during the perception of meaningless syllables: EEG evidence of motor recruitment. *Neuropsychologia*, *47*(12), 2558–63. doi:10.1016/j.neuropsychologia.2009.05.001

Creutzfeldt, O., Grunewald, G., Simonova, O., & Schmitz, H. (1969). Changes of the basic rhythm of the EEG during the performance of mental and vasomotor task. In C. R. Evans & T. B. Mulholland (Eds.), *Attention in Neurophysiology* (pp. 148–168). London: Butterworths.

Cuellar, M., Bowers, A., Harkrider, A. W., Wilson, M., & Saltuklaroglu, T. (2012). Mu suppression as an index of sensorimotor contributions to speech processing: Evidence from continuous EEG signals. *International Journal of Psychophysiology*, *85*(2), 242–248.

Cuevas, K., Cannon, E. N., Yoo, K. H., & Fox, N. A. (2014). The Infant EEG Mu Rhythm: Methodological Considerations and Best Practices. *Developmental Review*, *34*(1), 26–43. doi:10.1038/nature13314.A

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. a, Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, *9*(1), 28–30. doi:10.1038/nn1611

Davies, M., & Stone, T. (1995). *Mental Simulation: Evaluations and Applications*. Oxford: Blackwell.

Dawson, M. (2010, August 31). Are you high or low functioning? Examples from autism research. Retrieved from http://autismcrisis.blogspot.co.uk/2010/08/are-you-high-or-low-functioning.html

Dawson, G., & Adams, A. (1984). Imitation and social responsiveness in autistic children. *Journal of Abnormal Child Psychology*, *12*(2), 209–25.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. doi:10.1016/j.jneumeth.2003.10.009

Dewey, D. (1991). Praxis and sequencing skills in children with sensorimotor dysfunction. *Developmental Neuropsychology*, *7*(2), 197–206. doi:10.1080/87565649109540487

Dewey, D., Cantell, M., & Crawford, S. G. (2007). Motor and gestural performance in children with autism spectrum disorders, developmental coordination disorder, and/or attention deficit hyperactivity disorder. *Journal of the International Neuropsychological Society*,

13(2), 246–56. doi:10.1017/S1355617707070270

Dinstein, I., Heeger, D. J., Lorenzi, L., Minshew, N. J., Malach, R., & Behrmann, M. (2012). Unreliable evoked responses in autism. *Neuron*, *75*(6), 981–91. doi:10.1016/j.neuron.2012.07.026

Dockrell, J., Ricketts, J., Charman, T., & Lindsay, G. (2010). Profiles of need and provision for children with language impairments and autism spectrum disorders in mainstream schools: a prospective study. Report conducted for the Department of Education. Retrieved from https://www.gov.uk/government/publications/profiles-of-need-and-provision-for-children-with-language-impairments-and-autism-spectrum-disorders-in-mainstream-schools-a-prospective-study

Dohmen, A., Chiat, S., & Roy, P. (2013). Nonverbal imitation skills in children with specific language delay. *Research in Developmental Disabilities*, *34*(10), 3288–3300. doi:10.1016/j.ridd.2013.06.004

Donald, M. (2005). Imitation and Mimesis. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation, Volume 2: From Neuroscience to Social Science - Imitation, Human Development, and Culture.* (pp. 298–315). Cambridge, Massachusetts: MIT Press.

Dumas, G., Soussignan, R., Hugueville, L., Martinerie, J., & Nadel, J. (2014). Revisiting mu suppression in autism spectrum disorder. *Brain Research*, *1585*, 108–19. doi:10.1016/j.brainres.2014.08.035

Dziuk, M. a, Gidley Larson, J. C., Apostu, a, Mahone, E. M., Denckla, M. B., & Mostofsky, S. H. (2007). Dyspraxia in autism: association with motor, social, and communicative deficits. *Developmental Medicine and Child Neurology*, *49*(10), 734–9. doi:10.1111/j.1469-8749.2007.00734.x

Edwards, L. A. (2014). A meta-analysis of imitation abilities in individuals with autism spectrum disorders. *Autism Research*, *7*, 363–380. doi:10.1002/aur.1379

Engel, A. K., & Fries, P. (2010). Beta-band oscillations--signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–65. doi:10.1016/j.conb.2010.02.015

Etten, H. M. Van, & Carver, L. J. (2015). Does Impaired Social Motivation Drive Imitation Deficits in Children with Autism Spectrum Disorder ?, *Review Journal of Autism and Developmental Disorders*, *2*(3) 310–319. doi:10.1007/s40489-015-0054-9

Fabbri-Destro, M., Cattaneo, L., Boria, S., & Rizzolatti, G. (2009). Planning actions in autism. *Experimental Brain Research*, *192*(3), 521–525. doi:10.1007/s00221-008-1578-3

Fan, Y.-T., Decety, J., Yang, C.-Y., Liu, J.-L., & Cheng, Y. (2010). Unbroken mirror neurons in autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *51*(9), 981–8. doi:10.1111/j.1469-7610.2010.02269.x

Farrant, B. M., Maybery, M. T., & Fletcher, J. (2010). Socio-emotional engagement, joint attention, imitation, and conversation skill: Analysis in typical development and specific language impairment. *First Language*, *31*(1), 23–46. doi:10.1177/0142723710365431

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–91.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology*, *61*(6), 825–850. doi:10.1080/17470210701623605

Fletcher-Watson, S., Leekam, S. R., Benson, V., Frank, M. C., & Findlay, J. M. (2009). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia*, *47*(1), 248–257. doi:10.1016/j.neuropsychologia.2008.07.016

Folstein, S., & Rutter, M. (1977). Infantile autism: a genetic study of 21 twin pairs. *Journal of Child Psychology and Psychiatry*, *18*(January), 297–322.

Folstein, S., Santangelo, S. L., Gilman, S. E., Piven, J., Landa, R., Lainhart, J., … Wzorek, M. (1999). Predictors of Cognitive Test Patterns in Autism Families. *Journal of Child Psychology and Psychiatry*, *40*(7), 1117–1128. doi:10.1111/1469-7610.00528

Fombonne, E., Bolton, P., Prior, J., Jordan, H., & Rutter, M. (1997). A family study of autism: Cognitive patterns and levels in parents and siblings. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *38*(6), 667–683. doi:10.1111/j.1469-7610.1997.tb01694.x

Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N., Vanderwert, R. E., … van IJzendoorn, M. H. (2015). Assessing Human Mirror Activity With EEG Mu Rhythm: A Meta-Analysis. *Psychological Bulletin*. doi:10.1037/bul0000031

Freitag, C. M., Kleser, C., von Gontard, A., & von Gontardf, A. (2006). Imitation and language abilities in adolescents with Autism Spectrum Disorder without language delay. *European Child & Adolescent Psychiatry*, *15*(5), 282–91. doi:10.1007/s00787-006-0533-8

Friedrich, E. V. C., Sivanathan, A., Lim, T., Suttie, N., Louchart, S., Pillen, S., & Pineda, J. A. (2015). An Effective Neurofeedback Intervention to Improve Social Interactions in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *45*(12), 4084–4100. doi:10.1007/s10803-015-2523-5

Frith, U. (1989). *Autism : explaining the enigma*. Oxford, UK: Blackwell Science.

Gabig, C. S. (2008). Verbal working memory and story retelling in school-age children with autism. *Language, Speech, and Hearing Services in Schools*, *39*(4), 498–511. doi:10.1044/0161-1461(2008/07-0023)

Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Social Neuroscience*, *3*(3-4), 317–33. doi:10.1080/17470910701563608

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain : A Journal of Neurology*, *119 (Pt 2*, 593–609.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the mind-reading. *Trends in Cognitive Sciences*, *2*(12), 493–501. doi:10.1016/S1364-6613(98)01262-5

Gastaut, H. 1952. Étude électrocorticographique de la réativité des rhythmes rolandiques. *Revue Neurologique, 87*:176-182.

Gastaut, H. J., & Bert, J. (1954). EEG changes during cinematographic presentation; moving picture activation of the EEG. *Electroencephalography and Clinical Neurophysiology*, *6*(3), 433–444. doi:10.1016/0013-4694(54)90058-9

Gastaut, H., Bacher, F., Bert, J. et al .1959a. Étude des corrélations entre les variables d électroencéphalographiques et psychométriques (M.M.P.I. Rosenweig). *Revue Neurologique, 101:*376-384

Gaustaut, H, Dongier, S., and Dongier, M. 1959b. Electroencephalographie et nevroses. *Revue*

Neurologique. 101:435-436.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*, 336–360. doi:10.1016/0749-596X(90)90004-J

Gathercole, S. E., Willis, C. S., Baddeley, a D., & Emslie, H. (1994). The Children's Test of Nonword Repetition: a test of phonological working memory. *Memory*, *2*(2), 103–127. doi:10.1080/09658219408258940

Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *NeuroImage*, *35*(4), 1674–84. doi:10.1016/j.neuroimage.2007.02.003

Gernsbacher, M. A., Dawson, M., & Goldsmith, H. H. (2005). Three reasons not to believe in an autism epidemic. *Current Directions in Psychological Science*, *14*(2), 55–58. doi:10.1111/j.0963-7214.2005.00334.x

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385. doi:10.1093/cercor/7.4.374

Gillberg, C. (2010). The ESSENCE in child psychiatry: Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations. *Research in Developmental Disabilities*, *31*(6), 1543–51. doi:10.1016/j.ridd.2010.06.002

Gillberg, C., & Steffenburg, S. (1987). Outcome and prognostic factors in infantile autism and similar conditions: a population-based study of 46 cases followed through puberty. *Journal of Autism and Developmental Disorders*, *17*(2), 273–287. doi:10.1007/BF01495061

Gola, M., Magnuski, M., Szumska, I., & Wróbel, A. (2013). EEG beta band activity is related to attention and attentional deficits in the visual performance of elderly subjects. *International Journal of Psychophysiology*, *89*(3), 334–41. doi:10.1016/j.ijpsycho.2013.05.007

Grafton, S. T. (2009). Embodied Cognition and the Simulation of Action to Understand Others. *Annals of the New York Academy of Sciences*, *1156*(1), 97–117. doi:10.1111/j.1749-6632.2009.04425.x

Gray, D. E. (1994). Coping with autism: Stresses and strategies. *Sociology of Health & Illness*, *16*(3), 275–300. doi:10.1111/1467-9566.ep11348729

Green, D., Charman, T., Pickles, A., Chandler, S., Loucas, T., Simonoff, E., & Baird, G. (2009). Impairment in movement skills of children with autistic spectrum disorders. *Developmental Medicine and Child Neurology*, *51*(4), 311–6. doi:10.1111/j.1469-8749.2008.03242.x

Gutsell, J. N., & Inzlicht, M. (2010). Empathy constrained: Prejudice predicts reduced mental simulation of actions during observation of outgroups. *Journal of Experimental Social Psychology*, *46*(5), 841–845. doi:10.1016/j.jesp.2010.03.011

Gutsell, J. N., & Inzlicht, M. (2013). Using EEG mu-suppression to explore group biases in motor resonance. In B. Derks, D. Scheepers, & N. Ellemers (Eds.), *The Neuroscience of Prejudice and Intergroup Relations* (pp. 279–298). New York: Psychology Press. doi:10.4324/9780203124635

Hallgren, K. a. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34. doi:10.1016/j.biotechadv.2011.08.021.Secreted

Hamilton, A. (2008). Emulation and mimicry for social interaction: a theoretical approach to imitation in autism. *Quarterly Journal of Experimental Psychology*, *61*(1), 101–15. doi:10.1080/17470210701508798

Hamilton, A. (2013). Reflecting on the mirror neuron system in autism: A systematic review of current theories. *Developmental Cognitive Neuroscience*, *3*(max 6), 91–105. doi:10.1016/j.dcn.2012.09.008

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience*, *9*(10), 1218–20. doi:10.1038/nn1770

Hari, R., & Forss, N. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(December), 15061–15065.

Hari, R., & Salmelin, R. (1997). Human cortical oscillations : a neuromagnetic view through the skull. *Trends in Neurosciences*, *20*(1), 44–49.

Hatton, D., & Sideris, J. (2006). Autistic behavior in children with fragile X syndrome: prevalence, stability, and the impact of FMRP. *American Journal of Medical Genetics Part A*, *1813*, 1804–1813. doi:10.1002/ajmg.a

Heimann, M., Strid, K., & Smith, L. (2006). Exploring the relation between memory, gestural communication, and the emergence of language in infancy: A longitudinal study. *Infant and Child Development*, *15*(3), 233–249.

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., & Mazziotta, J. C. (2003). The essential role of Broca's area in imitation. *European Journal of Neuroscience*, *17*(5), 1123–1128. doi:10.1046/j.1460-9568.2003.02530.x

Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences*, *5*(6), 253–261.

Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews*, *34*(4), 575–583. doi:10.1016/j.neubiorev.2009.11.007

Heyes, C. (2013). A new approach to mirror neurons: Developmental history, system-level theory and intervention experiments. *Cortex*, *49*(10), 2946–2948. doi:10.1016/j.cortex.2013.07.002

Heyes, C. (2016). Homo imitans? Seven reasons why imitation couldn't possibly be associative. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150069. doi:10.1098/rstb.2015.0069

Hill, A. P., van Santen, J., Gorman, K., Langhorst, B. H., & Fombonne, E. (2015). Memory in language-impaired children with and without autism. *Journal of Neurodevelopmental Disorders*, *7*(1), 19. doi:10.1186/s11689-015-9111-z

Hill, E. L. (2001). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, *36*(2), 149–171. doi:10.1080/13682820010019874

Hobson, H. M., & Bishop, D. V. M. (in press) *Mu suppression - a good measure of the human*

mirror neuron system? To appear in *Cortex.*

Hobson, R., & Lee, A. (1998). Hello and goodbye: A study of social engagement in autism. *Journal of Autism and Developmental Disorders*, *28*(2), 117–127.

Hobson, R. P., & Hobson, J. a. (2008). Dissociable aspects of imitation: a study in autism. *Journal of Experimental Child Psychology*, *101*(3), 170–85. doi:10.1016/j.jecp.2008.04.007

Honaga, E., Ishii, R., Kurimoto, R., Canuet, L., Ikezawa, K., Takahashi, H., … Takeda, M. (2010). Post-movement beta rebound abnormality as indicator of mirror neuron system dysfunction in autistic spectrum disorder: an MEG study. *Neuroscience Letters*, *478*(3), 141–5. doi:10.1016/j.neulet.2010.05.004

Howlin, P. (2010). Evaluating psychological treatments for children with autism-spectrum disorders. *Advances in Psychiatric Treatment*, *16*(2), 133–140. doi:10.1192/apt.bp.109.006684

Iacoboni, M. (2005). Neural mechanisms of imitation. *Current Opinion in Neurobiology*, *15*(6), 632–7. doi:10.1016/j.conb.2005.10.010

Iglewicz, B., & Banerjee, S. (2001). A simple univariate outlier identification procedure. *Proceedings of the Annual Meeting of the American Statistical Association*, (2001), 5–9.

Ingersoll, B. (2010). Brief report: Pilot randomized controlled trial of reciprocal imitation training for teaching elicited and spontaneous imitation to children with autism. *Journal of Autism and Developmental Disorders*, *40*(9), 1154–1160.

Ingersoll, B., & Lalonde, K. (2010a). The impact of object and gesture imitation training on language use in children with autism spectrum disorder. *Journal of Speech, Language and Hearing Research*, *53*(August), 1040–1051.

Ingersoll, B., & Meyer, K. (2011a). Do object and gesture imitation skills represent independent dimensions in autism? *Journal of Developmental and Physical Disabilities*, *23*(5), 421–431. doi:10.1007/s10882-011-9237-1

Ingersoll, B., & Meyer, K. (2011b). Examination of correlates of different imitative functions in young children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*(3), 1078–1085. doi:10.1016/j.rasd.2010.12.001

Ingersoll, B., & Schreibman, L. (2006). Teaching Reciprocal Imitation Skills to Young Children with Autism Using a Naturalistic Behavioral Approach: Effects on Language, Pretend Play, and Joint Attention. *Journal of Autism and Developmental Disorders*, *36*(4), 487–505.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Med*, *2*(8), e124. doi:10.1371/journal.pmed.0020124

Jenson, D., Bowers, A. L., Harkrider, A. W., Thornton, D., Cuellar, M., & Saltuklaroglu, T. (2014). Temporal dynamics of sensorimotor integration in speech perception and production: independent component analysis of EEG data. *Frontiers in Psychology*, *5*(July), 1–17. doi:10.3389/fpsyg.2014.00656

Jenson, D., Thornton, D., Saltuklaroglu, T., & Harkrider, A. (2014). Speech perception, production, and the sensorimotor mu rhythm. *Proceedings of the 2014 Biomedical Sciences and Engineering Conference*, 1–4. doi:10.1109/BSEC.2014.6867736

Jerome, A. C., Fujiki, M., Brinton, B., & James, S. L. (2002). Self-Esteem in Children With

Specific Language Impairment. *Journal of Speech, Language & Hearing Research*, *45*(4), 700. doi:10.1044/1092-4388(2002/056)

Kayser, J., & Tenke, C. (2005). Reference-free quantification of EEG spectra: combining current source density (CSD) and frequency principal components analysis (fPCA). *Clinical Neurophysiology*, *116*(12), 2826–46. doi:10.1016/j.clinph.2005.08.007

Kayser, J., & Tenke, C. (2006a). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clinical Neurophysiology*, *117*(2), 348–68. doi:10.1016/j.clinph.2005.08.034

Kayser, J., & Tenke, C. (2006b). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates. *Clinical Neurophysiology*, *117*(2), 369–80. doi:10.1016/j.clinph.2005.08.033

Kimber, A. C. (1990). Data Analysis for Possibly Censored Data from Skewed Distributions. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *39*(1), 21–30.

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An Investigation of Language Impairment in Autism: Implications for Genetic Subgroups. *Language and Cognitive Processes*, *16*(2-3), 287–308. doi:10.1080/01690960042000058

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research. Brain Research Reviews*, *29*(2-3), 169–95.

Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., & Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neuroscience Letters*, *244*(2), 73–76. doi:10.1016/S0304-3940(98)00122-0

Koziol, L. F., Budding, D. E., & Chidekel, D. (2012). From Movement to Thought: Executive Function, Embodied Cognition, and the Cerebellum. *The Cerebellum*, *11*(2), 505–525. doi:10.1007/s12311-011-0321-y

Króliczak, G., Piper, B., & Frey, S. (2011). Atypical lateralization of language predicts cerebral asymmetries in parietal gesture representations. *Neuropsychologia*, *49*(7), 1698–1702. doi:10.1016/j.neuropsychologia.2011.02.044.Atypical

Kuhl, P. K. (2007). Is speech learning "gated" by the social brain? *Developmental Science*, *10*(1), 110–120. doi:10.1111/j.1467-7687.2007.00572.x

Kuhlman, W. N. (1978). Functional topography of the human mu rhythm. *Electroencephalography and Clinical Neurophysiology*, *44*(1), 83–93. doi:10.1016/0013-4694(78)90107-4

Kumar, S., Riddoch, M. J., & Humphreys, G. (2013). Mu rhythm desynchronization reveals motoric influences of hand action on object recognition. *Frontiers in Human Neuroscience*, *7*. doi:10.3389/fnhum.2013.00066

Laakso, M.-L., Poikkeus, a.-M., Katajamaki, J., & Lyytinen, P. (1999). Early intentional communication as a predictor of language development in young toddlers. *First Language*, *19*(56), 207–231. doi:10.1177/014272379901905604

Laws, G., & Bishop, D. (2004). Pragmatic language impairment and social deficits in Williams syndrome: a comparison with Down's syndrome and specific language impairment. *International Journal of Language & Communication Disorders*, *39*(1), 45–64. doi:10.1080/13682820310001615797

Le Bel, R. M., Pineda, J. a, & Sharma, A. (2009). Motor-auditory-visual integration: The role of the human mirror neuron system in communication and communication disorders. *Journal of Communication Disorders*, *42*(4), 299–304. doi:10.1016/j.jcomdis.2009.03.011

Le Couteur, A., Bailey, A., Goode, S., Pickles, A., Gottesman, I., Robertson, S., & Rutter, M. (1996). A Broader Phenotype of Autism: The Clinical Spectrum in Twins. *Journal of Child Psychology and Psychiatry*, *37*(7), 785–801. doi:10.1111/j.1469-7610.1996.tb01475.x

Leighton, J., Bird, G., Charman, T., & Heyes, C. (2008). Weak imitative performance is not due to a functional "mirroring" deficit in adults with Autism Spectrum Disorders. *Neuropsychologia*, *46*(4), 1041–9. doi:10.1016/j.neuropsychologia.2007.11.013

Leonard, L. B. (2014). *Children with Specific Language Impairment*. MIT Press.

Lepage, J.-F., Saint-Amour, D., & Théoret, H. (2008). EEG and neuronavigated single-pulse TMS in the study of the observation/execution matching system: are both techniques measuring the same process? *Journal of Neuroscience Methods*, *175*(1), 17–24. doi:10.1016/j.jneumeth.2008.07.021

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review*, *94*(4), 412–426. doi:10.1037/0033-295X.94.4.412

Leyfer, O. T., Tager-Flusberg, H., Dowd, M., Tomblin, J. B., & Folstein, S. E. (2008). Overlap between autism and specific language impairment: comparison of Autism Diagnostic Interview and Autism Diagnostic Observation Schedule scores. *Autism Research : Official Journal of the International Society for Autism Research*, *1*(5), 284–96. doi:10.1002/aur.43

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1–36. doi:10.1016/0010-0277(85)90021-6

Lindgren, K., Folstein, S., Tomblin, J., & Tager–Flusberg, H. (2009). Language and reading abilities of children with autism spectrum disorders and specific language impairment and their first-degree relatives. *Autism Research*, *2*(1), 22–38. doi:10.1002/aur.63.Language

Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*, *13*(3), 110–4. doi:10.1016/j.tics.2008.11.008

Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language assessment and development in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(8), 1426–1438.

Marshall, P. J., & Meltzoff, A. (2011). Neural mirroring systems: Exploring the EEG mu rhythm in human infancy. *Developmental Cognitive Neuroscience*, *1*(2), 110–123. doi:10.1016/j.dcn.2010.09.001.Neural

Marshall, P. J., Young, T., & Meltzoff, A. N. (2011). Neural correlates of action observation and execution in 14-month-old infants: An event-related EEG desynchronization study. *Developmental Science*, *14*(3), 474–480. doi:10.1111/j.1467-7687.2010.00991.x

Martineau, J., Cochin, S., Magne, R., & Barthelemy, C. (2008). Impaired cortical activation in autistic children: is the mirror neuron system involved? *International Journal of Psychophysiology*, *68*(1), 35–40. doi:10.1016/j.ijpsycho.2008.01.002

Marton, K. (2009). Imitation of body postures and hand movements in children with specific language impairment. *Journal of Experimental Child Psychology*, *102*(1), 1–13. doi:10.1016/j.jecp.2008.07.007

Marton, K., Abramoff, B., & Rosenzweig, S. (2005). Social cognition and language in children with specific language impairment (SLI). *Journal of Communication Disorders*, *38*(2), 143–62. doi:10.1016/j.jcomdis.2004.06.003

Mathewson, K. J., Jetha, M. K., Drmic, I. E., Bryson, S. E., Goldberg, J. O., & Schmidt, L. a. (2012). Regional EEG alpha power, coherence, and behavioral symptomatology in autism spectrum disorder. *Clinical Neurophysiology*, *123*(9), 1798–1809. doi:10.1016/j.clinph.2012.02.061

Matson, J. L., & Kozlowski, A. M. (2011). The increasing prevalence of autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*(1), 418–425. doi:10.1016/j.rasd.2010.06.004

Matthews, N., Gold, B. J., Sekuler, R., & Park, S. (2013). Gesture imitation in schizophrenia. *Schizophrenia Bulletin*, *39*(1), 94–101. doi:10.1093/schbul/sbr062

McDuffie, A., Yoder, P., & Stone, W. (2005). Prelinguistic predictors of vocabulary in young children with Autism spectrum disorders. *Journal of Speech, Language, and Hearing Research*, *48*(5), 1080–1097.

McEwen, F., Happe, F., Bolton, P., Rijsdijk, F., Ronald, A., Dworzynski, K., & Plomin, R. (2007). Origins of individual differences in imitation: Links with language, pretend play, and socially insightful behavior in two-year-old twins. *Child Development*, *78*(2), 474–492.

McFarland, D. J., Miner, L. A., Vaughan, T. M., & Wolpaw, J. R. (2000). Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topography*, *12*(3), 177–186. doi:10.1007/s00221-009-2062-4

Meltzoff, A. N. (1999). Origins of theory of mind, cognition and communication. *Journal of Communication Disorders*, *32*(4), 251–69.

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Development*, *54*(3), 702–9.

Meltzoff, A. N., & Moore, M. K. (2002). Imitation, memory, and the representation of persons. *Infant Behavior & Development*, *25*, 39–61.

Mervis, C. B., & Velleman, S. L. (2011). Children with Williams Syndrome: Language, Cognitive, and Behavioural Characteristics and their Implications for Intervention. *Perspectives on Language Learning and Education*, *18*(3), 98–107. doi:10.1016/j.surg.2006.10.010.Use

Mizuhara, H. (2012). Cortical dynamics of human scalp EEG origins in a visually guided motor execution. *NeuroImage*, *62*(3), 1884–95. doi:10.1016/j.neuroimage.2012.05.072

Moore, A., Gorodnitsky, I., & Pineda, J. (2012). EEG mu component responses to viewing emotional faces. *Behavioural Brain Research*, *226*(1), 309–16. doi:10.1016/j.bbr.2011.07.048

Moreno, I., de Vega, M., & Leon, I. (2013). Understanding action language modulates oscillatory mu and beta rhythms in the same way as observing actions. *Brain and Cognition*, *82*(3), 236–242.

Mostofsky, S. H., Dubey, P., Jerath, V. K., Jansiewicz, E. M., Goldberg, M. C., & Denckla, M. B. (2006). Developmental dyspraxia is not limited to imitation in children with autism spectrum disorders. *Journal of the International Neuropsychological Society*, *12*(03), 314–326. doi:10.1017/S1355617706060437

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The Influence of Vocabulary Size, Phonotactic Probability, and Wordlikeness On Nonword Repetition of Children With and Without Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *48*, 1033–1047. doi:10.1044/1092-4388(2005/072)

Muthukumaraswamy, S. D., & Johnson, B. W. (2004). Changes in rolandic mu rhythm during observation of a precision grip. *Psychophysiology*, *41*(1), 152–6. doi:10.1046/j.1469-8986.2003.00129.x

Muthukumaraswamy, S. D., & Johnson, B. W. (2004). Primary motor cortex activation during action observation revealed by wavelet analysis of the EEG. *Clinical Neurophysiology*, *115*(8), 1760–6. doi:10.1016/j.clinph.2004.03.004

Muthukumaraswamy, S. D., Johnson, B. W., Gaetz, W. C., & Cheyne, D. O. (2006). Neural processing of observed oro-facial movements reflects multiple action encoding strategies in the human brain. *Brain Research*, *1071*(1), 105–12. doi:10.1016/j.brainres.2005.11.053

Muthukumaraswamy, S. D., & Singh, K. D. (2008). Modulation of the human mirror neuron system during cognitive activity. *Psychophysiology*, *45*(6), 896–905. doi:10.1111/j.1469-8986.2008.00711.x

Muthukumaraswamy, S., Johnson, B., & McNair, N. (2004). Mu rhythm modulation during observation of an object-directed grasp. *Cognitive Brain Research*, *19*, 195–201. doi:10.1016/j.cogbrainres.2003.12.001

Nation, K. (2008). Developmental language disorders. *Psychiatry*, *7*(6), 266–269. doi:10.1016/j.mppsy.2008.04.003

Nation, K., & Hulme, C. (2011). Learning to read changes children's phonological skills: Evidence from a latent variable longitudinal study of reading and nonword repetition. *Developmental Science*, *14*(4), 649–659. doi:10.1111/j.1467-7687.2010.01008.x

Newman, D. L., Moffitt, T. E., Caspi, A., & Silva, P. A. (1998). Comorbid mental disorders: Implications for treatment and sample selection. *Journal of Abnormal Psychology*, *107*(2), 305–311. doi:10.1037/0021-843X.107.2.305

Niedermeyer, E., & Silva, F. da. (2005). *Electroencephalography: basic principles, clinical applications, and related fields* (5th ed.). Philadelphia: Lippincott Williams & Wilkins.

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charmand, T., Simonoff, E., … Andrew, P. (in press). The impact of NVIQ on prevalence and clinical presentation of language disorder: evidence from a population study. *Journal of Child Psychology and Psychiatry*, 1–37.

Nyström, P. (2008). The infant mirror neuron system studied with high density EEG. *Social Neuroscience*, *3*(3-4), 334–47. doi:10.1080/17470910701563665

Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., & Pineda, J. a. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Brain Research. Cognitive Brain Research*, *24*(2), 190–8. doi:10.1016/j.cogbrainres.2005.01.014

Oberman, L. M., McCleery, J. P., Ramachandran, V. S., & Pineda, J. a. (2007). EEG evidence for mirror neuron activity during the observation of human and robot actions: Toward an analysis of the human qualities of interactive robots. *Neurocomputing*, *70*(13-15), 2194–2203. doi:10.1016/j.neucom.2006.02.024

Oberman, L. M., Pineda, J. a, & Ramachandran, V. S. (2007). The human mirror neuron system: a link between action observation and social skills. *Social Cognitive and Affective*

Neuroscience, *2*(1), 62–6. doi:10.1093/scan/nsl022

Oberman, L. M., & Ramachandran, V. S. (2007). The simulating social mind: the role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin*, *133*(2), 310–327. doi:10.1037/0033-2909.133.2.310

Oberman, L. M., Ramachandran, V. S., & Pineda, J. A. (2008). Modulation of mu suppression in children with autism spectrum disorders in response to familiar or unfamiliar stimuli: the mirror neuron hypothesis. *Neuropsychologia*, *46*(5), 1558–65. doi:10.1016/j.neuropsychologia.2008.01.010

Olineck, K. M., & Poulin-Dubois, D. (2005). Infants' Ability to Distinguish Between Intentional and Accidental Actions and Its Relation to Internal State Language. *Infancy*, *8*(1), 91–100.

Onton, J., Westerfield, M., Townsend, J., & Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neuroscience and Biobehavioral Reviews*, *30*(6), 808–22. doi:10.1016/j.neubiorev.2006.06.007

Pellegrino, G. Di, Fadiga, L., & Fogassi, L. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, *91*, 176–180.

Perry, A., & Bentin, S. (2009). Mirror activity in the human brain while observing hand movements: a comparison between EEG desynchronization in the mu-range and previous fMRI results. *Brain Research*, *1282*, 126–32. doi:10.1016/j.brainres.2009.05.059

Perry, A., & Bentin, S. (2010). Does focusing on hand-grasping intentions modulate electroencephalogram μ and α suppressions? *NeuroReport*, *21*(16), 1050–1054.

Perry, A., Bentin, S., Bartal, I. B.-A., Lamm, C., & Decety, J. (2010). "Feeling" the pain of those who are different from us: Modulation of EEG in the mu/alpha range. *Cognitive, Affective & Behavioral Neuroscience*, *10*(4), 493–504. doi:10.3758/CABN.10.4.493

Perry, A., Troje, N. F., & Bentin, S. (2010). Exploring motor system contributions to the perception of social information: Evidence from EEG activity in the mu/alpha frequency range. *Social Neuroscience*, *5*(3), 272–84. doi:10.1080/17470910903395767

Pfurtscheller, G. (1992). Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest. *Electroencephalography and Clinical Neurophysiology*, *83*, 62–69.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857. doi:10.1016/S1388-2457(99)00141-8

Pfurtscheller, G., & Neuper, C. (1994). Event-related synchronisation of mu rhythms in the EEG over the hand area in man. *Neuroscience Letters*, *174*(43), 93–96.

Pfurtscheller, G., Neuper, C., Andrew, C., & Edlinger, G. (1997). Foot and hand area mu rhythms. *International Journal of Psychophysiology*, *26*(1-3), 121–35.

Pfurtscheller, G., Neuper, C., & Krausz, G. (2000). Functional dissociation of lower and upper frequency mu rhythms in relation to voluntary limb movement. *Clinical Neurophysiology*, *111*(10), 1873–1879. doi:10.1016/S1388-2457(00)00428-4

Pfurtscheller, G., Stancák, a., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band — an electrophysiological correlate of cortical idling: A review. *International*

Journal of Psychophysiology, *24*(1-2), 39–46. doi:10.1016/S0167-8760(96)00066-9

Piaget, J. (1951). *Play, dreams and imitation in childhood*.

Pickard, K. E., & Ingersoll, B. R. (2014). Brief Report: High and Low Level Initiations of Joint Attention, and Response to Joint Attention: Differential Relationships with Language and Imitation. *Journal of Autism and Developmental Disorders*. doi:10.1007/s10803-014-2193-8

Pineda, J. A., & Hecht, E. (2008). Mirroring and mu rhythm involvement in social cognition: are there dissociable subcomponents of theory of mind? *Biological Psychology*, *80*(3), 306–14. doi:10.1016/j.biopsycho.2008.11.003

Pineda, J. A. (2005). The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing". *Brain Research Reviews*, *50*(1), 57–68. doi:10.1016/j.brainresrev.2005.04.005

Pineda, J. A., Brang, D., Hecht, E., Edwards, L., Carey, S., Bacon, M., … Rork, A. (2008). Positive behavioral and electrophysiological changes following neurofeedback training in children with autism. *Research in Autism Spectrum Disorders*, *2*(3), 557–581. doi:10.1016/j.rasd.2007.12.003

Pineda, J. A., Carrasco, K., Datko, M., Pillen, S., & Schalles, M. (2014). Neurofeedback training produces normalization in behavioural and electrophysiological measures of high-functioning autism. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*, 20130183.

Pineda, J. A., Giromini, L., Porcelli, P., Parolin, L., & Viglione, D. J. (2011). Mu suppression and human movement responses to the Rorschach test. *NeuroReport: For Rapid Communication of Neuroscience Research*, *22*(5), 223–226.

Pineda, J. A., Pelton, H., Aragon, O., Bai, J., Erhart, M., Chambers, D., … Linderman, D. (2011). Behavioral and Electrophysiological Characterization of Induced Neural Plasticity in the Autistic Brain. In V. Eapen (Ed.), *Autism - A Neurodevelopmental Journey from Genes to Behaviour* (pp. 445–462). InTech.

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin*, *131*(4), 592–617. doi:10.1037/0033-2909.131.4.592

Poon, K., Watson, L., Baranek, G., & Poe, M. (2012). To what extent do joint attention, imitation, and object play behaviors in infancy predict later communication and intellectual functioning in ASD? *Journal of Autism and Developmental Disorders*, *42*(6), 1064–1074. doi:10.1007/s10803-011-1349-z.To

Preston, S. D., & Waal, F. B. M. De. (2002). Empathy: Its ultimate and proximate bases. *Behavioural and Brain Sciences*, *25*, 1–72.

Puzzo, I., Cooper, N. R., Cantarella, S., Fitzgerald, P. B., & Russo, R. (2013). The effect of rTMS over the inferior parietal lobule on EEG sensorimotor reactivity differs according to self-reported traits of autism in typically developing individuals. *Brain Research*, *1541*, 33–41. doi:10.1016/j.brainres.2013.10.016

Puzzo, I., Cooper, N. R., Vetter, P., & Russo, R. (2010). EEG activation differences in the pre-motor cortex and supplementary motor area between normal individuals with high and low traits of autism. *Brain Research*, *1342*(2008), 104–110. doi:10.1016/j.brainres.2010.04.060

Quadrelli, E., & Turati, C. (2015). Origins and development of mirroring mechanisms: A

neuroconstructivist framework. *British Journal of Developmental Psychology*, n/a–n/a. doi:10.1111/bjdp.12110

Raymaekers, R., Wiersema, J. R., & Roeyers, H. (2009). EEG study of the mirror neuron system in children with high functioning autism. *Brain Research*, *1304*, 113–121.

Renzi, E. De, Motti, F., & Nichelli, P. (1980). Imitating gestures: A quantitative approach to ideomotor apraxia. *Archives of Neurology*, *37*, 6–10.

Reproducibility and reliability of biomedical research: improving research practice. (2015). Report prepared by the Academy of Medical Sciences. Retrieved from http://www.acmedsci.ac.uk/viewFile/56531416e2949.pdf

Riby, D. M., Brown, P. H., Jones, N., & Hanley, M. (2012). Brief report: faces cause less distraction in autism. *Journal of Autism and Developmental Disorders*, *42*(4), 634–9. doi:10.1007/s10803-011-1266-1

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research*, *39*(6), 1239–1257. doi:10.1044/jshr.3906.1239

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research*, *38*(4), 850–863. doi:10.1044/jshr.3804.850

Riches, N. G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2011). Non-word repetition in adolescents with specific language impairment and autism plus language impairments: a qualitative analysis. *Journal of Communication Disorders*, *44*(1), 23–36. doi:10.1016/j.jcomdis.2010.06.003

Ritter, P., Moosmann, M., & Villringer, A. (2009). Rolandic alpha and beta EEG rhythms' strengths are inversely related to fMRI-BOLD signal in primary somatosensory and motor cortex. *Human Brain Mapping*, *30*(4), 1168–87. doi:10.1002/hbm.20585

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, *21*(5), 188–194.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 II Area F5 and the control of distal movements. *Experimental Brain Research*, *71*, 491–507. doi:10.1007/BF00248742

Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, *27*, 169–92. doi:10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., & Fabbri-Destro, M. (2010). Mirror neurons: from discovery to autism. *Experimental Brain Research.*, *200*(3-4), 223–37. doi:10.1007/s00221-009-2002-3

Rizzolatti, G., Fabbri-destro, M., & Cattaneo, L. (2009). Mirror neurons and their clinical relevance. *Nature Clinical Practice Neurology*, *5*(1), 24–34. doi:10.1038/ncpneuro

Rodgon, M. M., & Kurdek, L. A. (1977). Vocal and gestural imitation in 8-, 14-, and 20-month-old children. *The Journal of Genetic Psychology*, *131*(1), 115–123. doi:10.1017/CBO9781107415324.004

Rogers, S., Cook, I., & Greiss-Hess, L. (2005). *Mature imitation task*. M.I.N.D. Institute, University of California, Davis (Unpublished coding manual.).

Rogers, S. J., Hepburn, S. L., Stackhouse, T., & Wehner, E. (2003). Imitation performance in

toddlers with autism and those with other developmental disorders. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *44*(5), 763–81.

Rogers, S. J., & Pennington, B. F. (1991). A theoretical approach to the deficits in infantile autism. *Development and Psychopathology*, *3*(02), 137. doi:10.1017/S0954579400000043

Rogers, S. J., Young, G. S., Cook, I., Giolzetti, A., & Ozonoff, S. (2008). Deferred and immediate imitation in regressive and early onset autism. *Journal of Child Psychology and Psychiatry*, *49*(4), 449–457.

Rogers, S. J., & Williams, J. H. (2006). *Imitation and the Social Mind: Autism and Typical Development*. Guilford Press.

Rossi, S., Tecchio, F., Pasqualetti, P., Ulivelli, M., Pizzella, V., Romani, G. ., … Rossini, P. . (2002). Somatosensory processing during movement observation in humans. *Clinical Neurophysiology*, *113*(1), 16–24. doi:10.1016/S1388-2457(01)00725-8

Rumiati, R. I., Weiss, P. H., Tessari, A., Assmus, A., Zilles, K., Herzog, H., & Fink, G. R. (2005). Common and differential neural mechanisms supporting imitation of meaningful and meaningless actions. *Journal of Cognitive Neuroscience*, *17*(9), 1420–31. doi:10.1162/0898929054985374

Russell, J. (1997). *Autism as an executive disorder*. New York: Oxford University Press.

Rutter, M., & Thapar, A. (2014). Genetics of Autism Spectrum Disorders. In *Handbook of Autism and Pervasive Developmental Disorders, Fourth Edition* (pp. 411–423). Hoboken, New Jersey: John Wiley & Sons, Inc.

Ruysschaert, L., Warreyn, P., Wiersema, J. R., Oostra, A., & Roeyers, H. (2014). Exploring the Role of Neural Mirroring in Children with Autism Spectrum Disorder. *Autism Research*, 1–10. doi:10.1002/aur.1339

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism*, *6*(1), 2. doi:10.1186/2040-2392-6-2

Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation*, *110*(6), 417–438.

Schmitt, A. J., & Wodrich, D. L. (2004). Validation of a Developmental Neuropsychological Assessment (NEPSY) through comparison of neurological, scholastic concerns, and control groups. *Archives of Clinical Neuropsychology : The Official Journal of the National Academy of Neuropsychologists*, *19*(8), 1077–1093.

Schoppenhorst, M., & Brauer, F. (1980). The significance of coherence estimates in determining central alpha and mu activities. *Electroencephalography and Clinical Neurophysiology*, *48*, 25–33.

Semel, E., Wiig, E., & Secord, W. (2006). *CELF 4: clinical evaluation of language Fundamentals* (Fourth Edi.). San Antonio: Psychological Corp.

Sevlever, M., & Gillis, J. M. (2010). An examination of the state of imitation research in children with autism: Issues of definition and methodology. *Research in Developmental Disabilities*, *31*(5), 976–84. doi:10.1016/j.ridd.2010.04.014

Siegel, L. S. (1982). Early Cognitive and Environmental Correlates of Language Development

at 4 Years. *International Journal of Behavioral Development*, *5*(4), 433–444. doi:10.1177/016502548200500403

Sieratzki, J. S., & Woll, B. (2002). Toddling into language: Precocious language development in motor-impaired children with spinal muscular atrophy. *Lingua*, *112*(6), 423–433. doi:10.1016/S0024-3841(01)00054-7

Sigman, M., & McGovern, C. W. (2005). Improvement in cognitive and language skills from preschool to adolescence in autism. *Journal of Autism and Developmental Disorders*, *35*(1), 15–23. doi:10.1007/s10803-004-1027-5

Silas, J., Levy, J. P., Nielsen, M. K., Slade, L., & Holmes, A. (2010). Sex and individual differences in induced and evoked EEG measures of action observation. *Neuropsychologia*, *48*(9), 2417–26. doi:10.1016/j.neuropsychologia.2010.03.004

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–66. doi:10.1177/0956797611417632

Singh, F., Pineda, J., & Cadenhead, K. S. (2011). Association of impaired EEG mu wave suppression, negative symptoms and social functioning in biological motion processing in first episode of psychosis. *Schizophrenia Research*, *130*(1-3), 182–186.

Skuse, D., Warrington, R., Bishop, D., Chowdhury, U., Lau, J., Mandy, W., & Place, M. (2004). The developmental, dimensional and diagnostic interview (3di): a novel computerized assessment for autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*(5), 548–558. doi:10.1097/00004583-200405000-00008

Slaughter, V., & McConnell, D. (2003). Emergence of joint attention: relationships between gaze following, social referencing, imitation, and naming in infancy. *The Journal of Genetic Psychology*, *164*(1), 54–71. doi:10.1080/00221320309597503

Smith, I. M., & Bryson, S. E. (1994). Imitation and action in autism: a critical review. *Psychological Bulletin*, *116*(2), 259–73.

Snowling, M. J., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes. Some comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics*, *12*, 369–373. doi:10.1017/S0142716400009279

Sonnby-Borgstrom, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology*, *43*(5), 433–443. doi:10.1111/1467-9450.00312

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*(2), 245–251. doi:10.1037/0033-2909.87.2.245

Sterman, M., Macdoland, L., & Stone, R. (1974). Biofeedback training of sensorimotor EEG in man and its effects on epilepsy. *Epilepsia*, *15*, 395–416.

Stipacek, a., Grabner, R. H., Neuper, C., Fink, a., & Neubauer, a. C. (2003). Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load. *Neuroscience Letters*, *353*(3), 193–196. doi:10.1016/j.neulet.2003.09.044

Stone, W. L., Ousley, O. Y., & Littleford, C. D. (1997). Motor imitation in young children with autism: what's the object? *Journal of Abnormal Child Psychology*, *25*(6), 475–85.

Stone, W. L., & Yoder, P. J. (2001). Predicting Spoken Language Level in Children with

Autism Spectrum Disorders. *Autism*, *5*(4), 341–361. doi:10.1177/1362361301005004002

Stone, W., Lemanek, K., & Fishel, P. (1990). Play and imitation skills in the diagnosis of autism in young children. *Pediatrics*, *86*(2), 267–272.

Stothard, S. E., Snowling, M. J., Chipchase, B. B., & Kaplan, C. A. (1998). Language-Impaired Preschoolers: A Follow-Up Into Adolescence. *Journal of Speech Language and Hearing Research*, *41*(April 2016), 407–418. doi:10.1044/jslhr.4102.407

Strid, K., Heimann, M., & Tjus, T. (2013). Pretend play, deferred imitation and parent-child interaction in speaking and non-speaking children with autism. *Scandinavian Journal of Psychology*, *54*(1), 26–32. doi:10.1111/sjop.12003

Strid, K., Tjus, T., Smith, L., Meltzoff, A. N., & Heimann, M. (2006). Infant recall memory and communication predicts later cognitive development. *Infant Behavior & Development*, *29*(4), 545–53. doi:10.1016/j.infbeh.2006.07.002

Szatmari, P., Bryson, S. E., Boyle, M. H., Streiner, D. L., & Duku, E. (2003). Predictors of outcome among high functioning children with autism and Asperger syndrome. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *44*(4), 520–528. doi:10.1111/1469-7610.00141

Szatmari, P., Jones, M., & Tuff, L. (1993). Lack of cognitive impairment in first-degree relatives of children with pervasive developmental disorders. *Journal of the American …*, *32*(6).

Szenkovits, G., Peelle, J. E., Norris, D., & Davis, M. H. (2012). Individual differences in premotor and motor recruitment during speech perception. *Neuropsychologia*, *50*(7), 1380–1392. doi:10.1016/j.neuropsychologia.2012.02.023

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1430), 303–14. doi:10.1098/rstb.2002.1198

Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: evidence from Williams syndrome. *Cognition*, *76*(1), 59–90. doi:10.1016/S0010-0277(00)00069-X

Tai, Y. F., Scherfler, C., Brooks, D. J., Sawamoto, N., & Castiello, U. (2004). The Human Premotor Cortex Is "Mirror" Only for Biological Actions. *Current Biology*, *14*(2), 117–120. doi:10.1016/j.cub.2004.01.005

Tallal, P., & Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature*, *241*(5390), 468–9. doi:10.1038/241468a0

Tamura, T., Gunji, A., Takeichi, H., Shigemasu, H., Inagaki, M., Kaga, M., & Kitazaki, M. (2012). Audio-vocal monitoring system revealed by mu-rhythm activity. *Frontiers in Psychology*, *3*(July), 225. doi:10.3389/fpsyg.2012.00225

Tangwiriyasakul, C., Verhagen, R., van Putten, M. J. a M., & Rutten, W. L. C. (2013). Importance of baseline in event-related desynchronization during a combination task of motor imagery and motor observation. *Journal of Neural Engineering*, *10*(2), 026009. doi:10.1088/1741-2560/10/2/026009

Taylor, L. J., Maybery, M. T., & Whitehouse, A. J. O. (2012). Do Children with Specific Language Impairment have a Cognitive Profile Reminiscent of Autism? A Review of the Literature. *Journal of Autism and Developmental Disorders*, *42*, 2067–2083. doi:10.1007/s10803-012-1456-5

Tessari, A., Canessa, N., Ukmar, M., & Rumiati, R. I. (2007). Neuropsychological evidence for a strategic control of multiple routes in imitation. *Brain*, *130*, 1111–26. doi:10.1093/brain/awm003

Tessari, A., & Rumiati, R. I. (2004). The strategic control of multiple routes in imitation of actions. *Journal of Experimental Psychology. Human Perception and Performance*, *30*(6), 1107–16. doi:10.1037/0096-1523.30.6.1107

Tomasello, M. (1992). The social bases of language acquisition. *Social Development*, *1*(1), 67–87. doi:10.1111/1467-9507.ep12953134

Tomblin, J. B. (2008). Validating diagnostic standards for SLI using adolescent outcomes. In C. F. Norbury, J. B. Tomblin, & D. V. M. Bishop (Eds.), *Understanding Developmental Language Disorders: From Theory to Practice* (pp. 99–114). Hove: Psychology Press.

Tomblin, J. B., Hafeman, L. L., & O'Brien, M. (2003). Autism and autism risk in siblings of children with specific language impairment. *International Journal of Language & Communication Disorders*, *38*(3), 235–50. doi:10.1080/1368282031000086363

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of Specific Language Impairment in Kindergarten Children. *Journal of Speech Language and Hearing Research*, *40*(6), 1245. doi:10.1044/jslhr.4006.1245

Toth, K., Munson, J., Meltzoff, A. N., & Dawson, G. (2006). Early predictors of communication development in young children with autism spectrum disorder: joint attention, imitation, and toy play. *Journal of Autism and Developmental Disorders*, *36*(8), 993–1005. doi:10.1007/s10803-006-0137-7

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex*, *41*(3), 399–433. doi:10.1016/S0010-9452(08)70276-4

Ulloa, E. R., & Pineda, J. A. (2007). Recognition of point-light biological motion: Mu rhythms and mirror neuron activity. *Behavioural Brain Research*, *183*(2), 188–194.

van Elk, M., van Schie, H. T., Zwaan, R. A., & Bekkering, H. (2010). The functional role of motor activation in language processing: motor cortical oscillations support lexical-semantic retrieval. *Neuroimage*, *50*(2), 665–677. doi:S1053-8119(10)00004-2 [pii]\r10.1016/j.neuroimage.2009.12.123

Vanvuchelen, M., Roeyers, H., & De Weerdt, W. (2007). Nature of motor imitation problems in school-aged boys with autism: a motor or a cognitive problem? *Autism*, *11*(3), 225–40. doi:10.1177/1362361307076846

Vanvuchelen, M., Roeyers, H., & De Weerdt, W. (2011). Development and initial validation of the Preschool Imitation and Praxis Scale (PIPS). *Research in Autism Spectrum Disorders*, *5*(1), 463–473.

Virji-Babul, N., Moiseev, A., Cheung, T., Weeks, D., Cheyne, D., & Ribary, U. (2008). Changes in mu rhythm during action observation and execution in adults with Down syndrome: Implications for action representation. *Neuroscience Letters*, *436*(2), 177–180.

Vivanti, G. (2014). The Importance of Distinguishing Propensity Versus Ability to Imitate in ASD Research and Early Detection. *Journal of Autism and Developmental Disorders*, *45*(4), 1119–1120. doi:10.1007/s10803-014-2254-z

Vivanti, G., Nadig, A., Ozonoff, S., & Rogers, S. J. (2008). What do children with autism attend to during imitation tasks? *Journal of Experimental Child Psychology*, *101*(3), 186–205.

doi:10.1016/j.jecp.2008.04.008

Vukovic, M., Vukovic, I., & Stojanovik, V. (2010). Investigation of language and motor skills in Serbian speaking children with specific language impairment and in typically developing children. *Research in Developmental Disabilities*, *31*(6), 1633–44. doi:10.1016/j.ridd.2010.04.020

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. *Psychological Science*, *4*(3), 274–290. doi:10.1111/j.1745-6924.2009.01132.x

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence* (First Edit.). San Antonio: Psychological Corp.

Whitehouse, A. J. O., Barry, J. G., & Bishop, D. V. M. (2008). Further defining the language impairment of autism: is there a specific language impairment subtype? *Journal of Communication Disorders*, *41*(4), 319–36. doi:10.1016/j.jcomdis.2008.01.002

Williams, D., Payne, H., & Marshall, C. (2012). Non-word Repetition Impairment in Autism and Specific Language Impairment: Evidence for Distinct Underlying Cognitive Causes. *Journal of Autism and Developmental Disorders*. doi:10.1007/s10803-012-1579-8

Williams, J. H. G., Whiten, A., & Singh, T. (2004). A Systematic Review of Action Imitation in Autistic Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *34*(3), 285–299. doi:10.1023/B:JADD.0000029551.56735.3a

Williams, J. H. G., Whiten, A., & Singh, T. (2004). A Systematic Review of Action Imitation in Autistic Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *34*(3), 285–299. doi:10.1023/B:JADD.0000029551.56735.3a

Williams, J. H., Whiten, a, Suddendorf, T., & Perrett, D. I. (2001). Imitation, mirror neurons and autism. *Neuroscience and Biobehavioral Reviews*, *25*(4), 287–95.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. doi:10.3758/BF03196322

Woodruff, C. C., & Maaske, S. (2010). Action execution engages human mirror neuron system more than action observation. *Neuroreport*, *21*(6), 432–5. doi:10.1097/WNR.0b013e3283385910

Woodruff, C. C., Martin, T., & Bilyk, N. (2011). Differences in self- and other-induced Mu suppression are correlated with empathic abilities. *Brain Research*, *1405*, 69–76. doi:10.1016/j.brainres.2011.05.046

Wray, C., Norbury, C. F., & Alcock, K. (2015). Gestural abilities of children with specific language impairment. *International Journal of Language & Communication Disorders*, n/a–n/a. doi:10.1111/1460-6984.12196

Young, G. S., Rogers, S. J., Hutman, T., Rozga, A., Sigman, M., & Ozonoff, S. (2011). Imitation from 12 to 24 months in autism and typical development: a longitudinal Rasch analysis. *Developmental Psychology*, *47*(6), 1565–78. doi:10.1037/a0025418

Zambrana, I. M., Ystrom, E., Schjolberg, S., & Pons, F. (2013). Action imitation at 11/2 years is better than pointing gesture in predicting late development of language production at 3 years of age. *Child Development*, *84*(2), 560–573.

Appendix A: Formula for calculating partial correlations.

$$r_{12.3} = \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{(1-{r_{13}}^2)(1-{r_{23}}^2)}}$$

Where:

- $r_{12.3}$ = correlation between imitation and language with developmental level/age partialled out.
- r_{12} = correlation between language and imitation
- r_{13} = correlation between language and developmental level/age
- r_{23} = correlation between imitation and developmental level/age

Appendix B: Studies presented in forest plot of correlations between language and imitation in typically developing children

Abbreviations used include: MCDI – MacArthur Communicative Development Inventory

Table B.1: Studies reporting correlations between language and imitation in typically developing children

Paper	Participants	Results	Language tasks	Imitation tasks	Motor tasks	Forest Plot Reference
Charman et al. (2000)	13 TD infants (note: only 12 children did the imitation task), seen at 20 and 44 months.	When previous language ability and IQ are partialled out, imitation at 20 months significantly correlated with expressive language at 44 months (r = .60), but not receptive language (r = .49).	Reynell Developmental Language Scales.	Instrumental imitation task (Meltzhoff, 1988)	None.	1-2 Charman 2000
Laakso, Poikkeus, Katajamaki & Lyytinen (1999)	111 infants, with data collected at 14, 18 and 24 months.	Imitation of mother's object-directed actions not concurrently related to language production at 14 months (r=.34) or language comprehension at 14 months (r = .17), or of language comprehension at 18 months (r = .13). However, imitation at 14 months was predictive of language production at 24 months (r = .27).	MCDI, Reynell Developmental Language Scales and the Bayley Scales of Infant Development.	Based on parent report and observation of mother-child interactions.	None.	5-8 Laakso 1999
Olineck & Poulin-Dubois (2005)	30 18-month-olds and 26 14-month olds.	Understanding of intention (indexed by selective imitation of intentional action) not concurrently related to vocabulary, but did predict mental state vocabulary at 30 months (r = .42) (partialled for age). Negative result for the MCDI is not reported.	MCDI, plus a questionnaire about the children's understanding of internal state words. Internal state vocabulary questionnaire was administered at 30 months.	Imitation of intentional action and accidental actions.	None.	10 Olineck 2005

Table B.1: Studies reporting correlations between language and imitation in typically developing children						
Paper	**Participants**	**Results**	**Language tasks**	**Imitation tasks**	**Motor tasks**	**Forest Plot Reference**
Siegal (1982)	46 preterm and 50 full term infants, at 8 months, 12 months and 4 years.	Gestural imitation at 8 months and 12 months correlated with language at 4 years (composite of comprehension and expression scores). For imitation at 8 months, *r* =.32, and at 12 months *r* = 0.21.	Reynell Developmental Language Scales	Uzgiris-Hunt Scale, which includes a "Gesture Imitation" task, which includes familiar and unfamiliar actions.	Motor Development Index (fine and gross motor development)	11-13 Siegal 1982
Zambrana, Ystrom, Schjolberg & Pons (2013)	42,517 children aged 18 months, and for 28,107 of the same children at 36 months. (42, 227 in final analysis – 20,775 girls and 21,452 boys).	Reduced imitation at 18 months outperforms poor pointing as a non-verbal marker of a liability for late language production (between 18 and 36 months of age). Imitation at 18 months correlated with language production (*r* = .40), and language comprehension (*r* = 0.66) at 18 months. Imitation also correlated with language production at 36 months (*r* = 0.47).	Questionnaire - 7 questions relating to general language level.	Questionnaire - parents were asked: Does your child imitate you (e.g., you make a face–will your child imitate it?) Does your child copy the activities you do, such as wiping up a spill, sweeping, shaving?	None.	16-18 Zambrana 2013
Strid et al. (2006)	26 children, at 9 months, 14 months and 4 years.	Deferred imitation did not correlate with vocabulary (*r* = .08), nor the verbal scale of the McCarthy (*r* = .25).	Peabody Picture Vocabulary Test, and the verbal subscale of the McCarthy Scales of Children's Abilities.	Deferred imitation.	Motor subscale of McCarthy Scales of Children's Abilities.	22-23 Strid 2006
Heimann et al. (2006)	30 Swedish infants, at 6, 9, and 14 months.	Deferred imitation at 9 months was positively but not significantly related to vocal comprehension at 14 months, (*r* = 0.29). Deferred imitation at 14 months was significantly correlated with concurrent vocal comprehension (*r* =0.57).	Swedish version of MCBI (SECDI).	Deferred imitation.	None.	24-25 Heimann 2006

Table B.1: Studies reporting correlations between language and imitation in typically developing children						
Paper	**Participants**	**Results**	**Language tasks**	**Imitation tasks**	**Motor tasks**	**Forest Plot Reference**
Bates et al. (1989)	41 children, aged 13-15 months.	Total imitation did not significantly correlate with words comprehend (r = .05) or words produced (r = .25).	"Language and Gesture Inventory" given to parents to fill in. Documented words produced and words comprehended by the child.	Nine imitation trials. Each child was asked to imitate three gestures presented with a) supportive language, b) neutral language, and c) contradictory language. Block as a place holder in actions.	None.	26-27 Bates 1989
McEwen et al. (2007)	5,206 twins, aged 2 years.	Imitation significantly correlated with vocabulary (r = .27). Authors also calculated heritability of imitation.	MCDI.	Parents were asked to test their child's imitation of gestures and facial movements (from PARCA - Parent Report of Child Abilities).	None.	9 McEwen 2007
Slaughter & McConnell (2003)	Participants: 60 TD infants, aged 8 - 14 months.	Imitation did not correlate with vocabulary for names (r =.009), vocabulary for instrumental actions (r = -.014) nor or any other measure. Correlations were controlled for age.	MCDI (note possible floor effects). Infants' vocabularies were divided into names and instrumental words in the analyses.	Novel actions with objects (only two trials).	None.	19-20 Slaughter 2003
Farrant, Maybery & Fletcher (2010)	Two samples of typically developing children (samples sizes 94 and 93), and 30 children with SLI, aged 4 to 7.	For both samples of typical children, the correlation between conversational skill and imitation was significant (r = .43, for both samples of TD children.)	Parental questionnaire about conversational skill.	Retrospective questionnaire, with 3 questions pertaining to imitation.	None.	3-4 Farrant 2010

Table B.1: Studies reporting correlations between language and imitation in typically developing children						
Paper	**Participants**	**Results**	**Language tasks**	**Imitation tasks**	**Motor tasks**	**Forest Plot Reference**
Vukovic et al. (2010)	30 children with SLI and 30 TD children, 4-7 years.	In both typical and SLI samples, imitation of complex movement was correlated with expressive vocabulary (r = .53) and comprehension (r = .63), but not articulation (r = .32). Note that non-imitation motor tasks also correlated with the language measures in the TD group.	Token Test (comprehension), Boston Naming test (vocabulary), test of articulation, and story-generation task.	Imitation of simple movements and complex movements. The first contained 10 tasks were concerned with movements of the hands and 10 tasks concerned with movements of the arms (from the shoulders and from the elbows). Complex movements included 16 tasks, involving finer movements with the fingers.	McCarthy's Scales of Children's Abilities, including the Coordination of legs subtest and Coordination of arms subtest.	13-15 Vukovick 2010
Strid, Heimann & Tjus (2013)	23 TD and 20 ASD. Mean age of TD group was 35 months.	Language age did not correlate with deferred imitation in the typical group (*sho* = .13).	Peabody Picture Vocabulary Test. The Swedish version of the MCDI (SECDI) was also used for the children with autism.	Deferred imitation.	None.	21 Strid 2013

Appendix C: Studies presented in forest plot of correlations between language and imitation in children with an autism spectrum disorder.

Abbreviations used include: MCDI - MacArthur Communicative Development Inventory; CELF – Clinical Evaluation of Language Fundamentals; PIPS – Preschool Imitation and Praxis Scale.

Table C.1: Studies reporting correlations between language and imitation in children with an ASD

Paper	Participants	Results	Language tasks	Imitation tasks	Motor tasks	Forest Plot Reference
Dawson & Adams (1984)	15 children, aged 4-6 years.	Significant correlation between imitation and verbalization.	Verbalization during free play.	Uzgiris-Hunt Developmental Scales and spontaneous imitation. The UH scale includes familiar simple actions, complex familiar actions, unfamiliar gestures visible to the child and unfamiliar gestures invisible to the child.	None.	1 Dawson 1984
Stone, Ousley & Littleford (1997)	26 children, aged 2 years. At time 1, average age was 31.5 months. At time 2, average age was 45.6 months.	Total imitation correlated with language at Time 1 (r =.49), and at Time 2 (r =.55). Body imitation alone correlated with CDI words (at Time 1 r =.49; at Time 2 r = .43). Correlations between object-based actions and the CDI did not reach significance.	MCDI	Motor Imitation Scale	None.	2 -7 Stone 1997
Stone & Yoder (2001)	35 children, at age 2 and age 4.	Imitation at age 2 correlated with expressive language aggregate at age 4 (r = 0.55). Once previous language ability was partialed out, r = 0.38.	MCDI, SICD-R (Sequenced Inventory of Communication Development-Revised) and Preschool Language Scales-3.	Motor Imitation Scale.	None.	8 -9 Stone & Yoder 2001

Table C.1: Studies reporting correlations between language and imitation in children with an ASD						
Paper	**Participants**	**Results**	**Language tasks**	**Imitation tasks**	**Motor tasks**	**Forest Plot Reference**
Carpenter, Pennington & Rogers (2002)	12 children with autism and 11 children with developmental delays, matched group-wise for chronological age, verbal mental age, and nonverbal mental age. Aged 4 years.	Pass-fail scores for the imitation task correlated with referential language (r = .67). Instrumental imitation task was not correlated with referential language (r = -.09).	Children's spontaneous, non-echoed productions of referential words were coded during the entire session. Children passed referential language if they produced at least one referential word.	An imitative learning task, which included both instrumental and arbitrary imitation. They also performed a manual/facial imitation task but do not investigate correlations. (Note: issues with ceiling effects in comparison group.)	None.	10 - 11 Carpenter 2002
Rogers, Hepburn, Stackhouse & Wehner (2003)	24 children with autism (aged 34 months), 18 children with FXS, 20 children with other developmental disorders, and 15 typically-developing children.	Correlations were partialled for developmental level. Expressive language did not correlate with manual imitation (r = .07), object imitation (r = .11), oral imitation (r = .27), or total imitation scores (r =.27) for the children with autism. Imitation did show relationship with the severity of autism, joint attention skills, and fine motor skills.	Mullen Scales of Early Language.	Manual, object-based and oro-facial (not verbal) imitation.	Merrill-Palmer Scale for visual-spatial problem-solving abilities. Praxis battery designed to challenge motor planning and execution abilities.	12 - 15 Rogers 2003
Luyster, Kadlec, Carter & Tager-Flusberg (2008)	164 children, aged 18 to 33 months.	Non-verbal cognitive ability, gestures and imitation were significant predictors of expressive language. Imitation correlated with composite expressive language (r = .57), and receptive language (r = .55).	Receptive and expressive scores from the Mullen Scales, Vineland (Communication domain) and MCDI.	9 tasks involving imitating examiner's manual and oral-facial movements and manipulating objects (three manual acts, three oral-facial actions, and three actions on objects).	Mullen Scales of Early Learning and the Vineland Adaptive Behaviour Scales. A composite score for motor skills was created from these.	16 -17 Luyster 2008

Table C.1: Studies reporting correlations between language and imitation in children with an ASD						
Paper	**Participants**	**Results**	**Language tasks**	**Imitation tasks**	**Motor tasks**	**Forest Plot Reference**
Vanvuchelen, Roeyers & De Weerdt (2011)	47 children with autism, aged 1.9-4.5 years. (Also saw 498 children without disabilities).	In the sample of children with ASD, PIPS scale score correlated with receptive language (r = .66) and expressive language (r =.61), as well as nonverbal mental age, and gross and fine motor age.	Dutch version of MCDI (N-CDI), or the Reynell Developmental Language Scales (RTOS).	PIPS, the imitation scale being tested. PIPS includes meaningful, non-meaningful, object-directed, body, and facial tasks.	Subtests from Peabody Developmental Motor Scales-2.	18 -19 Vanvuchelen 2011
Ingersoll & Meyer (2011b)	23 children, aged 2-3 years.	Imitation correlated with expressive and receptive language, but these were no longer significant after partialling out developmental level (expressive language and MIS, r = .15; receptive language and MIS, r = .3; expressive language and the UIA, r = .12; receptive language and the UIA, r = .27). The MCDI remained significantly correlated with the MIS (r = .37) and the UIA (r = .41) even when developmental level was partialled out.	Preschool Language Scales (PLS-4), and the MCDI.	Motor Imitation Scales and an unstructured imitation assessment (UIA). All used actions on objects.	None.	20 -25 Ingersoll 2011b
Ingersoll & Meyer (2011a)	27 children, mean age of 38.7 months.	No correlations with language for object-based imitation (with PLS-4 r = .21; with MCDI r = .31) or gesture imitation (with PLS-4 r = .03; with MCDI r = .32). Composite imitation scores correlated with the CDI (r =.36), but not with PLS-4 (r = .14). Nonverbal IQ was partialled out of these correlations.	Preschool Language Scales (PLS-4), MCDI.	Motor Imitation Scales.	None.	26 -31 Ingersoll 2011

Table C.1: Studies reporting correlations between language and imitation in children with an ASD						
Paper	**Participants**	**Results**	**Language tasks**	**Imitation tasks**	**Motor tasks**	**Forest Plot Reference**
McDuffie, Yoder & Stone (2005)	29 infants with autism.	Commenting was the only significant predictor of comprehension after controlling for cognitive delay. Both commenting and motor imitation were significant predictors of production. Motor imitation without objects at Time 1 correlated with MCDI comprehension at Time 2 (r=.38), and production (r=.59).	MCDI	MIS	Raw score for the non-imitative fine motor items of the Mullens Scales of Early Learning.	32 -33 McDuffie 2005
Sallows & Graupner (2005)	24 infants with autism, 24-42 months, with follow up after 1 and 3 years of treatment	The strongest pre-treatment predictors of outcome were imitation, language, daily living skills, and socialization. Nonverbal imitation on the ELM correlated with language at 3 year follow up (r = .69).	Reynell Developmental Language Scales and the CELF.	Imitation learning as part of Early Learning Measure. The ELM included 10 each of verbal imitation, nonverbal imitation, following verbal instructions, and expressive object labelling.	Motor Scale of Vineland Adaptive Behaviour Scales.	34 Sallows 2005
Strid, Heimann & Tjus (2013)	20 ASD, 23 TDs. The ASD group were split by speaking and non-speaking for some analysis.	Both the speaking and non-speaking autism groups showed reduced deferred imitation. The correlation between deferred imitation and language age was not significant (r = .10). Significant correlation between deferred imitation and mental age for the autism group.	Peabody Picture Vocabulary Test (PPVT – third edition; Dunn & Dunn, 1997) and a Swedish version of the MCDI (which ever was most appropriate for the language level of the child).	Deferred object-based imitation. The objects were replicas of the ones originally used by Meltzoff (1985; 1988).	None.	35 Strid 2013

Appendix D: Nonword lists used in verbal imitation tasks.

Table D.1 lists the three nonword lists used in the Spy Password games. This includes the original list of words taken from the CNRep (Children's Test of Nonword Repetition) (Gathercole, Willis, Baddeley, & Emslie, 1994).

Table D.1: Nonword Lists

CNRep		List 1		List 2	
dɒpɛleɪt	prɪstɜrækʃʌnəl	tʃɪməlɑd	drɪntɜrækʃənəl	pɪgɜras	skantɜrakʃənəl
glɪstɜrɪŋ	ʌndɜbræntjuænd	grantɜrin	ændɪpræktjuɛst	brastɜrɪg	ɛntoustɪntjuand
pɛnəl	trʌmpʌtin	bænəl	klæstɜroun	tɛnəl	ʃrɪndoubat
difɜmɪkeɪʃən	slædɪŋ	tebɒlɪvæʃən	brufɪd	dɒsɒdɪleɪʃʌn	grɒlɒp
kɑntræmpɑnɪst	kɒmisɪteɪt	madrɪntoʊmɛnt	mægunɜrɪs	saldɪŋkɜrand	rɪpɪdɜrɪl
hæmpʌnt	tælfæst	dʒɛmtɛst	bʌznælt	fɪnpɒnd	lɒglɛst
riʌtɜpeɪʃʌn	lɒdənɛɪpɪʃ	gɑdɒlɛmeɪʃən	tɪfɪkeɪlɪʃ	mɪlɑndɪkaɪʃən	tɪbɪralʌm
pɜplɪstɜroŋk	bæræzən	rɒndɒntɜrɛst	ʃɒnɑtɪv	gandafrɪlɪsk	fɪloubɒt
blɑntɜsteɪpiŋ	kɒmɜrin	spɒldɜstɜnɛs	sɒrəleɪn	flastɜstɪgəm	havɪdɛt
sɛprɪteniʌl	ɛmplɪfɔvənt	hɑpfɔdɪkɔrɑn	tʌmptoʊbɛlænd	pɑgratɒmɜrɪŋ	sɛndrounapɛst
dɛtrʌtʌpɪlɪk	θɪkɜri	hɑnsɛridoʊmən	noʊpɜfoʊ	mɪsgatʃɜreθɪk	kɪdɜbau
glɪstaʊ	vɒltjulærɪti	spɒndi	gæntʌfæsɜri	drɪmli	bɒndɜlɑikasi
frɛskoʊvɛnt	vɜsʌtræʃənɪst	prɪntʌpɑst	kædɜtrɪʃənɪst	slafrɒnɒmp	fɛbɜtrouʃənɪst
bænɪfɜ	rubɪd	rɪdɛmeɪ	jɑsɛθ	ʃakɒnɜ	wɜdɛk
stɒpoʊgrætɪk	bræstɜrɜ	slɒdoʊkrɑdɪk	klæmpɜroʊ	gratɜfrɪgɪn	slɑntɜrou
wugʌlæmɪk	dɪlɜ	dɒfɪdeɪʃəs	θɪbi	gɛroubeɪlənt	mɪpeɪ
bælʌp	pɛnɛrɪfəl	nɑmɪt	fɪnɒgɜfʌl	velɒg	gɛlɛdɜfʌl
kʌnfræntjuʌli	bænaʊ	mɪstræntɒmɛli	dʒɛgu	lɪstrɒntamɒli	ralɔ
fɛnɜraɪzɜ	prɪndəl	fɒzɛlɒti	frændəl	yɒkoudeɪli	stɪmbəl
æltjupeɪtɔri	skɪtɪkʌlt	ɛstɜrɛloʊvɜ	blʊnɜrɛst	ɒspamɒnɪti	tragapɒnd

Appendix E: Adult Nonword Repetition Pilot Analyses

Two additional nonword repetition lists were constructed to match the nonwords in the CNRep. To check baseline performance on the 3 lists did not differ, 22 typical adults were tested on the three nonword lists in a nonword repetition paradigm. All the lists were tested in one session, and not in the context of the spy password game – adults heard the words from the laptop and repeated them to laptop. Because it was anticipated that most adults would be at ceiling on this task, pink noise was added to the recordings to make the task harder and ensure performance was not at ceiling.

Repetitions were coded correct or incorrect, rather than using the LD algorithm to code phonemic distance between the stimulus and responses. The mean scores for the different syllable lengths and lists was analysed using a two-way ANOVA, with list and syllable entered as factors. There was no significant difference in overall performance between the 3 lists, and no significant interaction effects between list and syllable (indicating that lists were matched overall, and the 4 syllable lengths within each list were also sufficiently matched). Figure E.1 shows the performance of the participants, and Table E.1 the mean scores.

Table E.1: Adult nonword repetition performance

	Nonword Lists		
Word Length	*List2*	*List2*	*List3*
Two Syllables	0.69 (0.16)	0.70 (0.21)	0.70 (0.17)
Three Syllables	0.57 (0.12)	0.61 (0.16)	0.63 (0.17)
Four Syllables	0.45 (0.15)	0.50 (0.17)	0.40 (0.18)
Five Syllables	0.53 (0.19)	0.44 (0.15)	0.43 (0.16)

Table E.1: *Mean scores for the different word lists and word lengths. Numbers in parentheses represent standard deviation.*

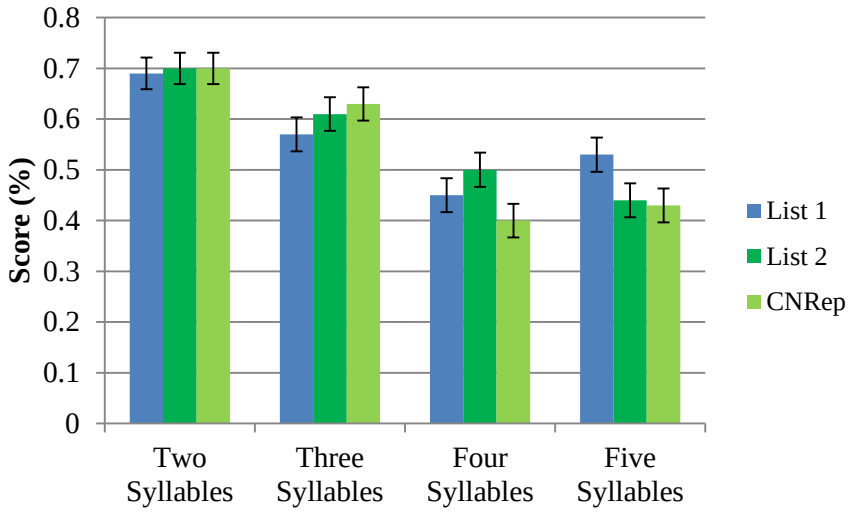


Figure E.1: Mean score on the different nonword lists and word lengths from adult pilot data. *Error bars represent standard error.*

Appendix F: Power analysis for registered report

(Submitted during registration of report.)

Previous effect sizes have not been routinely reported in mu suppression literature. This makes selecting an adequate effect size for the basis of a power analysis difficult. However, if mu suppression is a valid biomarker for social cognitive processes, or indeed autism, (Cheng, Tzeng, Decety, lmada, & Hsieh, 2006; Perkins, Stokes, McGillivray, & Bittar, 2010), and a potential target for neurofeedback therapy (Pineda, Juavinett, & Datko, 2012), then it is reasonable to expect a large effect size. In this context, our selection of a medium effect size for mu suppression is conservative: if we do not observe the phenomenon under this assumption, its use in clinical contexts is undermined.

A priori analyses were conducted using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). A power analysis was carried out to determine the sample size necessary to detect medium effects (f = 0.25, as outlined by Cohen, 1988), with power of 0.9. As repeated measures ANOVAs typically violate the sphericity assumption, our power analysis considered corrections for sphericity, and was based on the most conservative correction for nonsphericity. This is when the nonsphericity correction (denoted ε) is equal to 1/1-*m*, where *m* signifies the number of measurements. The anticipated correlation among the repeated measures was set at 0.6. In previous work, the correlation between occipital and central electrodes has been moderate (see Tangwiriyasakul et al., 2013), and we anticipate that the correlation between different conditions will be moderate to high. We feel this estimation of correlation is justified, especially as the analysis was based on the most conservative nonsphericity corrections possible.

Power analyses demonstrated that to detect a medium effect of condition (videos with hand and object, hand without object and kaleidoscope patterns, and the own movement

condition) 52 participants are needed. To detect a medium effect of site (occipital versus central) 36 participants are needed. To detect an interaction between these two variables, 61 participants are needed (note that many previous studies have used much smaller sample sizes: (Fan et al., 2010; Muthukumaraswamy & Johnson, 2004; Muthukumaraswamy et al., 2006; Oberman et al., 2008). The power of paired t-tests to investigate significant effects arising from the ANOVA was also computed; for two-tailed tests, with power of 0.9, 44 participants are required to detect a medium effect (d = 0.5,Cohen, 1988). The predicted power of these statistical tests with the intended sample size of 61 participants is displayed in Table F.1. Note that as analyses with the trial-by-trial baseline do not include the own movement condition, which makes the ANOVA for this baseline slightly different. Additional analyses have been conducted for this ANOVA and are detailed in Table F.1.

We planned that if data had to be discarded for a participant due too many rejected trials (see the previous section Analysis Plan for details on how trials will be rejected and how much data will be allowed to be discarded before a participant is removed from the analysis), further participants would be recruited to ensure the analysis is sufficiently powered. Participants who do not perform the movements during the own movement condition will be reminded by the experimenter to do so, but if the participant continues to not perform the movements they will be excluded from the analysis. If participants need to discontinue the EEG (in the case of sickness) their incomplete recording will be removed from the analysis. A complete record of participants who had to leave the experiment prematurely, or who failed to complete the movement condition, or whose EEG did not produce a sufficient number of good trials, will be kept and the final number of participants that had to be excluded will be reported.

Table F.1 Power of planned statistical tests

Proposed statistical test	**Power**
Repeated measures ANOVA - main effect of condition	0.94
Repeated measures ANOVA for trial-by-trial baseline- main effect of condition	0.96
Repeated measures ANOVA - main effect of site	0.99
Repeated measures ANOVA - interaction between site and condition	0.90
Repeated measures ANOVA for trial-by-trial baseline - interaction between site and condition	0.91
Paired t-tests (two-tailed)	0.97

Table F.1: *The power (1-β) of the planned statistical tests. For these calculations, the following parameters were used: N=61, ε = 1/m-1 (where m is the number of measurements. For the main effect of condition, this is 4. For the interaction, this is 8. For the trial-by-trial baseline these are 3 and 6, respectively.), the correlation between measures = 0.6, and standard medium effect sizes (f =.25, d = .5).*

Appendix G: Post-EEG recording questionnaire

Questionnaire

Which type of video did you find most interesting to watch?

Which type of video did you find least interesting to watch?

On a scale of 1-5, please rate how engaged you felt you were while watching the blank videos. A score of "1" represents no engagement, while a score of "5" represents a lot of engagement.

On a scale of 1-5, please rate how engaged you felt you were while watching the videos where hands interacted with objects. A score of "1" represents no engagement, while a score of "5" represents a lot of engagement.

On a scale of 1-5, please rate how engaged you felt you were while watching videos where hands interacted without objects. A score of "1" represents no engagement, while a score of "5" represents a lot of engagement.

On a scale of 1-5, please rate how engaged you felt you were while watching the videos of kaleidoscope patterns . A score of "1" represents no engagement, while a score of "5" represents a lot of engagement.

Could you imagine yourself performing the content of the videos? Please rate how hard you think it would be to perform the videos for each video type. A score of 1 would mean that the videos were very easy to perform, while a score of 5 means the videos are hard to perform.

Videos of the hand and the pencil:

Videos of just the hand:

Videos of the kaleidoscope patterns:

Did you notice anything other than the videos of hands or kaleidoscope patterns while you were doing the experiment? If you did, what did you notice?

How many times did you notice these extra stimuli?

Appendix H: Supplementary unregistered analyses for mu suppression registered report

Responses to post-recording questionnaire

Chi-square tests showed significant effects of condition when participants selected their most interesting (χ^2 (3) = 61.82, p <.001) and least interesting (χ^2 (3) =96.05, p <.001) condition. One-sample z tests comparing these proportions show that the kaleidoscope videos were rated as most engaging significantly more than hand-object videos (z = -6.47, p <.001), and hand-object videos significantly more than hand-no object videos (z = 6.21, p <.001). Hand-no object videos were selected as least interesting significantly more than hand-object videos (z = -2.53, p =.006), and rest more than hand-no object videos (z = 13.72, p <.001).

One-sample z tests also showed that the difference in participant's reports of feeling able to imitate the content of the videos significantly differed between kaleidoscope and hand-object videos (z = 20.49, p <.001) and hand-no object videos (z = 20.89, p <.001).

Gender effects in mu suppression

Independent t-tests were conducted comparing males (N=19) and females (N=42) on changes in activity at the central regions during the hand videos. No significant differences were found, for ether video type, for any baseline technique.