Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Humanities
Department of Language and Literature

Camilla E. Crawshaw

The Critical Role of Language in Theory of Mind Development

Viewed through the Lens of Autism and Parallel Deficits across Other Atypical Populations

Master's thesis in English Linguistics and Language Acquisition (ENG3920)

Trondheim, May 2018

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

Abstract

Theory of mind (ToM) is a vitally important area for study not only within psycholinguistics but also across neuroscience and other diverse fields, because of its implications for the study of social cognition, and the diagnosis and treatment of developmental and clinical disorders. While this theoretical construct has existed within the literature for some time, a definitive explanation for its emergence, development, and acquisition has yet to be reached. This thesis conducts a qualitative, critical review of the theoretical and research literature with the aim of establishing the role of language as a critical facilitator in the development of a ToM. The review begins with an analysis of the theoretical and historical background of ToM, examining the various accounts that have been proposed to explain it, including executive function (EF). It then turns to a contemplation of the developmental perspective on ToM, and its investigation within the research literature. Building upon and contrasting with the findings of the first main chapter (§2), the second (§3) presents and evaluates various language accounts of ToM, with reference to the evidence from language-impaired populations, concluding that language acquisition is necessary on a number of levels for ToM development. The final main chapter (§4) investigates this relationship in the context of autism, a developmental disorder where deficits in both ToM and language ability have been found, inferring that these are causally related. The thesis concludes by arguing that the relationships between ToM and language, and ToM and EF, are more dynamic than have previously been considered. Finally, it proposes that these relationships can be better understood within the framework of a tripartite model, representing the co-developmental, multidirectional interactions between each component. These findings are significant for future work not only on ToM and social cognition, but also with atypically developing and clinical populations.

Acknowledgements

Firstly, I would sincerely like to thank my supervisor Professor Mila Dimitrova Vulchanova for all the advice, guidance, and feedback that she has given me throughout the process of writing this thesis. I am very grateful for all of this time that she has made available to me.

Secondly, a special thank you goes out to my MPhil pals. All of those hours of coffee, cake, lunches, and work sessions in the lesesal and library have sustained and supported me through the researching and writing of this thesis. (Let's hope that the 'curse' of the MPhil laptops has now been lifted…)

Finally, I would like to express my profound gratitude and appreciation to my friends, family, and of course Gareth, for all of their support and encouragement not only during the course of this Master programme but all of my years of study. Thank you for making this accomplishment possible.

So long Norway, and thanks for all the fish!

Table of Contents

§1 Introduction

This thesis will aim to establish that language plays a critical role in the development and acquisition of theory of mind (ToM) by examining language accounts and comparing them to other accounts of ToM, such as those concerning the role of executive function (EF). This theoretical debate will also be considered in the context of autism as a domain of interest, and with reference to parallel deficits across other atypically developing populations. ToM is a critically important area of study, because it spans the diverse fields of psychology, linguistics, sociology, education, neuroscience, medicine, biology, and genetics. Achieving an explanation of ToM development, regarding its underlying cognitive structures and neural networks, related processes, and what causes or mitigates impairment, will have ramifications for future work within each of these research fields. A number of individuals in atypically developing populations experience various degrees of delay or impairment in their ToM ability, which can have an impact upon their prospective quality of life, so an explanation of ToM also has vital theoretical and practical implications for the development of effective clinical intervention strategies.

The thesis is divided into three main chapters: §2 ToM; §3 ToM and Language; and §4 Domain of Interest: Autism. The first main chapter (§2) functions as a backdrop to the following chapters: introducing the concept of ToM, considering the non-linguistic accounts proposed to explain it, and clarifying both its historical and current statuses within the theoretical and research literature. The second main chapter (§3) presents the language account of ToM, exploring and analysing the perspectives advanced by theorists working on different areas of language. It will also introduce the support for both meta-representational and general-language developmental accounts of ToM by referring to work from the research literature with the specific language impairment (SLI) and late-signing deaf populations. The final main chapter (§4) builds upon the language account of ToM introduced in §3, presenting autism as a valuable population within which to investigate this account, and evaluating the theoretical perspectives on autism, ToM, language, and EF within the context of the findings from the research literature. The critical discussion in §5 and conclusions in §6 will draw this thesis to a close, summarising the state of the research and arguments put forward throughout the main chapters, as well as discussing and highlighting the directions and areas that will be significant for future research.

§2 Theory of Mind (ToM)

2.1 What is Theory of Mind?

The ability to impute mental states (such as thoughts, emotions, intentions, knowledge, beliefs, and desires) to oneself and others, as well as interpret, explain, predict, and manipulate behaviour on the basis of this attribution, is essential to human social interaction and vital for successful participation within society; this faculty has come to be known as 'theory of mind' (ToM) and is an integral part of socio-cognitive functioning, closely intertwined with language, executive function, and memory [Astington & Baird, 2005; Bradford et al., 2015; Doherty, 2008; Frith & Frith, 2005; Griffin and Baron-Cohen, 2002; Frank et al., 2015; Meert et al., 2017; Premack & Woodruff, 1978; Sodian & Kristen, 2010]. When first presenting the concept of ToM, Premack and Woodruff employed the term 'theory' because they viewed the process of attributing unobservable mental states as the building of 'a system of inferences', which could then be utilised to predict the behaviour of other entities [1978, p.515; Doherty, 2008; Perner, 1999]. There have been a number of theories concerning the development of ToM, traditionally 'modular theory' (2.2.1), 'simulation theory' (2.2.2), and 'theory theory' (2.2.3), but also more wide-ranging accounts placing executive function and neural systems (2.2.4), or language (§3) at the core of ToM development, each of which will be discussed in further detail below. Since its emergence in the 1970s and through subsequent, growing popularity, the term ToM has been repurposed within theoretical literature both more broadly as an umbrella term for cognitive empathy, perspective-taking, or 'folk psychology', as well as more specifically for 'mentalising', 'mind-reading', or having an 'intentional stance' [Astington & Baird, 2005; Doherty, 2008; Frank et al., 2015; Frith & Frith, 2005; Griffin & Baron-Cohen, 2002; Meert et al., 2017]. ToM has also been used to refer to a hypothetical, domain-specific, cognitive structure facilitating mentalising abilities, the area of research exploring the development of said abilities, and the theoretical explanations for this development [Astington & Baird, 2005, p.4]. As a consequence, this conflation of distinct but interrelated concepts has proved problematic for assessing the outcomes of research within the context of the literature, as well as for the refinement of the theory itself.

Belief and desire are arguably the core mental states, because a person's understanding of reality and their preferred version of it will govern how they feel and act [Doherty, 2008]. Possessing a ToM enables a person to acquire knowledge about the beliefs and desires of other people, because they can recognise that the knowledge and experiences of another person are different to their own [Frith & Frith, 2005]. The advantage of this ability is that a person can explain and predict people's behaviour on the basis of this knowledge, understanding that people's behaviour is determined by their beliefs about reality and not reality itself [Doherty, 2008; Frith & Frith, 2005]. Being able to grasp the concept of a conflict between reality and belief is what allows a person to manipulate other people's beliefs and behaviour by making them aware of new information, concealing this information, or disseminating false information [Doherty, 2008; Frith & Frith, 2005]. In these ways, a ToM facilitates not only activities such as teaching but also tactical or deliberate deception (e.g. cheating), which may be one of the earliest indicators of a developing ToM [Doherty, 2008; Frith & Frith, 2005]. Deception is used to manipulate both beliefs and behaviour, but it does not signify a direct representation of ToM ability, because if a child merely wants to influence someone's behaviour, then it cannot be assumed that they understand the concept of belief [Doherty, 2008]. However, deception is important in ToM research because of its association with false belief (see 2.3.1). Perner argued that consciousness required some level of ToM, because it concerned not only one's awareness of what occurs in the world but also how one is connected to the world, including how experiencing events leads to knowledge and memory of these events being encoded, and intending to act causes action [Perner, 1999]. In this way, ToM development is necessary for self-referential awareness and to enable episodic memory and executive control [Perner, 1999]. Across the literature, it is generally believed that people's perceptions of belief and desire correspond to propositional attitudes, which are situations or states of affairs that may be true or false [Doherty, 2008, p.36-37]. Propositional attitudes are separated into four components: (1) the agent (the person possessing the belief or desire); (2) the proposition (the thing believed or desired); (3) the attitude (the type of mental state involved); and (4) the anchor (the part of reality that the proposition corresponds to) [Doherty, 2008, p.36]. The proposition relates to how a person represents the world and the attitude to how a person uses this representation (see 2.2.3 for more on representations) [Perner, 1999]. Importantly, the matter of whether a propositional attitude is true or false is determined by its accuracy in describing the mental state of the person in question and not the objective state of reality [Doherty, 2008].

2.2 The Theory behind ToM

Initially, the principal theories concerning the development and acquisition of ToM consisted of modular theory (2.2.1), simulation theory (2.2.2), and theory theory (2.2.3), with still further diverging accounts within each main paradigm. Thinking in the literature has subsequently evolved to span other accounts involving executive function (EF), neural structures and pathways (2.24), and language (§3). The following subsections will discuss the leading proponents, ideas, and constructs of each early theory, as well as EF accounts, before returning to a critical analysis of their instantiation in the research literature in 2.3. Although this thesis seeks to establish the role of language in ToM acquisition, it is essential to consider the antecedent and competing theoretical accounts, because many do not preclude language from being critically involved in the ToM developmental process. It is important to evaluate the state of the research literature concerning the accounts that do not place a focus on the role of language, before contemplating how and where language accounts may fit in. The role of language will be returned to in §3.

2.2.1 Modular Theory

Modular theory postulates a specialised piece of neural architecture or a 'mental module' dedicated to ToM processing and reasoning [Doherty, 2008; Mahy et al., 2014]. According to modular theorists [Fodor, 1983; (Baron-Cohen, 1995; Leslie, 1994; Scholl & Leslie, 1999) cited in Sodian & Kristen, 2010; (Baron-Cohen, 1998; Leslie et al., 2004) cited in Mahy et al., 2014], ToM (and language) acquisition transpires through neurological processes and the maturation of a system of biologically innate, domain-specific and modular mechanisms [Mahy et al., 2014; Sodian & Kristen, 2010]. Such theorists also hold that while experience and environmental factors may play a role in activating the operation of, or influencing the expression of, said mechanism, they do not alter its basic nature, which is likely to have been shaped through the course of evolution [Doherty, 2008; Mahy et al., 2014; Sodian & Kristen, 2010]. Fodorian modules were proposed to be encapsulated as well as: (1) to process specific inputs within a specialised system; (2) to operate mandatorily and spontaneously; (3) to be fast-performing as a possible consequence of encapsulation and mandatory operation; (4) to restrict output to 'low-level concepts'; (5) to exhibit specific and representative failure patterns; and (6) to develop in a characteristic pace and sequence [Doherty, 2008, p.50-51; Fodor, 1983; Mahy et al., 2014]. Perhaps the most prominent account of ToM modular theory has been presented by Leslie and colleagues [Doherty, 2008; Leslie, 1994; Leslie et al., 2004; Mahy et al., 2014]. Leslie proposed that a specialised theory of mind mechanism (ToMM) formed the 'specific innate basis' for our ability to acquire a theory of mind, damages to which were responsible for apparently characteristic ToM deficits in autism (§4), as per the Fodorian modular principle (5) [1994, p.214]. An inhibitory selection processor (SP), proposed to aid the ToMM, was claimed to be in operation by the second year of life, and Leslie et al. went on to hypothesise that early ToM was a 'modular–heuristic process' of domain-specific learning, where the ToMM-SP system became increasingly effective at coping with the executive demands of ToM reasoning [2004; Leslie, 1994; Mahy et al., 2014].

The appeal of this theory is that it could offer an explanation for the relatively culturally uniform development of ToM, as well as an account of the developmental deficits in autism (see §4) but subsequent data has suggested it is perhaps too simplistic and requires reassessment [Doherty, 2008, p.50; Mahy et al., 2014]. Some studies have suggested an early ability to attribute false beliefs about location and identity, as well as false perceptions among infants of two years of age [Baillargeon et al., 2010], which would be consistent with the proposal of an early-developing ToMM [Mahy et al., 2014]. This may, however, be more indicative of the two-step developmental sequence of desire- and belief-understanding at preschool age that has been found in research on ToM acquisition [Sodian & Kristen, 2010] (more in 2.3). Furthermore, much of Leslie's work on the ToMM-SP system is grounded in the concepts of pretence and pretend play, which, problematically, have recently been shown to engage distinct neural substrates to those engaged by false belief [Meinhardt et al., 2012], an important measure of ToM capability (see 2.3). Given the strict criteria for Fodorian modules, Mahy et al. argued that Leslie and colleagues' accounts of a ToMM would thus be the most easily falsifiable of ToM theories [2014, p.70]. Neurological research has revealed evidence for a common neural network implicated in mental state processing and reasoning, including areas such as the cortical midline structures (CMS): comprising the medial prefrontal cortex (MPFC), adjacent rostral anterior cingulate cortex (raCC), and medial posterior parietal cortices (MPPC), such as the posterior cingulate and precuneus; and the bilateral temporal parietal junction (TPJ), as well as the superior temporal sulcus (STS), and orbitofrontal-amygdala-temporal circuit [Bradford et al., 2015; Fine et al., 2001; Frank et al., 2015; Griffin & Baron-Cohen, 2002; Happé & Frith, 2014; Mahy et al., 2014; for further citations; Schurz et al., 2015] (more in 2.2). However, while CMS areas have been consistently implicated in ToM research, they have also been implicated in research on other executive functions and abilities, and thus cannot qualify for the ToMM under Fodorian modular principle (1) [Mahy et al., 2014, p.70]. The TPJ presented a potentially more plausible candidate for domain-specificity, but did not appear to qualify under Fodorian modular principle (6) because it appeared to be differentially recruited over age groups, as well as being recruited for non-ToM reasoning tasks even among adults [Mahy et al., 2014].

2.2.2 Simulation Theory

Simulation theory holds that mental state attribution is facilitated through introspection, projection, and reflection: using one's own psychological states and experiences to build a working model of another's mind and perspective, and implementing role-playing processes to predict their resulting behaviour [Doherty, 2008; Mahy et al., 2014; Perner, 1999; Sodian & Kristen, 2010]. Simulation theories arguably originate from Cartesian ideas of mental transparency: that we are directly capable of knowing and reasoning about our own minds through free introspection, and thus can extrapolate this knowledge to others' by taking and simulating their perspective [Perner, 1999]. These accounts postulate that constructing a 'theory' about another entity's mental states is unnecessarily inefficient when we can already infer from our own working model [Doherty, 2008]. However, they rely on the fact that people tend to think and respond in similar manners [Doherty, 2008], which may not be strictly true, given work on different cognitive styles (see 'weak central coherence', 4.4.1). Simulation theorists take a social-constructivist approach, placing more emphasis on the role of social experience in the development of ToM than cognitive theories such as 'theory theory' (see 2.2.3) [Dunn & Brophy, 2005; Harris, 1992, 2005; Sodian & Kristen, 2010]. Such theorists believe that participation in role-play activities – such as pretence – improves children's mentalising abilities and that this is facilitated through social interaction and conversation by establishing joint attention, where dyadic mother-child interactions followed by triadic mother-child-object interactions steadily foster ToM capacity [Astington & Baird, 2005; Dunn & Brophy, 2005; Farrant et al., 2006; Harris, 1992, 2005; Sodian & Kristen, 2010]. These theories also fit within a language account of ToM; at least to the extent that language functions as the medium for inter-individual communication, but this will be discussed further in 3.1.

Simulation theory is attractive because it slots nicely into our intuitions about socio-affective cognition, empathy, and emotional interaction. It also offers a potential explanation for the research data, which illustrates that as children develop, they become increasingly able to navigate the challenges presented by different tasks [Doherty, 2008]. Harris tried to conceptualise this by proposing that children's mental states could be framed in terms of two default-setting categories, referring to: (1) the current state of reality, and (2) the representation of mental states associated with that reality [Doherty, 2008, p.44; Harris, 1992]. In order to make inferences about another person's mental states, children would then need to alter their default settings to parallel the mental states of the other person; in this view, ToM tasks become more difficult when more sets of default settings require alteration [Doherty, 2008; Farrant et al., 2006; Harris, 1992]. Like Leslie and colleague's work on the ToMM, Harris' account is based on investigation into children engaging in pretence, which cannot necessarily be equated with other measures of ToM such as false belief (see 2.3.1) [Doherty, 2008; Leslie, 1994; Leslie et al., 2004; Meinhardt et al., 2012]. A problem with Harris' proposals and wider simulation theory is that it was assumed that children are able to introspect accurately and are aware of their own beliefs, desires, and emotions, something that does not appear to hold up against the research data [Doherty, 2008]. Gopnik & Astington used the unexpected contents false belief task (designed by Hogrefe, Wimmer & Perner, 1986; see 2.3.1) in their (1988) study, where, despite having recently experienced holding a false belief, children were unable to remember what they had previously thought to be inside the container once the contents had been revealed [1988; Doherty, 2008]. This work suggested that simulation theorists should not assume that we have direct and transparent access to our own minds, although the findings might also reflect working memory constraints. Empathy and emotion might present a plausible case for simulation theory, given that one can respond emotionally to purely imagined situations and could therefore simulate other people's emotions by imagining oneself in their position [Doherty, 2008]. However, some theoretical knowledge seems necessary for any form of simulation to be initiated (even just the construct that someone will respond in the same way as oneself), and if one mistakenly attributes knowledge vs. ignorance (and vice-versa) to another, then it could lead to misjudging their state of mind or emotion [Doherty, 2008]. Recent work has suggested that empathy and emotional or affective perspective taking may also require some level of ToM to function, and even that they are all part of a broader perspective-taking (PT) system [Bird & Viding, 2014; Farrant, 2015]. Furthermore, there are a number of social psychology studies (e.g. Milgram's 1963 work on obedience cited in Doherty, 2008) that depict participants behaving very differently in practice to how they themselves, as well as contemporary psychologists, would have predicted or simulated within a hypothetical thought experiment [Doherty, 2008].

Simulation theories could still fit within the research literature's picture of a protracted developmental trajectory of ToM, with the gradual deployment of increasingly sophisticated simulation processes over time, in addition to studies connecting children's capacities for ToM and imagination [Mahy et al., 2014, p.71]. In order to support simulation theory accounts, however, Mahy et al. reasoned that such links would require evidence from neuroscientific studies demonstrating that these processes, as well as processes involved in mapping representations of 'self' to 'other', share a common underlying neural system [2014]. They also argued that longitudinal neurological studies would need to establish a developmental timeline showing that implicated neural systems become more efficient and automatic over time [Mahy et al., 2014]. As mentioned in 2.2.1, the CMS system, incorporating the MPFC and medial precuneus (MP), has been implicated in neurological research regarding ToM and is thought to be involved in intentional and evaluative processes mapping representations of 'self' to 'other', because individuals appear to activate similar brain areas when considering both the self and others [Mahy et al., 2014: for further citations]. Studies have found activation of the CMS' MPFC and MP in adult and adolescent self-reflection and belief-reasoning tasks, which might indicate an involvement in simulation processes via self-perception and perspective-taking (2.3.3) [Mahy et al., 2014]. Some studies have also found a change in recruitment of the dorsal and ventral MPFC from child to adult when simulating and reasoning about similar and dissimilar others, suggesting that adults may have developed specialised simulation processes [Mahy et al., 2014]. The mirror neuron system (MNS) was also identified as another candidate for a simulation process network, because it had been shown to be co-activated for the actions, intentions, and emotions of the self and other in both adults and children during the perceiving and carrying out of identical actions [Mahy et al., 2014, p.72; Lamm & Majdandžić, 2015]. It was proposed that the MNS is made up of regions in the inferior frontal gyrus (IFG) comprising the pars opercularis and adjacent ventral premotor cortex, in addition to the rostral inferior parietal lobule [Mahy et al., 2014, p.72]. Based on research on the MNS in both human and non-human primates, as well as infant parsing of intentional action, some authors believed that the MNS could enable mental state simulation by facilitating a direct mapping of goals, intentions, and actions from 'other' to 'self' [Mahy et al., 2014, p.72; Lamm & Majdandžić, 2015]. It was thought that this sort of mapping might then aid the child's acquisition process of concepts like intention, desire, and belief [Mahy et al., 2014; Lamm & Majdandžić, 2015]. However, these ideas have met with significant criticism in more recent work, which has suggested that the MNS' responses are limited to physical, manual, and motor actions, and may even be restricted to the hand areas [Lamm & Majdandžić, 2015]. It is also likely that mental state representations can exist separately to simulations of the other, but this will be considered further in the following section.

2.2.3 Theory Theory

Theory theorists propose that our understanding of the mind is theory-like, constructed from a framework of interrelated representations within the brain, and defined by their relationships with one another and to reality [Doherty, 2008; Frith & Frith, 2005; Mahy et al., 2014, Perner, 1999, Sodian & Kristen, 2010]. According to this view, children build their ToM by gathering evidence concerning the relationship between actions and mental states in the same way that a scientist would use data to inform their theory [Mahy et al., 2014]. Their representations correspond to mental states, concepts, and propositional attitudes, which are causally linked and radically updated as both an outcome and instigator of conceptual change and development [Doherty, 2008; Frith & Frith, 2005; Mahy et al., 2014]. In this sense, theory theory emphasises the role of a child's personal experiences in concept formation [Mahy et al., 2014; Sodian & Kristen, 2010], and could fit within language accounts of ToM.

Of these, Perner's account is one of the most prominent; he argued that children acquired a representational understanding of mind (RUM) by 4 years-of-age (around the time they begin to pass false belief tests: see 2.3), where they could frame mental states by means of propositional attitudes, prior to which they could only be credited with a theory of behaviour [Doherty, 2008; Perner, 1991]. Perner posited that propositional attitudes did not, however, characterise higher order representations, distinguishing between different levels of representation in his three-stage model: primary, secondary, and meta-representation [1991; 1999]. From birth to 18 months-of-age, children have been credited with a 'single veridical updating' model of the world, where their primary goal is to develop authentic representations of the objects and entities in their environment; these are their primary representations [Doherty, 2008, p.37; Stanovich & Toplak, 2012; Whiten & Suddendorf, 2001]. After 18 months, children overcome their difficulties with concepts like invisible displacement and can begin to represent non-present situations: these multiple models can reflect past, possible-future, desired, and pretend or hypothetical situations, corresponding to secondary representations [Doherty, 2008; Stanovich & Toplak, 2012; Whiten & Suddendorf, 2001]. Being able to accommodate multiple models is necessary to comprehend anything beyond the present scenario, and crucial for desire, memory, planning, and language (§3) [Doherty, 2008]. From 3-4 years-of-age onwards, children start to become aware that these models correspond to mental representations; this forms the basis of meta-representation [Doherty, 2008; Stanovich & Toplak, 2012; Whiten & Suddendorf, 2001]. Perner asserted that a person's knowledge about the physical world consisted of representations of reality, whereas their knowledge of the mind extended to comprehension of how these representations of the world are used by other entities [1999, p.4]. From this perspective, forming meta-representations involves understanding that representations can be decoupled from reality, and is necessary to understand belief [Doherty, 2008; Frith & Frith, 2005; Stanovich & Toplak, 2012; Whiten & Suddendorf, 2001]. To theory theorists, the distinction between employing secondary representations and meta-representation parallels the divergence of pretence and belief [Doherty, 2008]. In order to engage in and understand pretence, the child requires access to multiple models [Lillard, 1993] but only to the extent of secondary representations, as the child merely needs to be aware that everyone correctly evaluates a false proposition in the same way [Doherty, 2008]. Whereas to be able to understand false belief, the child must access meta-representations to recognise that another person is evaluating a certain proposition as true when they themself evaluate it as false, as well as appreciating the cascade effects on the person's behaviour [Doherty, 2008, p.40]. This distinction also facilitates the appreciation of different perspectives (see 2.3.3) [Doherty, 2008]. Theory theory predicts a similar developmental trajectory for understanding one's own and other people's minds, whereas simulation theory (2.2.2) predicts difficulties with understanding other people's minds but not one's own [Perner, 1999]. Here, evidence has supported the theory-theory account, suggesting that children can predict other people's false beliefs around the time that they can remember their own [Gopnik & Astington, 1988; Perner, 1999]. This also suggests that the phenomenon of experiencing difficulty with remembering one's own false beliefs is not due to working memory constraints.

Developmental psychologists have often suggested that theory theory provides the best explanation for existing behavioural data, especially the gradual process of grasping increasing conceptual complexity through experience [Astington & Baird, 2005; Doherty, 2008; Mahy et al., 2014, p.74 for further citations]. Since theory theory's rationale positions conceptual change as the driving force of ToM development, and it is unclear how conceptual change manifests within the brain, Mahy et al. suggested that theory theory could be the hardest ToM account to evaluate within a neuroscientific context [2014, p.75]. They did, however, propose that theory-theory accounts could be substantiated if evidence were found to suggest: (a) neural networks or structures sustaining domain-specific conceptual structures; or (b) neural changes correlating with conceptual changes during development [Mahy et al., 2014, p.75]. Perner highlighted similar weaknesses in the theory-theory account, pointing out that no one had yet specified a clear framework of theoretical laws underlying ToM, as well as that it did not account for self-knowledge [1999]. Mahy et al. reviewed a number of studies, proposing that neurological evidence taken to support modularity theories could also be relevant to theory theory, in particular data implicating the TPJ [2014]. This is because its increasing selectivity is problematic for modularity theory but could easily be accommodated within theory theory as an example of age-related conceptual change [Mahy et al., 2014]. They further suggested that a lack of specialisation or dedicated neural architecture for ToM could also fit within theory theory, as no position is taken against multi-purpose recruitment of any given neural system [Mahy et al., 2014]. To support their analysis, Mahy et al. cited various fMRI and EEG research implicating the TPJ when switching between and integrating internal and external representations, something that could underpin the system of conceptual updates purported to facilitate ToM development [2014]. They also pointed to research indicating that both common and unique neural regions are recruited to support desire and belief reasoning, which offers plausible evidence for theory theory accounts proposing that an understanding of belief develops out of an understanding of desire [Mahy et al., 2014]. Evidence for developmental neural change (either in recruitment or connectivity) that could be associated with progressive ToM understanding is sparse in the research literature, but Mahy et al. submitted white matter connectivity changes, recognised to be cultivated by experience, as a promising area for investigating correlations between neural and conceptual changes [2014]. They also highlighted research from other neuroimaging fields such as ERP work on implicit and explicit memory that could parallel behavioural research indicating both an implicit and explicit level of ToM [Mahy et al., 2014, p.75-76]; this might also be linked to the distinction between secondary- and meta-representation. Another candidate that has been proposed is the default mode network, connecting a number of brain regions that overlap with the MPFC, MP, posterior cingulate and TPJ, which is activated during higher level socio-cognitive tasks, introspection, and self-referential thought [Mahy et al., 2014]. Its connections undergo a lengthy maturation process which would be consistent with theory theory's predictions of an extended conceptual change trajectory [Mahy et al., 2014]. In future work, Mahy et al. argued that it would be necessary to take a 'network-wide approach' involving diverse brain regions as well as to undertake longitudinal research beginning at a very early age, in order to better evaluate the theory-theory account of ToM [2014, p.76]. Theory theory perspectives could also be accommodated within a language account of ToM, given that they do not preclude multi-purpose recruitment of neural networks, systems, or architecture, and that conceptual change could be driven by increasingly sophisticated language acquisition, but this will be returned to in later sections.

2.2.4 Executive Function Accounts

Some accounts propose that children's age-related advances in their higher-order cognitive faculties, their 'executive functions', are responsible for their developing ToM abilities [Carlson & Moses, 2001; Doherty, 2008; Mahy et al., 2014; Sodian & Kristen, 2010]. The umbrella term executive function (EF) encompasses a number of cognitive abilities related to self-control including: the control of actions, thoughts, and attention, as well as planning, decision-making, working memory, switching, inhibition and impulse control [Doherty, 2008; Durrleman & Franck, 2015]. A number of **theoretical stances** have been taken with respect to the relationship between EF & ToM, generally falling into one of four groups: **(1)** EF is necessary and/or sufficient for ToM development; **(2)** ToM is necessary for EF development; **(3)** they have a bidirectional interdependent relationship; and **(4)** ToM development is separate but masked by EF, because ToM tasks also employ EF components [Carlson et al., 2015; Carlson & Moses, 2001; Doherty, 2008; Fine et al., 2001; Mahy et al., 2014]. Despite the theoretical controversy over its nature, it remains clear that a robust and complex relationship does exist: children become more able to regulate their behaviour and think flexibly at the same stage that they begin to attribute mental states and beliefs to other entities [Bradford et al., 2015; Carlson et al., 2015; Doherty, 2008]. Being able to reflect upon one's own mental states might facilitate self-control and regulation, whereas EF might be required in switching attention from reality to focus on one's own mental states, inhibiting one's own perspective to visualise someone else's, and holding all of this in working memory [Carlson et al., 2015; Carlson & Moses, 2001; Doherty, 2008; Mahy et al., 2014]. Self-control is also necessary for managing distractions when deciding upon and attempting to accomplish goals [Doherty, 2008, p.129].[1] Defining EF is, however, problematic (see §5).

Returning to the theoretical stances, let us first consider stance **(4)**: ToM tasks (e.g. those concerning false belief: 2.3.1) do recruit EFs: children must understand that an agent's behaviour is grounded in their mental representations of the world; they need a certain level of working memory to follow a narrative, remember events, compare another's perspective to their own, and make inferences using this information; set-shifting may be necessary to switch between their own and someone else's perspective; and inhibition might be required to block their own representation of reality in order to consider another's [Carlson et al., 2015; Doherty, 2008, p.132; Durrleman & Franck, 2015]. However, Wellman et al.'s (2001) meta-analysis of ToM research indicated that stance (4)'s predictions do not hold up against the data: while ToM tasks do recruit EFs, younger children's ToM performance did not significantly improve when tasks were manipulated to considerably reduce executive demands, and appropriate guessing baselines were established, demonstrating that explanation and prediction of false belief (FB) were equally difficult [Doherty, 2008; Sodian & Kristen, 2010; Wellman et al., 2001]. Bradford et al. cited various research with adults which found that increasing executive demands decreased performance on ToM tasks, as well as that inhibition and processing speed abilities determined ToM performance [2015]. However, this data does not necessarily provide support for stance (4), and could arguably be more representative of stances (1) or (3).

With respect to stance **(1)**, there are three sub-theories: **(i)** EFs are sufficient to support ToM reasoning without requiring other processes; **(ii)** EFs are necessary (but not sufficient): **(a)** for the development of mental-state concepts; or **(b)** to frame the expression of mental-state concepts that children might already have [Carlson et al., 2015; Doherty, 2008; Mahy et al., 2014].

[1] Some accounts further propose a distinction between affective and non-affective EF [Bradford et al., 2015, for further citations], which might bear interesting implications for the determining of the relationship between EF and ToM.

In order to support theory **(1i)** that EFs are sufficient to account for ToM, Mahy et al. argued that neuroimaging studies would need to show that ToM tasks activated neural regions associated with EF but did not activate additional ToM-specific regions [2014]. To support the theories **(1ii.a)** or **(1ii.b)** that EF is necessary but not sufficient for ToM, Mahy et al. stated that ToM tasks should be shown by neuroimaging studies to consistently activate EF regions but could further activate ToM-specific regions [2014]. Here, the available evidence suggests that accounts proposing EF to be necessary for ToM (1ii) (and potentially that there is a bidirectional/interdependent relationship (3)) are more plausible than those proposing it to be sufficient (1i) [Bradford et al., 2015; Carlson et al., 2015; Carlson & Moses, 2001; Mahy et al., 2014]. In some ways, these theories make similar claims to those of modular theorists (2.2.1); it may be that ToM abilities are supported by a modular structure with separate components recruited for different mentalising processes such as distinguishing between 'self'- and 'other'-oriented mental state attribution [Bradford et al., 2015, p.21; Doherty, 2008, p.73]. If this is the case, then it is relevant that some aspects of ToM may recruit EFs more heavily, or that some EFs may be more heavily recruited than others in expressing ToM [Bradford et al., 2015]. Here, the research literature has returned to the role of switching, inhibitory control, and working memory in the development of ToM abilities [Bradford et al., 2015, p.24; Carlson et al., 2015; Carlson & Moses, 2001; Mahy et al., 2014]. However, while behavioural and neurological studies have shown that inhibitory control appears necessary for some forms of ToM reasoning, it does not appear to be sufficient, favouring the (1ii.a/b) accounts over that of (1i) [Mahy et al., 2014]. The mapping of inhibitory processes within the brains of both children and adults has been well-documented: these processes have been shown to activate areas in the bilateral ventral prefrontal cortex, right parietal lobe, and right dorsolateral prefrontal cortex [Mahy et al., 2014, p.73 for further citations]. Here, research has indicated that both common and distinct regions of the brain are recruited for FB and inhibitory EF tasks [Mahy et al., 2014, p.73]. Mahy et al. also cited further research designed to match tasks' EF demands, which demonstrated that ToM tasks more heavily recruited the bilateral TPJ than inhibitory control tasks, in addition to a study implicating the right TPJ in the processing of mental states and the left TPJ for reasoning about false signs [Mahy et al., 2014, p.73]. Other research contrasting low and high inhibition conditions of belief reasoning found modulated activation in ToM areas like the bilateral TPJ but also in executive control areas such as the ventrolateral prefrontal cortex, suggesting that different parts of a ToM reasoning process may recruit different EF regions [Mahy et al., 2014, p.74]. The above findings all favour the (1ii.a/b) 'necessary' accounts over the (1i) 'sufficient' account. However, Mahy et al. also cited an EEG study (conducted by Sabbagh et al. in 2009) investigating the relationship between EF and ToM in children, which indicated that the resting alpha waveform (thought to depict the developing brain's functional, domain-general, maturational changes) estimates in the dorsomedial prefrontal cortex and TPJ correlated positively with individual differences in FB understanding irrespective of age and EF task performance considerations [2014, p.74]. These findings may also provide support for a theory-theory account (2.2.3) of ToM. Moreover, this work suggested that EF was insufficient for ToM, but also that EFs cannot be assumed to be necessary for the expression of ToM concepts (1ii.b), indicating that more investigation into functional development would be needed to properly separate and verify the (1ii.a) and (1ii.b) accounts [Mahy et al., 2014]. Carlson et al. conducted a (2015) study, contrasting these (1ii.a) 'emergence' and (1ii.b) 'expression' accounts, and concluding from their data in favour of the (1ii.a) 'emergence' account [2015]. They recognised however that this account could still fit within the bidirectional view taken in stance (3), the domain-general theory-theory perspective, or within a representational language-mediated account (§3) [Carlson et al., 2015]. These perspectives could also all be accommodated within a tripartite multi-directional model, but this will be discussed in §5.

On the other hand, turning now to stance **(2)**, it is possible that ToM is necessary for the development of EF [Carlson et al., 2015; Doherty, 2008; Perner, 1999]. For those taking this position, meta-representation might be seen as necessary for certain aspects of EF, where mental entities ('schemas': Norman & Shallice, 1986) manage basic actions and cognitive processes [Doherty, 2008, p.134; Perner, 1999]. Perner cited Wimmer (1989) in making his argument that a better understanding of one's own mind facilitates a better control of one's actions, because actions are governed by the mind, and proposing that mastering ToM tasks, such as the FB task, led to increased executive control [1999]. This account speculates that ToM may have evolved as a self-monitoring mechanism, enabling better self-control and thus more complex behaviour [Carlson et al., 2015; Doherty, 2008, p.135]. In examining this stance, it is important to consider issues of automaticity within the context of Norman & Shallice's concept of schemas [Doherty, 2008; Norman & Shallice, 1986]. During tasks such as the card-sorting game, it is hypothesised that people create schemas to encode the required actions for said task, so that when a certain card appears, the appropriate schema is activated, resulting in the action taking place [Doherty, 2008]. In this view, schemas would automatically inhibit one another so that in response to environmental input within the context of working memory, the optimally relevant schema is selected [Doherty, 2008, p.134]. In order to prevent a situation where the correct schema is not automatically activated (e.g. the card-sorting condition changes) an overseeing system would be required to inhibit the old schema [Doherty, 2008, p.134]. In this sense, inhibition consists of two distinct stages: (i) assessing the contents of schemas (their action sequences), and (ii) inhibiting the schema with undesirable contents, while possessing the meta-awareness that activating a schema initiates the action sequence it contains [Doherty, 2008]. This requirement for meta-representation is paralleled in the FB task: children need to be able to represent the relationship between a belief and the situation it corresponds to, as well being aware that beliefs cause action [Doherty, 2008]. Therefore, it is reasonable to suggest that common developments in a person's capacities for meta-representation and perceiving causality, such as meta-linguistic awareness (3.1), might underlie their abilities to think about beliefs and inhibit schemas based on their contents [Doherty, 2008, p.135]. If this is the case, then these developments should be concurrent and correlated [Doherty, 2008]. This stance requires further investigation within the field of neuroscience but could potentially account for the work cited by Mahy et al. (2014), which indicated both common and distinct activation of neural regions for ToM and EF in FB tasks. Work has however suggested that while EF and ToM developments may bidirectionally enhance one another, EF may have causal primacy [Carlson et al., 2015], although these findings could still fit within the model discussed in §5.

The theories and evidence underpinning stances (1ii.a/b) and (2) are also very relevant to stance **(3)**; it still remains entirely possible that developments in ToM and EF occur interdependently within a complex sequential process. This could account for the disparities in the evidence proposed to support each theory. Accounts advocating that ToM is necessary for EFs like inhibition, as well as those suggesting that the ability to inhibit and remove oneself from a current reality or perspective in order to switch to considering someone else's [Carlson et al., 2015; Carlson & Moses, 2001; Doherty, 2008; Perner, 1999], could all be accommodated within a model of complex developmentally-staged sequences, incorporating the role of language (§3, §5). In order to examine this, however, competing theories will be required to give much clearer specifications concerning which abilities, as well as the level required of each of them, are necessary to facilitate the expression of others [Doherty, 2008]. It would also be crucial to collect neuroimaging evidence in support of this account, although possible task designs are beyond the scope of this thesis. Here, neurological research indicating that EF is necessary for some but not all types of ToM reasoning [Mahy et al., 2014] might be an avenue for further exploration.

Bradford et al. cited evidence from research (the work of Fine et al., 2001) reporting the case study of an adult patient with purportedly congenital amygdala damage, schizophrenia and possibly on the autistic spectrum, who was shown to be severely impaired in ToM abilities but demonstrated no impairment in EF [2015; Fine et al., 2001] (although this is at odds with EF deficit accounts of autism: see §4). This was taken as a possible indication for the lack of a link between ToM and EF abilities [Bradford et al., 2015]. It could be argued that this research provides evidence against stances (1i), (2), and (4), but that it might still be accommodated by stances (1ii.a/b), because if EF is necessary but not sufficient for ToM development, then the impairments could be within a ToM-unique region. Here, Fine et al. proposed the amygdala to be such a ToM-specific region [2001]. However, it arguably remains possible that stance (3) could accommodate this research data. The work by Fine et al. [2001] does not necessarily preclude an interdependent EF-ToM relationship, especially given the fact that the single subject passed first-order but not second-order FB tasks, because it is still unclear which ToM and EF abilities interact as well as at what level these abilities must be functioning in order to enable subsequent progression in others. Moreover, this study was flawed, given the comorbidity of disorders, and that it only examined one subject in adulthood, with the authors making a number of inferences based upon retrospective judgment. In the majority of the research literature, however, there is a clearly documented link between ToM and EF across clinical and non-clinical populations, cultures, and age-ranges, even though the specifics of this link remain unclear [Bradford et al., 2015; Carlson et al., 2015; Doherty, 2008; Mahy et al., 2014]. It has also been demonstrated that greater cognitive demands in terms of EFs result in more egocentric errors in implicit ToM tasks, with some suggestion that EFs such as inhibition and attention-direction (§5) may only be recruited when explicitly required and not automatically in tasks tapping implicit ToM [Bradford et al., 2015, p.23]. It remains apparent that further investigation will be necessary to establish the role of EF in ToM development.

2.2.5 Summary So Far and Future Directions for the Theory behind ToM

ToM is an important topic across psychology, linguistics, and neuroscience, and this research overlap has both aided and constrained the objective of realising one overarching explanation for the existence, development, and role of ToM. Research within the fields of psychology and neuroscience has focused on differing aspects of ToM: psychology on how and when ToM emerges, and neuroscience on where in the brain it is located or takes place [Mahy et al., 2014], while language accounts have often fallen under the radar and are markedly under-represented. As a consequence, each field's theory and data have only marginally influenced that of the other [Mahy et al., 2014; Schaafsma et al., 2015]. A more holistic and integrated approach could generate more robust results, bridging the divide, and a number of authors have specified some ways in which this has so far been prevented and how this might be addressed in future

[Mahy et al., 2014; Schaafsma et al., 2015]. Some authors have taken the position that a combination of processes and neural regions proposed by two or more separate theories could potentially engender ToM development [Astington & Baird, 2005; Doherty, 2008; Mahy et al., 2014]. Some accounts accommodate this theoretical combination by drawing distinctions between types or levels of complexity of ToM, proposing that different processes support differing functions [Doherty, 2008; Schaafsma et al., 2015]. Other accounts placing language at the centre of ToM acquisition [Astington & Baird, 2005; Nilsson & De López, 2016] are critically important; the examination of these will be the focus of §3. Furthermore, if any accounts of ToM are to be appropriately scrutinised within a neuroimaging research context, then the theories concerning the emergence and expression of ToM will need to be refined, because a number of them are not necessarily mutually exclusive [Mahy et al., 2014, p.76]. As a result, neural studies have so far generated both potential support for and arguments against each theory because of a lack of sufficient evidence either way [Mahy et al., 2014; Schaafsma et al., 2015]. They have also highlighted problems surrounding the articulation of the proposed mechanisms behind ToM acquisition and the conceptualisation of neural network maturity [Mahy et al., 2014; Schaafsma et al., 2015]. A number of authors have also underlined issues regarding task heterogeneity and construction in the research literature, as well as the disproportionate number of neuroimaging studies focusing on adult ToM processing, with few examining its developmental acquisition [Byom & Mutlu, 2013; Mahy et al., 2014, p.77; Perner, 1999; Schaafsma et al., 2015]. All of these issues have ramifications for work on ToM across the literature, emphasising the need for an interdisciplinary and holistic approach.

2.3 ToM: The Developmental Perspective and Research Context

ToM develops out of and alongside a number of other abilities, which are all associated with mentalising to varying extents. ToM and language abilities draw on this range of skills and they all share some degree of overlap with one another. It is therefore appropriate to consider the developmental perspective of ToM, and evaluate: false belief and its potential precursors; gaze-following and goal-orientation, as well as their relevance to belief-tracking, mental state attribution, and the establishment of joint attention; and finally, visual, spatial, cognitive, and affective perspective-taking. Moreover, numerous experimental tasks have been designed and refined in attempts to find evidence to substantiate the proposed accounts of ToM discussed above, as well as the language accounts discussed in §3. If any explanation for ToM development and acquisition is to be reached, it is thus critically important to examine each developmental paradigm's theoretical background, context, and instantiation within the research literature, as well as the specific experimental tasks designed and employed by researchers. This is particularly important given that the ways in which researchers frame their conceptualisations of ToM and select their experimental measures influence the course of research [Byom & Mutlu, 2013]. The following subsections will review these areas before presenting and analysing the language accounts of ToM in §3, as it will be necessary to refer back to these constructs.

2.3.1 False Belief and Potential Precursors: Pretence and Imaginary Play

Given that people's beliefs cannot be directly observed by the researcher, tests distinguishing between reality and mental representations of reality are required in order to diagnose them [Bradford et al, 2015; Doherty, 2008]. This can be done because of the existence of false belief (FB), in other words, a belief state (or proposition, see 2.1) that does not correspond to the actual state of reality [Dennett, 1978; Doherty, 2008; Perner, 1999]. In order to accurately attribute mental states to someone else, one must understand firstly that is possible for someone to possess a FB, and secondly that it is their FB which causes this other person to behave in a way that cannot be explained solely by looking at the objective reality [Doherty, 2008]. To achieve this, children must comprehend both that: (1) people can act according to the real world or a possible world (this is the case in a pretend scenario); and (2) that a person with a FB is trying to act in accordance with the real world but acts according to a possible world, because they misrepresent that possible world as the real world [Perner, 1999, p.7]. Tests of false belief involve presenting scenarios to children, and assessing whether they can recognise that other people can hold beliefs about reality, which are mistaken (and not the same as their own), as well as predict how these people will behave on the basis of these mistaken or different beliefs [Bradford et al., 2015; Farrant et al., 2006]. In typically developing children, this first-order FB understanding emerges at around 4-5 years-of-age, with second-order FB understanding found to be present from around 6-9 years-of-age (their differences are described later in this section) [Bradford et al., 2015; Doherty, 2008; Frith & Frith, 2005; Happé, 1993; Nilsson & De López, 2016; Perner, 1999; Sodian & Kristen, 2010; Wellman et al, 2001]. This co-occurs with a developmental shift that facilitates a more enhanced understanding of desire and the importance of perceptual access for knowledge, the abilities to attribute knowledge and ignorance, as well as distinguish between appearance and reality, and infer more complex emotions [Doherty, 2008, p.5; Perner, 1999].

Prior to this developmental shift, children tend to fail ToM FB tasks, signalling that they do not appear to understand belief, and instead predict actions based upon their own (egocentric) understanding of reality, experiencing difficulty when asked to explain why someone might act differently [Bradford et al., 2015; Doherty, 2008; Meert et al., 2017; Perner, 1999]. EF accounts of ToM might argue this to be the result of failing to inhibit one's own perspective and current state of knowledge [Bradford et al., 2015]. Some ability to reason about desire (grounded in a basic understanding of positive vs. negative emotions), as well as mirror self-recognition, and signs of empathic responses to another's distress, however, appear to develop significantly earlier at around 18 months-of-age [Doherty, 2008; Perner, 1999; Repacholi & Gopnik, 1997]. From the age of 18-months to 3-years, children develop their capacities for imitation, imaginary play, and pretence, while anticipatory looking presents at around 3 years-of-age, suggesting an early, implicit understanding of ToM [Perner, 1999]. An understanding of desire may function as a potential precursor of belief recognition, as it equips the child with the knowledge that certain reasons motivate people's actions, e.g. attempting to achieve their goals [Perner, 1999]. Once children recognise that: (a) someone must have a reason for behaving in a certain way, (b) there can be a distinction between appearance and reality, and (c) people can act according to possible worlds (e.g. pretence); it is possible that they might progress to inferring that people's beliefs trigger their actions and that the person in question possesses a FB [Byom & Mutlu, 2013; Doherty, 2008; Perner, 1999]. In adults, being able to interpret behaviour in this way is also important for gaining an insight into a person's mood or personality, which will affect their social interaction [Doherty, 2008]. Some understanding of desire is required to predict basic behaviour, and while beliefs about the world are often objectively true (and shared), desire can be more individual and subjective [Doherty, 2008]. However, some authors theorise that children and adults may access different levels of desire conceptualisation: (1) a simpler objective idea, equating it to a property of specific objects or actions; and (2) a more complex subjective and internal state, framing it as a property of situations by relating a propositional attitude to a mental representation of a potential future state [Doherty, 2008, p.56]. Accessing a more subjective percept of desire allows children to move past the idea that objects have inherent (un-) desirability, and potentially start to represent that people may perceive things differently [Doherty, 2008; Perner, 1991; Repacholi & Gopnik, 1997]. This early representational capacity might be linked with a developing meta-linguistic awareness (see 3.1).

Distinguishing between appearance and reality is the first step in transpassing outward appearances and in many ways parallels FB, which could be framed as the "belief-reality distinction" [Doherty, 2008, p.60], as they both involve assigning multiple representations to a single object, entity, or situation [Lillard, 1993]. Research has indicated that children begin accomplishing these tasks at about the same time and the unexpected contents FB task is an effective example of this association (also mentioned in: 2.2.2; and presented in greater detail along with other examples of FB tasks in the last two paragraphs of this section) [Doherty, 2008; Gopnik & Astington, 1988]. The only distinction lies in the factoring of another naïve entity into the equation [Doherty, 2008]. Problems with this task might lie in an inability to support dual representations (appearance vs. reality) or comprehend visual misrepresentation [Doherty, 2008; Flavell et al., 1983; Lillard, 1993; Perner, 1991]. Recognising that appearances can be misleading might thus be facilitated by understanding that appearances can influence beliefs and vice versa, as well as that 'shared' knowledge can sometimes be ambiguous [Doherty, 2008]. Being able to understand the concept of knowledge also requires some understanding of belief [Doherty, 2008]. Research has indicated that at around the same time that children pass FB tasks, they also start to piece together the relationship between knowledge and experience, comprehending that knowledge equates to a true belief validated through experience [Doherty, 2008; Povinelli & DeBlois, 1992]. In this way, they must understand that the source of a belief lies in an experience (either a perceptually direct, verbally reported, or mentally inferred one); prior to this, children appear to update their beliefs via experience but do not encode the source, seeming unable to recall how they came to know something [Doherty, 2008, p.64; Gopnik & Astington, 1988; Povinelli & DeBlois, 1992]. Gopnik & Astington's (1988) and Povinelli & DeBlois' (1992) findings also present a counter-argument to simulation theory accounts (2.2.2), which assume that direct and transparent access to one's own mind enables mentalising about the minds of others. Predicting others' inferences is a second-order mental-state task because it involves pairing mental-state concepts: knowledge and thought [Doherty, 2008]. Appreciating the idea of knowledge as well as potential distinctions between appearance and reality is the starting point for many types of deception (2.1) [Doherty, 2008]. Competence with deceptive actions like lying or concealing information (e.g. denial and secrecy) appears to correlate with performance on FB tasks, and it has been observed that children begin to behave deceptively from the age of 3-years [Doherty, 2008; Perner, 1999]. These capacities may originate from survival instincts and early competence may be indicative of ToM development [Doherty, 2008]. However, research into early 'pseudo-lying' has suggested that it may be aimed more at manipulating others' behaviour to avoid punishment rather than at truly influencing another's beliefs [Doherty, 2008, p.23].

It has been argued that a full understanding of belief requires children to recognise that propositions are evaluated irrespective of their objective truth [Doherty, 2008, p.43; Perner, 1999]. Although many theorists (particularly simulation theorists, 2.2.2) relate pretence to FB, these diverge in the sense that pretend play can occur without necessitating such a recognition, because pretence merely requires all participants to jointly evaluate a proposition as false [Doherty, 2008; Perner, 1999]. In this way, it may serve as a precursor or zone of proximal development to true ToM, because imaginary play and pretence might enhance the child's developing social awareness and cognition, just as this enhanced understanding would support more complex levels of pretend play [Doherty, 2008, p.91; Lillard, 1993]. Pretend play presents at approximately 18 months-of-age, typically beginning with substantial scaffolding by adults and with props, but later progressing to greater independence on the part of the child, requiring less support in the form of props and becoming a more social activity [Bretherton, 1984; Doherty, 2008, p.92]. Pretend play can include: (1) object substitution; (2) attribution of non-real properties to objects or situations; and (3) invention of non-existent objects [Bretherton, 1984; Doherty, 2008, p.92; Lillard, 1993]. The attribution of non-real properties to objects or situations feels particularly pertinent to ToM ability, as this seems to parallel the process of attributing mental states, which are not directly observable, to other entities. Children younger than 2 years-of-age, however, struggle to maintain their role within a pretend scenario, finding it hard to cope with transformations [Doherty, 2008; Kavanaugh & Harris, 1992]. Once they can follow pretend-scenario transformations performed by other participants, this implies that they are capable of inferring another person's intended pretence [Doherty, 2008; Kavanaugh & Harris, 1992]. This allows them to engage in joint pretence, important for developing skills in social interaction [Doherty, 2008]. More sophisticated forms of social pretence can then include the taking on and attributing of non-real emotions, roles, personalities, or mental representations, which might support ToM development by freeing representations from the confines of reality and allowing them to be modified [Bretherton, 1984; Lillard, 1993]; this might be similar to the process of engaging in a logical thought experiment. Here, Leslie's ToMM theory might also equate the capacity for pretence with that for general mental-state ascription [Leslie, 1994]. After 3-4 years-of-age, the ability to step out of pretend roles to engage in meta-communication begins to emerge [Doherty, 2008; Howes & Matheson, 1992]. This is very relevant to general ToM development as it parallels Perner's theory of a later-emerging capacity for meta-representation [1999], and also the concept of an emerging meta-linguistic awareness (3.1), relevant to language accounts of ToM. This ability may even be supported by early, general ToM development [Doherty, 2008]. Neither modular-, simulation-, nor theory theorists have yet fully accounted for how children acquire a conceptual representation of pretence [Doherty, 2008], although it remains possible that it might be supported by a developing understanding of representation as a result of the language acquisition process (3.1). Leslie asserted that children must have some method for separating representations of pretence and reality, possibly by tagging propositions with mental-state expressions, such as ‘x (person) pretends that y (object/situation) has z (property)’ [Leslie, 1994]. This could parallel belief attribution, where it might be the case that ‘x (person) believes that y (proposition) is z (true/false)’ [Doherty, 2008; Leslie, 1994]. However, some have argued that this framing may be “unnecessarily elaborate” and that pretend propositions would only need to be tagged as false, which would not mean that children perceive pretence as a mental state; instead, they may perceive it as an action [Doherty, 2008, p.96]. Perner et al. proposed that prior to understanding FB, 3-year-olds possessed a transitional concept of ‘prelief’ that did not distinguish between pretence and belief, but later allowed children to move from a conceptualisation of true and false actions to true and false beliefs once they were able to access meta-representation [1994; Doherty, 2008]. Harris’ simulation theory account (discussed in 2.2.2) also attempted to explain pretence, proposing that children have the capacity to imagine holding pretend mental-states, which they would achieve by altering their ‘default settings’ corresponding to reality [1992; Doherty, 2008]. Children with autism (§4) do not appear to engage in pretend play, although research has shown that they are capable of it [Doherty, 2008, p.93; Kavanaugh & Harris, 1992]. Specific deficits in autistic children’s understanding of FB have also been observed, but a high verbal mental age (VMA) in older autistic participants seems to offer them an alternative route to working out the answers to tasks via logical deduction [Baron-Cohen et al., 1985; Doherty, 2008; Frith & Frith, 2005; Happé, 1995a; Sodian & Kristen, 2010], which indicates that language plays an important role in facilitating ToM ability (more in §3 and §4).[2]

Although children do not demonstrate an explicit understanding of belief until they are around 4 years-of-age, a number of studies have demonstrated that they engage in anticipatory looking during FB tasks, which might signal an implicit understanding [Baillargeon et al., 2010; Doherty, 2008; Garnham & Ruffman, 2001; Meert et al., 2017]. This anticipatory looking to a correct location precedes success with an explicit FB task question by at least a year and may indicate a discrepancy between children’s confidence in or reliance upon implicit mental vs. explicit verbal reasoning [Doherty, 2008; Garnham & Ruffman, 2001; Meert et al., 2017]. This finding is important because it could offer an insight into how children’s concept of belief develops. Anticipatory looking is also linked to gaze-following and an understanding of goal direction (2.3.2) as well as causal relations. While infant anticipatory looking might be taken to support modular

[2] It has also been shown that developmental difficulties with FB tasks are not the result of demands posed by language or EFs, as tasks with a reduced cognitive or verbal load have proven to be comparably problematic for young children [Doherty, 2008; Sodian & Kristen, 2010; Wellman et al., 2001].

theorists' account of ToM [Baillargeon et al., 2010], it is also possible that the maturational shift from implicit to explicit understanding might reflect theory theorists' accounts which hold conceptual change responsible for ToM development.

A number of tasks have been designed to probe FB understanding. The **unexpected contents FB task** (briefly mentioned in 2.2.2) is one of these. This task was devised by Hogrefe et al. to give children a personal experience of having a FB and then investigate whether they were subsequently more capable of attributing FBs to others [1986; Doherty, 2008]. This is also known as the 'smarties' task, because during the task children are exposed to a familiar container (e.g. a tube of smarties) [Hogrefe et al., 1986; Doherty, 2008]. The children are questioned about what the contents might be, typically give the obvious answer (e.g. smarties), and are shown to be mistaken (e.g. a pencil), and are then questioned about what another person (e.g. their friend) would envision the contents to be [Hogrefe et al., 1986; Doherty, 2008]. Research has shown that children younger than 4-5 years-of-age tend to perform poorly on both the own- and other- belief conditions of this task, and are typically unable to recall what they originally imagined the contents to be [Doherty, 2008; Gopnik & Astington, 1988]. As previously discussed, the recall difficulties in this task do not provide support for simulation theory accounts, but it might be possible that these findings could be accounted for by deficits in or demands that are too high upon EFs (such as inhibitory, working memory, or switching loads) or meta-representational capacity (2.2.3). Another, and perhaps the most famous, is the **unexpected transfer FB task** (or the Sally-Ann/Sally-Tony task) designed by Dennett and implemented by Wimmer and Perner [Baron-Cohen et al., 1985; Dennett, 1978; Doherty, 2008; Wimmer & Perner, 1983]. The child is told a story about character A (e.g. Sally), who places an object (e.g. a marble) in a particular location (e.g. inside a basket). While character A is out of the room, character B (e.g. Ann) moves the object to another location (e.g. inside a cupboard). In addition to reality and memory questions, the child is then asked where character A will look for the object when they return to the room. This task is effective because it clearly distinguishes between reality and belief: what the child vs. character A knows or believes to be the case [Doherty, 2008]. Children from 4 years-of-age onwards tend to perform successfully on this task, prior to this, most 3-year-olds fail by answering with the new (and real) location of the object instead of the old (and believed) location [Doherty, 2008; Wimmer & Perner, 1983]. Autistic children at the same developmental stage tend to fail this task and there is a considerable delay before they start to perform successfully (more in §4) [Baron-Cohen et al., 1985; Happé, 1995a]. In order to probe higher- or second-order ToM and FB (false beliefs about beliefs) Perner & Wimmer developed the **ice-cream-van story** [1985]. Children begin to pass this kind of test from

6-7 years-of-age onwards, and by 10 years-of-age most perform successfully [Byom & Mutlu, 2013; Doherty, 2008; Nilsson & De López, 2016; Perner & Wimmer, 1985; Sodian & Kristen, 2010]. During the story, characters A & B (John & Mary) are both informed independently about the unexpected location transfer of an object (an ice-cream van); the child is subsequently asked where character A thinks character B will go to buy ice-cream [Perner & Wimmer, 1985]. It may be the case that difficulties with this task are due to increased EF demands in working memory, although other, less convoluted, versions of the unexpected transfer task still show a discrepancy in performance on first- and second-order beliefs [Doherty, 2008]. Happé also designed a second-order ToM task probing thoughts and feelings: **the strange stories task** [1994]. These stories depicted interactions involving sarcasm and double-bluffs, with similar performance statistics to the ice-cream van story, and even autistic children who passed second-order FB tasks struggled with this task [Doherty, 2008; Happé, 1994]. Reasoning about emotions parallels the process of reasoning about beliefs and requires children to grasp that emotions are belief-based; this forms the basis of children's capacity for empathy [Doherty, 2008]. Understanding emotion goes beyond other second-order belief-based tasks because it requires the child to evaluate not only the interaction of one person's mental state with that of another, but also the interaction of two different mental states within one person [Doherty, 2008, p.70].

Classic FB tasks have been criticised, because, like any cognitive test, they measure not only the targeted conceptual competence but also other cognitive abilities (such as EFs) that are necessary in order for the subject to perform it [Doherty, 2008]. In this way, the FB task also requires an operational working memory to follow the events of a narrative, linguistic and conceptual comprehension to understand the question being asked, and possibly the ability to inhibit their own representation of reality [Byom & Mutlu, 2013; Doherty, 2008]. Wellman et al.'s (2001) meta-analysis (discussed in 2.2.4) aimed to establish and investigate the factors determining FB task performance but found that performance did not appear to depend on the surface features of the tasks [Doherty, 2008; Wellman et al., 2001]. This implied that some conceptual or analytical development was required to succeed on the tasks, and while some variables were implicated as influential factors (temporal marking, motive, salience, participation, target object absence, country), they did not raise 3-year-old subjects' performance beyond that of chance [Doherty, 2008, p.14; Wellman et al., 2001]. The FB task also entails a number of assumptions, including that an agent's behaviour is grounded in their mental representations of the world [Doherty, 2008]. In this sense, classic ToM tasks also tap shared world knowledge as well as the abilities to interpret actions and perceive social and vocal cues [Byom & Mutlu, 2013].

It is also true that classic tasks generally test passive and reflective aspects of ToM, occurring within a lab setting, where subjects have a generous response-time and are not required to respond as if they were in the scenario themselves [Byom & Mutlu, 2013]. For these reasons, such tasks might overestimate ToM ability in subjects who would otherwise demonstrate impaired ToM in a reactive online setting such as their daily lives [Byom & Mutlu, 2013].[3] On the other hand, such explicit tests of FB might require too high a level of reflective meta-cognition and actually underestimate implicit ToM ability. It is also possible that a more implicit form of belief processing, such as belief-tracking, might be a different type of mentalising to explicit belief reasoning [Meert et al., 2017]. Byom & Mutlu made the insightful point that classic ToM experimental paradigms typically require a subject to make inferences from static images, textual stories, and video vignettes, which may only demonstrate an observer's 'third-party' understanding of ToM and FB [2013, p.4]. Research has indicated that there are significant differences in the way that ToM is reflected in terms of motivational consequences, and across neural processes, when the subject is socially and emotionally engaged in the scenario rather than just looking in [Byom & Mutlu, 2013]. Byom & Mutlu also emphasised that active involvement in a scenario facilitates the establishment of shared goals, intentions, and actions, which might help a subject to make ToM inferences by allowing them to refer to their first-hand experience of the scenario [Byom & Mutlu, 2013]. For these reasons, emerging interactive simulation tools, stimuli, and tasks, which test online ToM ability, may help to better clarify the less well-defined areas of the ToM research data [Byom & Mutlu, 2013].

2.3.2 Gaze Following, Goal-Orientation, and Joint Attention

Gaze cues play a vital role in social interaction and are a critically important method by which people can infer the mental states of others [Byom & Mutlu, 2013; Dunphy-Lelii & Wellman, 2004; Griffin & Baron-Cohen, 2002]. By following another person's gaze, one is able to infer the direction or object of their attention (§5) and thereby their intentions [Byom & Mutlu, 2013; Dunphy-Lelii & Wellman, 2004]. By assigning referential meaning to gaze, gaze-following scaffolds language acquisition: 18 month-olds interpret unknown words as a label for an object that a speaker is gazing at [Dunphy-Lelii & Wellman, 2004: citing Baldwin, 1993]. In these ways, gaze-following scaffolds not only language but also mentalising within a conversation, by allowing a person to monitor comprehension of their speech, invite feedback or a response, prompt

[3] For example, autistic people. This could also be relevant to Happé's (1995a) findings that autistic participants with a high VMA were able to 'hack out' solutions to FB tasks using verbal logic; it may be that with adequate time allowance in a lab setting, they perform more successfully on FB tasks.

conversational turn-taking, and resolve linguistic ambiguities arising from the use of non-literal or figurative language [Byom & Mutlu, 2013, p.3]. Gaze cues and the perception of changes in facial expressions are therefore also key factors in emotion recognition and belief attribution [Byom & Mutlu, 2013; Griffin & Baron-Cohen, 2002]. Understanding that someone sees something is clearly linked to understanding that someone knows something [Doherty, 2008; Dunphy-Lelii & Wellman, 2004; Garnham & Ruffman, 2001]. Humans appear to have evolved specific adaptations that make gaze-detection and signalling easier, such as wider eyes, white sclera as a backdrop highlighting the iris and pupil, and eye movement independent of the head [Doherty, 2008]. In early infancy, the most basic gaze-perception ability presents: distinguishing direct from averted gaze [Doherty, 2008, p.109; Happé & Frith, 2014]. Within their first few months, infants can follow another's head direction to an object if it is within their field-of-view and are able to follow gaze beyond this from around 12-months onwards [Doherty, 2008; Dunphy-Lelii & Wellman, 2004]. From 18-months onwards, they can start to follow eye direction as their only cue and if the gaze direction goes beyond their field of view, expect an object to be behind the occluding body [Doherty, 2008; Dunphy-Lelii & Wellman, 2004]. Research has indicated that infants are at the very least sensitive to surface features of direct-gaze, but that they can only begin to make explicit judgements about it from 3-4 years-of-age onwards, steadily improving over time [Doherty, 2008]. Gaze-following is also important for understanding another person's engagement within an attempted interaction: young infants can recognise an adult's lack of involvement with them and increase their attempts at communication, and from 2.5 years-of-age they can act according to past episodes of disengagement, attempting to communicate about those events [Doherty, 2008, p.123; Happé & Frith, 2014]. In this way, ascribing attention (§5) or engagement to another may be a low-level precursor to the ability to attribute other more complex mental states. Even though young children seem to be capable of spatially representing a relationship between an object and the eyes or head of the person looking at it, as well as orient their attention by using another person's gaze, this does not equate to having a full understanding of attention (§5) as an internally represented mental state [Doherty, 2008, p.127]. This understanding might only begin to emerge once they are capable of making explicit judgments about another's visual attention, in addition to representing other mental states such as belief and knowledge [Doherty, 2008].

Around the age of 10-12 months, social-referencing abilities involving gaze-cues emerge, as well as early language abilities like communicative gesturing (such as declarative pointing), which is designed to orient attention to an object or goal, and signal an intention [Bates & Dick, 2002; Byom & Mutlu, 2013; Griffin & Baron-Cohen, 2002]. Such abilities are important in establishing joint attention in dyadic and triadic interactions as well as in facilitating shared information, goals, intentions, and actions [Byom & Mutlu, 2013; Griffin & Baron-Cohen, 2002]. For these reasons, precursor facilities such as gaze-following, goal-orientation, and joint attention have been viewed as an early, basic ToM, through which the child encodes an adult's mental states of attention to, and emotion regarding, a state of affairs [Doherty, 2008; Griffin & Baron-Cohen, 2002, p.87; Sodian & Kristen, 2010]. Some authors have taken the position that prior to a true ToM, children take a teleological stance, which interprets actions as goal-directed but functions as a non-mentalistic and non-causal precursor; this position does not endow children with a full knowledge of belief, desire, and intentions [Griffin & Baron-Cohen, 2002, p.92]. Some nativist and modular theorists have appeared to equate understanding visual attention with mind-reading or mentalising, while other theorists have proposed that learning to follow gaze could occur in isolation from grasping its implications, via a process of reward, reinforcement, and conditioning [Doherty, 2008]. It may also be the case that the parent encourages and teaches gaze-following during joint attention episodes and triadic interactions [Doherty, 2008], using verbal stimuli. These accounts hold that once a child has learned to follow gaze, they can then begin to perceive that visual experience is subjective, by realising that other people might have the same or a different experience of seeing something as them [Doherty, 2008, p.120]. This is relevant for the consideration of visual perspective-taking (VPT) (see 2.3.3). Importantly for language-related accounts of ToM, early individual differences in joint attention capacity have also been strongly linked to later gesture and language skills [Bates & Dick, 2002].

2.3.3 Perspective-Taking

It has been argued that gaze-following represents an implicit form of perspective-taking (PT) because it is used to track perspectives, and the monitoring of it can serve as an indirect measure of ToM-related visual, spatial, cognitive, and affective perspective-taking (VPT, SPT, CPT, APT) [Surtees et al., 2016]. The ability to take another's perspective is a crucial aspect of social cognition, allowing people to represent what others see and how they see it, while simultaneously representing their own perspective [Pearson et al., 2013; Schurz et al., 2015]. APT involves explicitly inferring the emotions of other people and is arguably equally as important as CPT and VPT in order to fully appreciate another's mind [Farrant, 2015, p.2]. In these ways, PT is clearly related to ToM and mental-state attribution, because these abilities all require a person to recognise that another person can have a different representation of the world, either visually or mentally [Pearson et al., 2013; Schurz et al., 2015]. For these reasons, VPT has often been considered a low-level mechanism within a broader ToM system [Schurz et al., 2015]. Computing someone's visual perspective involves processes that are also important for spatial navigation [Frith & Frith, 2005]. To make this computation, a person must decode and evaluate spatial information within a social framework: this information consists of the relationships between egocentric and allocentric spatial coordinates of the self and other as well as any other relevant objects or entities within the environment [Frith & Frith, 2005; Pearson et al., 2013]. With its roots in spatial cognition, frames of reference, and the physiological representation of spatial relationships, investigating PT might therefore offer an avenue for examining the neural bases of ToM [Frith & Frith, 2005; Schurz et al., 2015]. There are also links between language and PT which should be considered, since language verbally mediates PT, by allowing two interlocutors to communicate, invite, and take one another's perspectives.[4]

Based on Flavell's work in the 1970s, two levels of VPT have been discussed in the literature: level 1 (VPT1), referring to the ability to infer *whether* someone can see an object ('perceptual' PT); and level 2 (VPT2), referring to the ability to judge *how* someone sees an object ('conceptual' PT) [Doherty, 2008; Farrant et al., 2006; Hamilton et al., 2009; Surtees et al., 2016]. VPT1 involves calculating whether a person's head or body is in the right position to see something and their eyes are open, as well as whether or not their line-of-sight is obstructed [Hamilton et al., 2009]. VPT2, on the other hand, requires an understanding that the same object can be perceived differently by two different people [Farrant et al., 2006; Hamilton et al., 2009; Sodian & Kristen, 2010]. VPT1 tasks generally involve line-of-sight measures, such as hiding an object from another entity and asking the subject to judge whether that other entity can see the object [Hamilton et al., 2009], or asking the subject to ascertain the number of dots on a room's wall that are visible to them and another entity facing or looking away from the target (the 'dots' task) [Samson et al., 2010; Surtees et al., 2016; Todd & Simpson, 2016]. VPT2 tasks involve the subject and another entity occupying different positions but facing the same object, and asking the subject to judge how the other entity perceives the object (e.g. what side, part, or angle of it that they see) such as the three-mountains or 'droodles' tasks, sometimes also asking what the other entity will do on the basis of this perspective [Hamilton et al., 2009; Farrant et al., 2006; Perner, 1999; Surtees et al., 2016]. Children begin to pass VPT1 tasks from around 2 years-of-age, whereas successful performance on VPT2 tasks occurs at around the same age as successful performance on FB tasks: 4-5 years-of-age [Doherty, 2008; Farrant et al., 2006; Frith & Frith, 2005; Hamilton et al., 2009; Sodian & Kristen, 2010; Surtees et al., 2016]. Some theorists have argued that this is

[4] This relationship will be discussed further with reference to the SLI population in 3.2.1.

because tasks requiring FB and VPT2 reasoning may both tap conceptual perspective-taking [Doherty, 2008]. Research has further indicated that VPT1 calculations are triggered relatively automatically, whereas VPT2 (and explicit judgments about both VPT1 & VPT2, which are subject to egocentric interference) require some degree of cognitive control [Surtees et al., 2016, p.97]. Children's errors on VPT tasks tend to be egocentric, which is thought to reflect the over-application of their own perspective [Surtees et al., 2016, p.97]. Developmental deficits in EFs such as inhibitory control and switching may also account for these difficulties. Work on VPT has also expanded into emotional (affective) PT and cognitive PT: mentalising about what people do or do not feel or know [Perner, 1999]. Interestingly here, research has indicated that experiencing certain emotions can impede VPT abilities: anxiety (relative to anger and neutral feelings) particularly impaired adults' spontaneous calculations of what another social entity could see during the 'dots' task [Todd & Simpson, 2016, p.88]. Since anxiety did not impair subjects' perspective calculations when a non-social entity (e.g. a standing lamp) was present, it was proposed that anxiety is particularly disruptive for ToM and general social cognition [Todd & Simpson, 2016].[5]

Neuroimaging research has implicated some common brain regions that are activated during both ToM and VPT tasks: the TPJ, MPFC, and precuneus [Hamilton et al., 2009; Schurz et al., 2015]. However, in recent meta-analyses, VPT tasks appeared to activate the right TPJ, ventral MPFC, and ventral precuneus, while ToM or mental-state reasoning tasks appeared to activate the bilateral posterior TPJ and dorsal MPFC [Schurz et al., 2015]. The VPT areas were also found to be involved in the spontaneous processing of 'other' (altercentric) perspective information during 'self' perspective judgements [Schurz et al., 2015; Surtees et al., 2016]. Schurz et al. asserted that the lack of a neural overlap was problematic when considered in the context of conceptual and developmental links between VPT and ToM, proposing that methodological differences within VPT and ToM research might account for inconsistencies in the findings [Schurz et al., 2015, p.386]. They further argued that VPT imaging studies have concentrated on explicit VPT and the contrast between 'self'- vs. 'other' perspective judgments, while ToM imaging studies have focused on the identification of relevant areas by examining 'other' mental-perspective judgments vs. non-mental control conditions [Schurz et al., 2015, p.386]. It was also suggested that, given the overlap in activations, an automatic 'self'- vs. 'other' contrast might be conflating aspects of mental-state reasoning in neural studies [Schurz et al., 2015]. However, if an implicit

[5] This finding has interesting implications for the consideration of autism, with which anxiety often co-occurs (more in §4, 4.1 and §5).

processing of ‘other’ perspectives recruits areas within the ToM brain network, then this could be taken as evidence in support of a neural and conceptual relationship between ToM and VPT [Schurz et al., 2015]. This ‘self’-’other’ distinction is also important to bear in mind when reviewing simulation theory accounts (2.2.2). In addition to parallel FB imaging research implicating the right TPJ in tasks which involved mental states creating perspective differences, Schurz et al. found evidence that VPT does engage ToM areas when scenes depict a difference in perspective, but not when subjects are required to adopt another’s perspective [2015, p.394-5]. Other research has found that altercentric information about ‘other’ perspectives is not spontaneously calculated in VPT2 tasks [Surtees et al., 2016]. The authors attribute this to issues of cognitive efficiency and the resultant cost of inflexibility regarding the type of perspective that can be processed, with reference to parallels from FB research [Surtees et al., 2016]. They also question whether spontaneous implicit PT, in comparison to explicit PT, actually encodes an equally rich level of information about the spatial relationships of an entity to the objects within that entity’s perspective, or whether the subject merely generates two basic sets corresponding to ‘seen by self’ or ‘seen by other’ [Surtees et al., 2016, p.103]. These questions necessitate further investigation, although the idea of implicit and explicit PT nicely parallels the concept of an implicit and an explicit ToM (2.2.3, 2.2.4, 2.3.1).

2.4 Deficits

Deficits in ToM are an important cognitive measure and test case for the examination of psychological and neurological disorders which are currently only behaviourally or symptomatically diagnosed. These include autism, schizophrenia, psychopathy, attention deficit hyperactivity disorder (ADHD) and social anxiety disorders, many of which share overlapping characteristics, for example deficits in ToM or pragmatic language competence [Bambini et al., 2016; Bradford et al., 2015; Corcoran et al., 1995; Craig et al., 2016; Cummings, 2014; Gadow, 2012; Happé & Frith, 2014; Hollocks et al., 2014; Jones et al., 2017; Joyce et al., 2017; Nilsson & De López, 2016; Rogers et al., 2006; Tick et al., 2016]. Deficits in ToM also correlate with deficits in language: children who have SLI (3.2.1) or are deaf with non-signing parents (3.2.2) have been shown to be significantly delayed in passing FB tasks, which indicates that language plays a significant role in facilitating ToM development and acquisition [Astington & Baird, 2005; de Villiers, 2005; Doherty, 2008; Farrant et al., 2006; Nilsson & De López, 2016; Perner, 1999; Tager-Flusberg, 2007]. Deficits in ToM and language also co-occur within people on the autistic spectrum (§4) [Astington & Baird, 2005; Doherty, 2008; Durrleman & Franck, 2015; Happé, 1993, 1995a, 1995b; Tager-Flusberg & Joseph, 2005]. Taking into account the theories, concepts, and findings discussed throughout §2, §3 will conduct an in-depth analysis of the relationship between language and ToM.

§3 Theory of Mind and Language

3.1 What is the Relationship between Language and ToM?

As previously discussed, ToM is a broad umbrella term encompassing a diverse range of interrelated processes. These include mental-state attribution, perspective-taking, and empathy, as well as a hypothetical cognitive structure supporting these processes, and further still, the relevant associated research areas and theories [Astington & Baird, 2005]. The term 'language' also engenders many wide-ranging connotations. Therefore, it is vital to bear such multifaceted senses in mind when examining the relationship between ToM and language, because different elements of each system may relate to one another in distinctive and complex ways [Astington & Baird, 2005; Doherty, 2008]. For humans, language serves as a multi-purpose tool; one that is used both for intra-individual representation (e.g. verbal thought) and inter-individual communication (e.g. dyadic interaction within discourse), the latter of which must be further divided into linguistic ability (both structural and pragmatic competences) and linguistic environment (e.g. participation within a social context) [Astington & Baird, 2005, p.6]. Astington & Baird emphasised that linguistic ability and linguistic environment influenced one another but might differentially relate to ToM, as social contexts determine the types of input children are exposed to when developing their linguistic competence, while their level of ability affects the type of communication they receive [2005, p.6]. Structurally, a distinction exists between: form, such as phonology, morphology, and syntax; and meaning, such as lexical and discourse semantics, and pragmatics; although these also overlap and interact [Astington & Baird, 2005; Doherty, 2008]. It is also important to differentiate between language in general and 'mentalistic' language, which are terms used to refer specifically to the mind as well as mental states and processes [Astington & Baird, 2005, p.6]. Mental verbs (such as 'think', 'know', 'want', 'believe', 'hope', etc.) constitute a particularly significant proportion of these terms and straddle the divide between semantics (the lexical encoding of mental state concepts) and syntax (the sentential complementation of mental state verbs) [Astington & Baird, 2005; Nilsson & De López, 2016].

Given the many varied components making up the systems of language and ToM, it is likely that they relate to one another in different ways and that these associations could change over time [Astington & Baird, 2005; Chahboun et al., 2016a]. The extent of the human capacities for language and ToM are those which arguably set us apart from other species, and it seems improbable and inefficient that these defining abilities could have evolved in an unrelated and isolated manner [Doherty, 2008, p.151]. Effective language use requires the speaker to consider the listener's potential mental states, and the listener to make inferences about those of the speaker [Doherty, 2008, p.151]. Since language is not only our means for communication but also for representation, so that even the expression of ToM is verbally through language, it filters through into everything that we do [Astington & Baird, 2005; Doherty, 2008].

Some theoretical perspectives take the position that language plays no special part in ToM acquisition and development, beyond being a facilitative instrument for its expression. Modularity theorists (2.2.1) hold the nativist perspective that ToM is innately specified but only evident after the child reaches a certain level of cognitive and linguistic development [Astington & Baird, 2005, p.8]. In this vein, some have argued that ToM tasks are verbal and thus require the subject to meet a particular threshold of language ability in order to succeed on the task [Astington & Baird, 2005]; these are similar to the EF accounts of ToM (discussed in 2.2.4), which posit that a certain level of EF is necessary for successful task performance. While both of these are potentially true, they do not exclude the possibility that aspects of language may be critical in shaping later ToM ability [Sodian & Kristen, 2010]. It may be the case that once implicit intuitions about ToM-related concepts have been established, the terms and means made available by language then function as the refining tool through which an explicit understanding is forged and tempered [Sodian & Kristen, 2010, p.195]. Other theorists have proposed that ToM ability relies on domain-general cognitive processes, whose implementation is supported by language, while further perspectives view language merely as the mode of input, by which children gain the information necessary for the construction of a ToM [Astington & Baird, 2005, p.8]. A number of theorists take the opposite perspective: that language plays a fundamental causal role in ToM development, especially during children's first few years (from 18 months- to 5-6 years-of-age), over the course of which their language, ToM, and EF skills are all rapidly co-developing and interacting [Astington & Baird, 2005; Doherty, 2008; Sodian & Kristen, 2010]. These echo Piaget's idea that language is critical in the development of logical thought and reasoning, because it establishes a bond between thoughts and words, supporting representational capability [Farrant et al., 2006]. Work with atypical populations has shown promising support for this type of perspective but 3.2 will return to this in greater detail.

Before examining the relationships between specific areas of language and ToM, it is important to discuss the nature of pictorial and linguistic representations, as well as the link between meta-linguistic awareness (mLA) and ToM. Language and pictures have been viewed as forms of non-mental representational media, which provide a conduit for ToM and general cognitive development [Doherty, 2008], although this does beg the question: can language really be considered 'non-mental'?. Evidence has also shown that a number of parallels exist between the capacity to reflect upon mental states and the capacity to reflect upon words and images (mLA) [Doherty, 2008]. As previously discussed, having a ToM requires an understanding of mental representation and misrepresentation: in order to understand belief and FB, one must recognise that beliefs represent a possible state of the world [Doherty, 2008; Frith & Frith, 2005]. Understanding that images and words correspond to objects in the world requires a related understanding of the representational nature of pictures and language [Doherty, 2008]. In this way, an understanding of pictorial representation and basic referential language could be equated with the capacity for secondary representations in Perner's three-stage RUM model (2.2.3). This capacity is also necessary in order to engage in pretence (2.3.1) and represent objects or events [Bretherton, 1984]. Understanding the nature of pictures is an important precursor to language ability, as the child must recognise that both make references to an entity or object in the real world, as well as that these references can be represented across multiple instances and in varied ways [Doherty, 2008]. Otherwise, they might be restricted to the concept that individual people (and other entities or objects) physically occur more than once and in different contexts [Doherty, 2008, p.78]. Recognising that a picture refers to something or someone in the real world is not the same as understanding the relationship between the reference and the referent, and the former is a precursor to the latter [Doherty, 2008]. Here, the distinction between the former and the latter could be conceived of as similar to the distinction between supporting primary and secondary representations in Perner's RUM model, which may be relevant to theory theory accounts of ToM (2.2.3). Research has shown that children gradually begin to comprehend the referential nature of pictures from around 9 months-of-age onwards: in experiments, most infants try to pick up the objects depicted by the image, but when the images are paired with the real object, they are twice as likely to touch the real object, [DeLoache & Burns, 1994; Doherty, 2008: referencing DeLoache et al., 1998]. This indicates at least a small recognition of the distinction between them. In DeLoache et al.'s (1998) study, these attempts decreased rapidly with age, with 15- and 19-month-olds much less likely to try to touch the depicted objects, and instead more likely to point at the object in the image, although some 2-year-olds are still occasionally seen making such attempts [Doherty, 2008, p.77]. This is about the time (around 18 months-of-age) when children become able to represent multiple models (secondary representations), which may reflect developments in early communicative ability using gesture, and is relevant to the interpretation of photographs of unique familiar entities or objects [Bates & Dick, 2002; Doherty, 2008]. When a child encounters both a photograph of their parent from another point in time and the same parent in the flesh, instead of assuming that there are numerous clones of their parent, interpreting the picture may be a cognitively similar process to that of recalling a memory [Doherty, 2008, p.77]. Both require the child to recognise that the same person can occur in more than one situation, so they must represent one mental model of the parent ‘in the here and now’ as well as a second mental model of the parent ‘in the picture’ or ‘in the memory’ [Doherty, 2008, p.77]. Research has demonstrated that most children achieve an understanding of correspondence at around 2-and-a-half years-of-age [DeLoache & Burns, 1994; Doherty, 2008]. In this study, children were required to use a picture in order to find a hidden object by applying their knowledge of the object’s location within the picture to the real situation [DeLoache & Burns, 1994]. Very few 24-month-olds (one in eight) managed to perform this task successfully (as opposed to 72% of 30-month-olds), whereas with element-to-element correspondence in scale models, 30-month-olds were unsuccessful while 36-month-olds performed successfully [DeLoache & Burns, 1994; Doherty, 2008]. Failing these tasks may be due to a lack of awareness or understanding that the picture and real scenario are relevant to one another, or to difficulties with drawing inferences [Doherty, 2008]. An understanding of geometric correspondence within scale models develops even later, with 3-year-olds performing at chance, 4-year-olds reasonably well, and 5-year-olds at ceiling [Doherty, 2008: citing Blades & Cooke, 1994, and Newcombe & Huttenlocher, 2000]. It may be significant that children begin to understand geometric correspondence at the same time that they pass the FB task [Doherty, 2008].

An understanding of the representational nature of language, or possessing meta-linguistic awareness (mLA), has been likened to the third and final stage of Perner’s RUM model: meta-representation [Doherty, 2008]. There are a number of conflicting perspectives about what actually constitutes mLA, but Doherty collated and presented three criteria: “(1) the ability to reflect on language as an object”; “(2) the ability to control, monitor, and plan linguistic processing”; and “(3) consciousness, awareness, or intentionality” [2008, p.81]. Doherty proposed that the second criterion equated to the EF aspects of language and that if the criterion of ‘consciousness’ were excluded, then the characterisation of mLA closely paralleled that of ToM [2008, p.81-82]. He argued that this was the case because: (a) consciousness is problematic to define, given that it is not yet fully understood or evaluable, and (b) ToM involves reflecting on mental states while mLA involves reflecting on language, both require EF for their expression and, potentially, emergence, and thus may only differ from one another in their domain of application [Doherty, 2008, p.82]. In order to pass the ToM FB task, children are required to recognise that mental states represent a potential state of the world, which is, in the FB scenario, different from the objective reality [Doherty, 2008]. To do so, they must possess the meta-representational capacity to distinguish between what and how something is represented [Doherty, 2008, p.82]. Language also represents situations in a particular way, which signifies that mLA requires meta-representation [Doherty, 2008, p.82].

3.1.1 Lexical Acquisition and Semantic Development

It is difficult to tease apart the various areas of language in order to relate them to ToM, because they all interact with one another to some extent; however, some theorists propose that children's acquisition of the meaning (semantics) of words is responsible for their ToM development [Astington & Baird, 2005; Doherty, 2008]. The accounts within this area take varying stances, ranging from: the more direct, which point to specific lexical items used to refer to the mind, mental states, and mental activities; to the more general, which propose that children acquire their mental-state concepts in and through discourse [Astington & Baird, 2005]. The latter accounts assert that it is discourse which provides the means to abstract these concepts, because these are semantically encoded within a culture's language [Astington & Baird, 2005]. Looking first to the more direct end of the theoretical spectrum: it is true that the meanings of mental state terms are significant for ToM development in a number of ways. The process of associating a linguistic term with a mental state may help to direct attentional focus to the mental state in question, as well as provide the means to identify and distinguish between both similar and dissimilar mental states [Doherty, 2008]. This type of linguistic distinction, for example, might be relevant to young children's developing analyses of pretence and belief, the behavioural expressions of which share many surface characteristics [Doherty, 2008]. Language is fundamentally a tool for interaction, so if children are exposed to mental-state terms within dyadic and triadic contexts, where the same words apply to both their own and others' mental states, this may enable them to map the behaviour and experiences of others to their own cognitive and affective experiences, and vice-versa [Astington & Baird, 2005; Doherty, 2008].

It is possible that children first acquire mental-state concepts before learning to properly use the associated terms, but evidence from the research suggests the opposite: that children begin to use the terms properly at around the time they begin passing ToM tasks [Astington & Baird, 2005; Doherty, 2008]. Children start to acquire these specific mental-state terms much earlier at around 2-3 years-of-age, beginning with perception ('see', 'look'), emotion ('happy', 'sad'), and desire ('want'), before progressing to terms of cognition ('know', 'think', 'remember') [Astington & Baird, 2005, p.9]. It has been argued that children's experiences with such terms in interactive contexts are what guide them in developing a conceptual understanding of unobservable mental states, and that this process takes considerable time [Astington & Baird, 2005]. Children's early use of mental-state terms suggests that they do not fully comprehend the associated concepts, and that they gain and refine this understanding of meaning through continued exposure [Astington & Baird, 2005; Doherty, 2008]. This might in turn foster their abilities to reflect upon and identify the mental states of themselves and others [Astington & Baird, 2005]. In studies conducted with toddlers, 2-year-olds frequently used desire terms but references to cognitive-state terms were rare and produced with significantly less frequency than other internal-state terms: only a third made references to their own or others' thoughts [Doherty, 2008, p.154: citing Bretherton & Beeghly, 1982]. However, later analysis of the contexts in which children used the language showed that these were not genuine references to mental states [Doherty, 2008]. For example, such references to mental-state terms like 'know' and 'think' were typically made within the chunked phrase "I don't know" or in sentences with "I think…" to express possibilities, disagreements or uncertainty [Doherty, 2008, p.154]. Other research utilising the CHILDES database has also found that, while genuine references to desire occur in late infancy, genuine references to belief only begin to occur from 3 years-of-age, rising steadily, and stabilising once children are around 5 years-of-age [Doherty, 2008, p.154-155: citing Wellman & Bartsch, 1994]. These trends within naturalistic conversations parallel those from the experimental data on children's awareness of mental-state concepts, suggesting that children begin to make genuine references to belief from around the time that they begin to perform successfully on the FB task [Doherty, 2008, p.155]. It is worth noting here that cross-linguistic research has revealed that different languages explicitly distinguish mental states to varying degrees, and that this explicit linguistic marking does not appear to significantly enhance children's understanding of the concepts [Doherty, 2008]. Any small advantages on FB test questions using verbs that explicitly marked beliefs as false appeared to be locally limited, and did not seem to carry over into later relevant task questions [Doherty, 2008]. Even so, this might be accounted for by the idea that a certain level of executive functioning is necessary to support both language and ToM capabilities, and that while language may yet prove to be crucially important in the ToM conceptual development process, the children at the age of the participants in the study might still need to catch up in terms of their EF development (see §5 for discussion of a tripartite model). Just because EF may be necessary in ToM development, it does not preclude language from also playing a critical role. Another possibility is that children may actually have a different understanding of these terms from adults: they may represent these explicit linguistic markings in the same way that they might tag a pretend scenario (2.3.1), associating the terms with actions that do not match the situations in which they occur [Doherty, 2008, p.157].

Regarding children's more general semantic development, evidence from the research has consistently shown that children's vocabulary levels and VMA correlate with their ToM ability [Astington & Baird, 2005; Doherty, 2008; Happé, 1995a]. Some have argued that this association is purely down to the fact that children with higher mental ages perform tasks more successfully than those with a lower mental age, but studies have demonstrated a strong relationship between vocabulary scores and FB task performance in both typically developing and autistic children (§4) even after factoring out mental age statistics [Doherty, 2008; Happé, 1995a]. It is also true that mental states are unobservable abstract entities, into which we can only gain insight through the second-hand observation of clues obtained from other people's facial expressions, speech, and behaviour [Astington & Baird, 2005, p.14]. Barsalou grounded conceptualisation in sensorimotor, bodily, physical environmental, and social environmental experiences [2016]. Concepts do not map one to one onto specific behaviours or situated conceptualisations and, as previously discussed, it remains possible that language could assist children in identifying and separating out these interrelated concepts [Astington & Baird, 2005; Barsalou, 2016; Doherty, 2008]. This linguistic influence on the conceptualisation of mental states might occur within the ostension paradigm, where children have mentalistic experiences that they then associate with and map to mental-state terms, or within the contextual paradigm, where children derive and construct the meaning of these terms from their pragmatic usage and the functions they serve in everyday social contexts [Astington & Baird, 2005, p.14: citing Montgomery, 2005; Barsalou, 2016]. Others take the view that while comprehension of mental-state concepts may emerge from communicative experiences, this emergence transpires via language supporting and inviting logical reasoning and inference about the behaviour of others, much in the way that it might during the developmental fast-mapping process of acquiring terms for physical objects [Astington & Baird, 2005: citing Baldwin & Saylor, 2005]. These build on Gricean principles concerning a speaker's intentional focus, as well as their communicative and referential intents [Astington & Baird, 2005]. Taking a different position, Jacques & Zelazo reframed ToM as cognitive flexibility in PT and the ability to support multiple (secondary) representations, with language serving a labelling function within this process, enabling self-reflection through psychological distance from the concept itself [2005; Astington & Baird, 2005].

Astington & Baird investigated whether linguistic manipulations of the mode in which FB tasks were presented (only verbal and only visual conditions vs. the standard verbal and visual condition) would affect children's performance, but found no effect of condition [2005]. This is interesting, as it does not indicate a specific effect of language on performance, yet does not preclude their conclusions regarding a more general role of language in supporting meta-representational capacity [Astington & Baird, 2005]. To this end, evidence from the research has been somewhat corroborative: some findings suggest that early language ability predicts later ToM performance while others indicate a more bidirectional interaction [Nilsson & De López, 2016; Sodian & Kristen, 2010].

When examining this meta-representational association between language or mLA and ToM, it is worth considering Doherty & Perner's (1998) work on synonymy and Doherty & Wimmer's (2005) work with Jastrow's ambiguous figures [referenced in Doherty, 2008]. Doherty & Perner predicted that mLA would emerge at around 4 years-of-age, similarly to ToM FB understanding [Doherty, 2008]. They argued that simulation and modular theories did not 'naturally extend' to mLA, and that EF accounts faced similar limitations, because such accounts would contend that the expression of ToM and mLA are subject to comparable EF demands, when the domains are different, so their tasks would not pose the same EF challenges [Doherty, 2008, p.82]. To investigate their predictions, they examined how children understood synonymy, proposing that to support such an understanding, children would need to recognise the relationship between linguistic forms and what they represent, as well as that two forms could represent the same object or entity [Doherty, 2008, p.82]. As a result of this requirement to distinguish between what and how something is represented (necessitating meta-representation), they proposed that understanding synonymy and FB might therefore be parallel processes [Doherty, 2008]. For this study they developed tasks deploying synonym pairs familiar to young children (e.g. lady/woman, cup/mug, coat/jacket, bunny/rabbit, etc.), where in one task children had to produce synonyms, and in another, judge whether or not another person had produced one correctly [Doherty, 2008, p.83]. In their analysis, they compared performance on these tasks with performance on FB tasks, and found them to be both of equivalent difficulty and strongly associated even after controlling for VMA [Doherty, 2008, p.84]. Doherty acknowledged that the synonym production task might have posed some of the same EF inhibitory demands as the FB task, but that the synonym judgment task did not do so and was even more strongly correlated with the FB task [2008, p.84]. While there is some debate surrounding the existence of true synonyms (relevant to the mutual exclusivity bias theory), Doherty further argued that their analogous research with homonyms and lexical hierarchical relations found the same strong associations with FB task performance after controlling for VMA [2008]. This indicated that understanding lexical overlaps (difference in form but equivalence in meaning or vice versa) required meta-representation [Doherty, 2008, p.85]. Doherty and Wimmer proposed that Jastrow's ambiguous figures (e.g. the duck-rabbit) were visually analogous to homonyms, as their interpretations both depend highly on context [cited in Doherty, 2008]. Previous work had found that children tended to experience reversals of the pictures a year after they tended to pass the FB task, which was taken as problematic for theories suggesting that a general recognition of meta-representation emerges at around 4 years-of-age [cited in Doherty, 2008, p.86]. This work hypothesised that in order to experience reversals the subject would need to be aware of the ambiguity (the existence of the two possible interpretations or forms), which would require meta-representation, however Doherty & Wimmer theorised that it might also require certain EF abilities not yet available to 4-year-old children [cited in Doherty, 2008, p.87]. Their work found that individual children were able to report the two possible interpretations at the same time they could pass the FB task, and that the two abilities were strongly correlated [Doherty & Wimmer, 2005; cited in Doherty, 2008]. Wimmer's later (2007) work established that a certain level of EF inhibitory control was necessary to experience reversals with the ambiguous figures [cited in Doherty, 2008]. Research on autism and visual illusions is also relevant here, as well as work on weak central coherence theory (WCC) (see 4.4.1) [Frith & Happé, 1994; Happé & Frith, 2006]. Doherty argued that his and Wimmer's findings demonstrated that developing the ability to understand the representational relationship between a figure and its referents (meta-representation) was associated with similar developments in ToM and mLA, and that this provided support for the theory theory account of ToM [Doherty, 2008]. Thus, specific aspects of vocabulary and semantic development may not be necessary for ToM development, yet more general meta-representational language capabilities may be necessary [Doherty, 2008]; this is relevant to theories concerning the roles of conversational pragmatics and discourse, which will be discussed in more detail in 3.1.3 and 3.1.4.

3.1.2 Syntax

In a similar way to those who theorise that specific aspects of lexical semantics can account for ToM development, there are also those who propose a specific role of syntax: the grammatical component pertaining to the structure of language. It is important to bear in mind, however, that syntax is inherently linked to semantics, and that it is this interface between meaning and a technical system of rules, rather than the individual components, which may be significant for ToM acquisition. Some authors have suggested that the ability to support and use complex grammatical structures (such as complementation) might be necessary in granting children the cognitive tools with which to reason about and judge the truth or falsity of others' belief contents [Doherty, 2008; Durrleman & Franck, 2015]. Cross-linguistically, mental-state verbs and expressions typically take complex forms and word orders, where propositions are embedded within sentences; becoming proficient with these constructions may be required to represent and keep track of the mental states of others in contrast to states of reality [Doherty, 2008, p.153; Nilsson & De López, 2016]. These kinds of abilities may be reflected in or supported by syntactic development [Doherty, 2008].

Sentences with subordinate clauses or 'sentential complements' functioning as their grammatical objects often feature a mental verb as the main verb, and these constructions allow a speaker to break down propositional attitudes into the proposition (the contents of the mental state) and the attitude (the type of mental state) [Astington & Baird, 2005, p.7; Doherty, 2008]. These sentences can remain true even when a false sentence is embedded within them [Astington & Baird, 2005], for example: "Sophie thinks that the unicorn is in the garden.", when it is actually in the stable. Whereas if someone stated, "The unicorn is in the garden.", and another person replied, "It is in the stable.", one of the sentences would be objectively true and the other objectively false. With sentential complementation, however, both sentences could be true [Astington & Baird, 2005], as in the case of: "Sophie thinks that the unicorn is in the garden and Guy knows it is in the stable.". In this way, complementation syntax may provide us with the necessary structure to support FB representation, and attribute different mental states to other people [Astington & Baird, 2005; Doherty, 2008; Nilsson & De López, 2016]. Children start to use these constructions from 2 years-of-age (soon after they begin producing mental-state verbs), but early use has been shown to be formulaic; true complementation is not mastered (and genuine references to belief are not made) until 1-2 years later [Astington & Baird, 2005, p.10; Doherty, 2008]. Some research has indicated that this mastery predicts FB task performance, which has been claimed to be evidence that FB understanding is supported by syntactic development [Astington & Baird, 2005; Doherty, 2008: both citing de Villiers & Pyers, 2002; Nilsson & De López, 2016]. The verb 'say' has been identified as a potential bootstrapping agent for acquiring an understanding of complementation, as its flexible and diverse application precedes that of mental verbs [Astington & Baird, 2005]. This is particularly pertinent to the verb 'think' as they are used in similar discourse contexts, and the truth value of the propositional content of ‘say’ complements is more tangibly falsifiable for young interlocutors [Astington & Baird, 2005].[6]

One of the most prominent accounts advocating the role of complementation syntax in ToM development was that of (J.) de Villiers’ ‘linguistic determinism’ [1995, 2005: referenced in Astington & Baird, 2005, and Nilsson & De López, 2016; 2000, 2002: referenced in Doherty, 2008]. This account emphasises the role of verbs which take realis object complements (propositions concerning actual events), e.g. belief and communication verbs, as opposed to verbs which take irrealis object complements (propositions concerning imaginary or future events), e.g. verbs of desire {want + infinitive in English, want-that in German} or pretence {pretend-that}, and general tensed object complements {that + finite verb} [in Astington & Baird, 2005]. De Villiers took this perspective because she asserted that belief and communication verbs specify a (partly semantic-criterion) point-of-view (POV) marker, which is then tagged onto their complement clauses [2005: in Astington & Baird, 2005]. In this way, de Villiers differentiated verbs of belief from verbs of desire or pretence, arguing that sentential complements of desire- and pretence-verbs have intensional noun phrase complements [in Astington & Baird, 2005]. She proposed that children extended their recognition of irrealis objects to all desire verb complements, but treated belief-verb (‘think’) complements as realis clauses from the outset, recognising that false propositions can be embedded within these clauses [in Astington & Baird, 2005]. This could be because children have already been exposed to the analogous communication-verb ‘say’, which can take false complements where the truth values of their propositional contents are obvious to the child, reflecting either what they know to be the case, or not [Astington & Baird, 2005].

However, views which designate syntax and sentential complementation as driving factors of ToM development seem implausible, as they discount the importance of mental-state concepts. Semantics clearly also plays a role. Although increased syntactic development may facilitate further conceptual development, syntax is not possible without semantics, just as representation is not possible without underlying concepts [Barsalou, 2016]. Thus, the role of language in early ToM acquisition more likely reflects their integration, in addition to other developments in general language and meta-representational capacity [Astington & Baird, 2005; Doherty, 2008; Lohmann et al., 2005]. These are largely the conclusions that can be drawn from the research literature. A number of cross-linguistic studies have challenged de Villiers’ theories, demonstrating that children understand desire before belief, regardless of whether they speak languages requiring

[6] This theory is important to consider when examining proposals that people on the autistic spectrum ‘hack out’ mentalising tasks using language (see §4 and §5) [Happé, 1995a; Tager-Flusberg & Joseph, 2005; Tager-Flusberg, 2007], as they might rely on syntactic cues to do so.

tensed that-complements for both desire and belief (such as German), or languages that do not require tensed that-complements for either (such as Mandarin) [Perner et al., 2003: cited in Doherty, 2008; Perner et al., 2005: in Astington & Baird, 2005]. These accounts take the position that the syntactic forms for the way we talk about the mind do not necessarily structure how we think about it [Astington & Baird, 2005, p.18]. Astington & Baird, however, made the important point that the two stances (Perner and colleagues vs. de Villiers) do not necessarily preclude one another, as their differences are chiefly rooted in how one perceives the relationship between syntax and semantics: as separate units or as an integrated whole [2005, p.18; Griffin & Baron-Cohen, 2002]. Work by Lohmann et al. has supported this more combinatorial and interactive perspective on syntax and semantics or discourse pragmatics [2005]. They demonstrated in a training study that both conversation about deceptive objects and training in complementation syntax improved 3-year-olds' FB understanding, but that the largest improvement effect occurred in the double condition, which they suggested was an indication that each condition contributed individually to ToM development [Lohmann et al., 2005; Astington & Baird, 2005]. They also found that manipulating objects in the absence of contextual discourse led to no training effect, which they proposed was evidence for a causal role of language in the advancement of social understanding [Lohmann et al., 2005; Astington & Baird, 2005].[7]

As discussed by Doherty, there are a number of reasons why it is unlikely that sentential complementation has to be mastered before any form of ToM development can proceed [2008]. Firstly, is difficult to prove whether syntactic or semantic development occurs first, as the syntax may develop for productive purposes, i.e. to express concepts which are already understood [Doherty, 2008]. Secondly, the aforementioned work by de Villiers & Pyers, which proposed that mastery of complementation syntax predicted FB task performance, was based upon children's ability to remember false complements [2002: cited in Astington & Baird, 2005; Doherty, 2008]. The task design was arguably very similar to that of a FB task, participants were required to remember and then report a FB, which would be problematic for subsequent, comparative analysis of the two tasks' data [Doherty, 2008]. It was argued by de Villiers & de Villiers (2000: cited in Doherty, 2008) that the false complement task did not require participants to attribute FBs (only remember them), and for this reason the task might be less demanding [Doherty, 2008]. Doherty asserted, however, that in order to extract the propositional content from the sentence or reconstruct this proposition in memory, children may have to represent the character's belief as

[7] This is also relevant to the discussion of autism and the development of clinical interventions (see §4 and §5).

well as understand FB [2008, p.158], which may pose extra demands on working memory. Doherty also pointed out that with a different analysis of the data, comparing only the performance data from the unexpected transfer FB task (2.3.1) with that of the memory-for-complements task, FB performance was consistently better, and early performance on the unexpected transfer FB task was a stronger predictor of later performance on the memory-for-complements task [2008, p.159]. From this angle, it appears that basic FB ability precedes or emerges simultaneously with the ability to remember false complements [Doherty, 2008, p.159]. This analysis reiterates the conclusions drawn above; it is perhaps more likely that syntax and semantics share a complex, staged, and interactive process of development [Doherty, 2008; Lohmann et al., 2005]. Thirdly, in line with the aforementioned work of Perner and colleagues, children's problems with complementation syntax only appear to extend to complements of belief or communication verbs, and not to verbs which do not require an understanding of misrepresentation [Doherty, 2008]. Therefore, it is probable that difficulty representing belief is more than a purely syntactic issue [Doherty, 2008]. Interestingly, Sodian & Kristen discussed a number of behavioural and neuroimaging studies, which suggested that adults process ToM more verbally than children, but also other studies of late-acquired aphasia (and loss of grammatical skills) demonstrating that adult ToM functions without the relevant syntactic structures [2010]. This indicates that syntactic aspects of language may be significant for a developing ToM but not necessarily for its expression in adulthood [Sodian & Kristen, 2010].

3.1.3 Pragmatics

Another relevant area for the examination of the relationship between ToM and language is pragmatics. Some authors assert a specific association between pragmatics and ToM, proposing that ToM development is responsible for pragmatic language competence [Happé, 1993, 1995b]. Others view pragmatics as the language component of ToM, arguing that they are so closely related, they are almost one and the same [Astington & Baird, 2005; Doherty, 2008]. There are also those who view pragmatics and ToM as related but independent processes, with pragmatics more rooted in semantic skills [Norbury, 2005], or as individually contributing components (alongside others) to specific aspects of pragmatic language competence, such as figurative language (FL) comprehension [Chahboun et al., 2016a,b]. Despite the controversy over the particulars, pragmatics and ToM are intricately interconnected.

Pragmatics refers to the appropriate use of language within communicative contexts and entails the application of general world knowledge to language [Astington & Baird, 2005; Doherty, 2008; Frank et al., 2015]. It requires interlocutors to make inferences based upon their world knowledge, and integrate their communication within a context, going beyond the literal interpretations of individual utterances to work out what a speaker is actually trying to convey, as well as gauge what a listener might infer from a speaker's utterance [Astington & Baird, 2005; Doherty, 2008; Frank et al., 2015]. In order to accomplish this, one must keep track of the mental states of everyone involved in a conversation, which is why Doherty argued that pragmatics is the application of ToM to language [2008, p.152]. Some communication is executed for the direct purpose of altering someone's behaviour, but in general, communication involves the exchange of information with the goal of influencing the mental states of others [Doherty, 2008, p.152]. This latter type of communication requires ToM, where the former does not, and it has been argued that complex languages would not have evolved without ToM, because the value of language would otherwise be limited [Doherty, 2008, p.152]. In this sense, ToM ability has been claimed to underlie the capacity for language, as the awareness and understanding of beliefs and intentions, which can be monitored and tracked through pragmatic language competence, are central features of ToM [Astington & Baird, 2005, p.13: referencing O'Neill, 2005]. Pragmatic language competence is also important in storytelling, because one must be aware of an audience's informational needs in order to tell a story, and one must understand things about the social world in order to understand a story [Barnes & Baron-Cohen, 2012]. There is also some evidence from the research suggesting that men and women comprehend and process pragmatic aspects of language and ToM differently, which has implications for our understanding of autism and the expression of empathy [Frank et al., 2015] (more in §4).

Happé discussed the relationship between pragmatics and ToM with reference to Sperber & Wilson's (1986) relevance theory, which she argued made the role of understanding intentions in human communication explicit [1993, p.101; 1995b]. The ability to recognise people's communicative intentions is especially important in the case of FL, e.g. simile, metaphor, and irony, many of which may not be understood in terms of their literal interpretations [Happé, 1993, 1995b]. Happé went on to examine the predictions made by relevance theory with regards to the level of FL competence achievable within no, first-order only, or second-order ToM ability experimental conditions [1993, 1995b]. Here, Happé established theoretical parallels between these levels of ToM ability and the levels of non-literal interpretation required to understand FL, in line with relevance theory [1993, 1995b]. From this perspective, similes can be interpreted wholly literally, as the inclusion of the word 'like' signals to the recipient of an utterance that they need to decide the degree to which two referents are similar [Happé, 1993, p.103]. In this account, being able to understand a simile would not require ToM, and no theoretical difference would exist between: “She was like a statue.” and “She was like her mother.” [Happé, 1993, p.103]. In contrast, Happé asserted that information about the speaker’s mental state and some understanding of intentions are vital to comprehending an utterance with a metaphor, because the propositional form is, in essence, an interpretation of the speaker’s thought [1993, p.104; 1995b]. Thus, the default literal interpretation of a metaphor would not appropriately reflect the speaker’s intended meaning, and understanding a metaphor would require at least first-order ToM [Happé, 1993]. This is similar to FB tasks, where objective reality is not helpful, but recognising the mental states of the characters involved is key for successful task performance [Happé, 1993]. Finally, Happé’s account asserted that understanding irony requires second-order meta-representational ToM because it involves thought about attributed thoughts [1993, p.104; 1995b]. Citing Sperber and Wilson (1981, 1986), Happé posited that ironic utterances quote or refer to an attributed thought by mentioning a possible thought and expressing an attitude towards it [1993, p.104]. For example, if someone said: “Well, that is just brilliant!”, while actually being unhappy about an event, then the speaker is mentioning the possible thought of being happy, while expressing a sardonic attitude towards it. Happé also discussed classical theoretical accounts of irony, which hold that ironic utterances are first interpreted literally, then examined for suitability within their given context, and then if this interpretation is not appropriate to the context, the reverse meaning is accessed by the listener [1993, p.104: citing Grice, 1975]. Happé argued that this view does not separate irony and metaphor in terms of the ToM demands involved and it thus predicts that only first-order ToM is required to understand irony [1993]. In Happé’s view, someone who could only attribute first-order mental states would be able to recognise the ‘informative intention’ of a speaker, but could only access their ‘communicative intention’ with second-order mental-state attribution [1993, p.102: referencing Sperber & Wilson, 1986].

Being able to access the communicative intentions of other people through language is critically important within everyday social interactions. A number of studies have shown that even with high verbal ability in high-functioning autism (HFA), difficulties persist with non-literal FL, e.g. joking, metaphor, irony, indirect requests, etc. (§4) [Chahboun et al., 2016b; Doherty, 2008; Happé, 1993, 1995b; Norbury, 2005; Vulchanova et al., 2015]. Happé investigated this relationship between relevance theory, FL, and ToM in autism, arguing that autistic people adopt a default literal interpretation of an utterance when attempting to access its propositional content [1993]. The conclusions drawn from this study’s data demonstrated support for Happé’s account proposing hierarchical links between ToM and FL competence [1993, 1995b], but these will be examined in more detail below.

Contradicting Happé's assertions, Norbury argued that while ToM ability might be necessary for metaphor comprehension, it is not sufficient, proposing instead that the understanding of metaphors was rooted in semantic ability [2005]. Results from her study indicated that being able to employ first-order ToM did not guarantee metaphor comprehension and that semantic skills were a stronger predictor of successful performance on a metaphor task [Norbury, 2005]. To better understand their claims, it is worth comparing the design of the two studies. Norbury's (2005) study involved a greater number of participants than Happé's (1993) study, and while Happé examined competence with similes, metaphors, and irony, Norbury only investigated competence with metaphors. Happé's (1993) study pre-grouped autistic participants according to performance on a battery of first- and second-order ToM tasks, with 6 participants in each ToM-level group, and included a control group of 14 participants with 'moderate learning difficulties' (MLD) [1993]. Norbury recruited 94 participants with "communication impairments" (SLI = 19; primary pragmatic language impairment (PLI) = 15; PLI and 'autistic behaviours' = 24; HFA = 18; 'Asperger's disorder' = 18), who she regrouped for consistency according to autistic status and language ability, in addition to 34 typically developing, age-matched peers [2005, p.386-387]. Norbury's regrouped participants fell into four categories: language impairment (LI = 28), autistic spectrum and language impairment (ASL = 31), autistic spectrum only (ASO = 29), and pragmatic difficulties without other language impairment or autistic symptoms (n = 6) [2005, p.388]. Norbury's participants ranged from 8-15 years-of-age, although her control group trended younger than the ASL group, while the age-ranges of Happé's participant groups overlapped with Norbury's but tended to be higher and were notably larger. While matched for other forms of IQ, Happé's groups were not matched for verbal IQ (VIQ) as each higher ToM stage group had a higher mean VIQ (no-ToM = 62.3; first-order ToM = 81.5; second-order ToM = 89.5), apart from the MLD group's mean VIQ of 55.6. In Norbury's study, the language-impaired ASL and LI groups had lower non-verbal IQ scores than the ASO and control groups, and were not measured in terms of VIQ but 'language scores', on which the LI and ASL groups were significantly poorer than the ASO and control groups. Furthermore, the autistic symptomatology was greater in the LI than control group, which is interesting in the context of Griffin & Baron-Cohen's (2002) proposal that semantic-pragmatic language disorder might be part of the autistic spectrum (4.3). Norbury found that only the children with language impairment, regardless of autistic symptomatology, were impaired on a metaphor task [2005, p.383]. This finding is intriguing, and if we follow Doherty's (2008) and

Astington & Baird's (2005) reasoning that pragmatic language competence and ToM are two sides of the same coin, then it could be argued that Norbury's (2005) account advocates language as being more influential upon (pragmatic competence and) ToM development than the other way around. Norbury's account is, however, limited in that she only discusses and contrasts theoretical perspectives on simile and metaphor with no mention of irony, particularly as a number of studies have shown that people with HFA can develop strategies for and succeed on first-order ToM (and VPT1) tasks, but struggle with second-order ToM (and VPT2) tasks (more in §4).[8]

It is also important to consider the developmental context when attempting to tease apart the relationship between pragmatic language and ToM. Acquiring FL occurs later than general language acquisition (e.g. vocabulary), accelerating from the first years of formal instruction at school, flattening out from 10-12 years-of-age but extending throughout adolescence into young adulthood, and many researchers agree that understanding FL requires mLA [Chahboun et al., 2016a,b; Norbury, 2005; Vulchanova et al., 2015]. Instances of FL span from singular words to whole sentences, and many factors are involved in the process of their comprehension and acquisition, including: frequency of usage, structural transparency, context of the encounter, and general linguistic competences [Vulchanova et al., 2015, p.2]. It has also been shown that the ability to provide word definitions facilitates FL comprehension [Chahboun et al., 2016a: citing Levorato & Cacciari 2002]. However, a number of issues make it difficult to clearly define the influence of semantics and lexical acquisition on pragmatic language competence and ToM. These stem from wider theoretical debates about conceptualisation, the lexicon, conceptual and lexical storage and retrieval (including where and how FL is stored, processed, and accessed), which are beyond the scope of this thesis. Returning to a developmental focus, research has shown that idioms (theoretically viewed as 'dead' metaphors) appear to be acquired gradually and non-linearly over an extended period of time, peaking at 11 years-of-age, with contextual sensitivity demonstrated only from 9 years-of-age [Chahboun et al., 2016a,b; Vulchanova et al., 2015, p.2]. The ability to understand similes and metaphors requires recognition of the comparison being made, which in turn requires sufficient world knowledge and broad semantic representations, as the relevant features of the referents being compared may not be their most salient [Norbury, 2005, p.384]. This could also require a certain level of inhibitory EF in order to suppress the salience of the irrelevant features, as well as switching to select and enhance the relevant features for comparison. Vulchanova and colleagues considered this sort of EF account,

[8] These accounts are also fundamentally based upon the idea that ToM is actually impaired in SLI and autism, although, as will be discussed in 3.2.1 and §4, this does appear to be the case.

arguing that FL comprehension goes beyond structural language competence, involving wider skills such as: inferencing, multimodal information integration, inhibition, etc. [2012a, 2012b; Chahboun et al., 2016a]. In this sense, the contexts underlying metaphorical language are crucially important and may affect how well a metaphor is understood; there are also various types of metaphor: (1) attributive (a comparison based on shared physical properties); (2) relational (a comparison involving abstract knowledge about non-physical properties, e.g. function); and (3) double (a comparison can be made both attributively or relationally) [Norbury, 2005, p.385]. These type differences may affect their salience. Young children have consistently been shown to find similes easier to comprehend than metaphors, probably because the word 'like' provides an additional, explicit syntactic cue [Happé, 1993; Norbury, 2005]. Research has also shown that irony is consistently misunderstood below 6 years-of-age, with unreliable performance extending as far as 13 years-of-age, with many children misjudging it as deception [Happé, 1993]. These older age-ranges for acquiring figurative language overlap with and parallel those for second-order ToM and mLA [Happé, 1993] (2.3.1), which is significant for meta-representational accounts of ToM, such as language and theory theory accounts.

Chahboun et al. conducted a study investigating linguistic factors, such as: structural language, receptive vocabulary size, conceptual reasoning, and semantic skills, in order to ascertain the extent to which each factor predicted the performance of participants with HFA on a task contrasting various metaphors (i.e. conventional vs. novel) with literal expressions [2016a]. They also aimed to pinpoint the role of ToM in FL comprehension, but found their task limited to this end [Chahboun et al., 2016a]. The authors argued that conventional metaphors are similar to idioms and other lexicalised expressions, meaning that they can be retrieved from the lexicon or long-term memory, whereas novel metaphors are more dynamic, requiring online processing [Chahboun et al., 2016a, p.29]. If figurative language is the application of ToM to language, then this online processing might require ToM, or it might rest upon semantic skills which facilitate ToM development. They found that their results in the ASD groups were consistent with those of Norbury (2005), where semantic skills determined metaphorical comprehension [Chahboun et al., 2016a, p.44]. However, the results in the control group lacked this association, and no associations with basic vocabulary and receptive grammar skills were discovered either [Chahboun et al., 2016a, p.44]. Chahboun et al. further found that vocabulary size negatively correlated with accuracy in the child-aged controls, which they presented as evidence for an interference between an expression's metaphorical nature and the accessibility of a figurative interpretation in cases where a word shared only a literal relationship with the prime [2016a, p.45].

Consequently, they proposed that autistic participants experienced a delay (or different trajectory) in the acquisition of FL competence compared to typically developing controls [Chahboun et al., 2016a, p.45]. They reasoned that the autistic participants were still in the stage of acquiring the (meta-)linguistic skills and competences needed to process metaphors, whereas the controls had already passed this point, with their basic language skills no longer impacting upon their FL processing [Chahboun et al., 2016a, p.45]. These findings are interesting, given that research on autism has revealed similar delays in ToM and FB performance [Happé, 1995a] (more in §4). Chahboun et al. further discussed evidence from the wider language acquisition literature, which supports the model of a changing force in predictors over the course of development and the skill in question, as well as the changing dimensionality of language abilities over time [2016a, p.45]. This is significant, as this idea parallels the aforementioned neuroimaging findings discussed by Sodian & Kristen (2010), which demonstrated that adult ToM can function in the absence of syntactic ability (as evidenced in late-acquired aphasia). Sodian & Kristen argued that these findings were an indication that, while language might be important for the development of ToM, it was not critical for its mature expression [2010]. It may be that a number of aspects and processes of FL and ToM become automatised as the typically developing population ages [Happé & Frith, 2014], with language ability vital for their early acquisition but diminishing in importance over time. This theory bears interesting implications and will be returned to in subsequent sections.

3.1.4 General Language Development and the Role of Discourse

While it may be difficult to pinpoint any contributions made by specific linguistic features to ToM development, it does seem increasingly likely that language as a whole plays some crucial part in the process. This perspective has been voiced by a growing number of authors, and it may be the case that semantics, syntax, and pragmatics all have a role to perform in ToM acquisition. Pragmatic competence facilitates children’s participation in communicative interactions, within these interactions they encounter mental terms framed by sentential complements, from which they develop an awareness of different perspectives and mental-state concepts, as well as acquire increasingly complex syntactic structures for the representation of FB [Astington & Baird, 2005, p.11; Nilsson & De López, 2016]. Lohmann et al. also took this stance, proposing that the relationship between language and ToM assumes different forms and directions contingent on the matter and developmental stage at hand [2005; Astington & Baird, 2005]. They asserted that the recognition of other people as agents with communicative intentions is a prerequisite for language acquisition, but that the reverse is also true: using language to participate in social interactions facilitates the appreciation of other people as agents with mental intentions that are rooted in their subjective desires and beliefs [Lohmann et al., 2005; Astington & Baird, 2005]. Therefore, in their view, language and ToM take turns at further promoting each other's progression over the course of development. Doherty propounded similar views, arguing that specific linguistic features (e.g. sentential complementation, etc.) might not be the driving force of ToM development, but that they could augment and refine it, either directly (as Astington & Baird, 2005; or Lohman et al., 2005 suggested) or through the direction of children's attention towards mental states [2008]. Some forms of conversation may also implicitly, or explicitly, invite children to engage in PT [Harris, 2005; Astington & Baird, 2005]. Griffin & Baron-Cohen (following the ideas of Dennett) also discussed the generative importance of language in allowing people to share the reasoning and causality behind the behaviour of themselves and others, which in turn further cultivates their abilities to predict and explain future behaviour [2002]. They considered this within the framework of cultural storytelling, referring to Dennett's 'freefloating rationales', which are reasons for behaviour that are not explicitly apparent to an individual organism but implicit in the context of a larger system [Griffin & Baron-Cohen, 2002, p.103]. Following Dennett's philosophy, their argument presented language as a tool for rendering intentionality and mental states explicit [Griffin & Baron-Cohen, 2002]. When communicating with someone, we are forced to frame concepts in concrete linguistic terms in order to convey them to the other person, which might then influence our own cognitive profiles [Griffin & Baron-Cohen, 2002].[9]

Language certainly wields some considerable influence in the emergence of ToM, and it is therefore noteworthy that early language ability has been found to predict later ToM ability but not vice versa [Doherty, 2008: referencing Astington & Jenkins, 1999, and Ruffman et al., 2003; Nilsson & De López, 2016] (briefly discussed in 3.11). Milligan et al. carried out a meta-analysis of data from 104 studies and 8891 children, concluding that language and FB abilities were strongly associated (regardless of children's age), their analysis covering general language, semantics, syntax, memory for complements, and receptive vocabulary, as well as diverse FB task designs [2007; Doherty, 2008]. They did, however, find receptive vocabulary less strongly associated with FB performance than the other measures [Milligan et al., 2007; Doherty, 2008], which interestingly reflects the (2016a) findings of Chahboun et al. that receptive vocabulary and metaphorical comprehension were not related. As Doherty argued, this may be because receptive

[9] This is interesting in the context of mentalising within therapy or intervention schemes, as perhaps the process of making mental states explicit allows for their mental realisation and processing – this, however, is beyond the scope of this thesis.

vocabulary measures are designed to probe a very specific area of language ability [2008, p.163]. Milligan et al. concluded from their analysis that there is a bidirectional relationship between general language and ToM development, but that there is a stronger effect of early language on later FB [2007; Doherty, 2008]. These findings bring us back to the theoretical accounts emphasising the role of discourse in the acquisition of ToM.

A number of authors take a socio-cultural approach in analysing the relationship between language and ToM, centring their focus on the input of conversational pragmatics. These accounts propose that children's ToM equates to 'social understanding' and is acquired through participation in 'the social world': more accurately framed as 'the mental world' or 'a community of minds' [Astington & Baird, 2005, p.11]. Language is viewed as the key to this world and the medium through which children join this community of minds [Astington & Baird, 2005]. It is then within this community that children learn how minds interact and encounter the fluidity and opacity of the mental states beneath people's emotions and intentions [Astington & Baird, 2005; Nilsson & De López, 2016]. There, early ToM skills assist them with their further acquisition of lexical items, semantic concepts, and syntactic structures, while continually expanding linguistic abilities improve the experiential and informative quality of their interactions with others [Astington & Baird, 2005; Sodian & Kristen, 2010]. Thus, early language capacity kickstarts a reciprocal relationship between language and ToM of staged co-dependent development. In this sense (and in keeping with Vygotskian principles) children's socially mediated experience of external verbal representations within discourse drives their development of internal verbal representations [Astington & Baird, 2005, p.12: citing Nelson, 2005].

Building upon this perspective, some of the research literature claims that children's acquisition of a ToM is brought about by their participation in conversation or discourse. Evidence from the research has shown that parental (particularly maternal) language about mental states underpins later ToM ability, with FB performance strongly predicted by the amount of conversation about feelings and causal relations that children have experienced, as well as the number of adults and older children that they have interacted with on a daily basis [Astington & Baird, 2005; Doherty, 2008; Dunn & Brophy, 2005; Sodian & Kristen, 2010]. A pronounced older sibling effect was also found in Ruffman et al.'s (1999) study, which may be attributable to various factors, for example: the stimulation of pretend play, fostering the development of FB understanding; or greater linguistic opportunities, exposing children to more complex mental-state language either directly from the older sibling or indirectly via observed parent-older sibling interactions [cited in Doherty, 2008, p.168]. This effect is interesting in the context of the wider literature on autism and SLI, where it has been found that the later-born siblings of children with ASD share some of their early non-verbal communicative deficits (e.g. in gesture) associated with subsequent language delay [Le Barton & Iverson, 2016; Peterson & Siegal, 2000].

Meins et al.'s (2002) work found that maternal language predicted ToM performance at 3-4 years-of-age, which they attributed to mothers' 'mind-mindedness': crediting infants with their own individual minds instead of responding to them only as things with needs to be satisfied [cited in Doherty, 2008, p.165]. This work measured the appropriateness of mental-state language used in mother-child dyadic interactions, analysing whether maternal comments were relevant to and an accurate reflection of their children's presumed mental states, and whether they sustained and prolonged the interaction [in Doherty, 2008]. The participants were later assessed with a battery of ToM (FB and appearance-reality) tasks, with VMA followed by maternal 'mind-mindedness' (at 6 months-of-age) emerging as the strongest predictors of performance [in Doherty, 2008]. Similar findings have been established across the research, demonstrating the positive effect on later ToM performance of increased maternal use of 'think' and 'know' constructions (sentential complements, 3.1.2) as well as modulated assertions expressing uncertainties or possibilities, such as 'I wonder what this/that is… It could be (x).', which highlight the gap between thought and knowledge [Doherty, 2008, p.166: citing Ruffman et al., 2002]. Dunn & Brophy further asserted that the characteristics of children and their communicative partners, and the quality of the relationship between them, influenced the frequency of their exposure to interactions where meaningful mental-state discourse occurred [2005; Astington & Baird, 2005].[10] Work by Dunn et al. (1991) as well as Ruffman et al. (1999) has also indicated that the early experience of social conflict, as well as parental resolution of it, promotes social understanding and later FB performance [cited by Doherty, 2008]. This transpires via language use about the causality of feelings, and encouragement to reflect on the emotional perspectives of others (particularly the victim) during disciplinary contexts [in Doherty, 2008]. A number of studies have also shown that primary caregivers actively moderate their language to scaffold children's developing understanding, which is important for their emerging ToM [Doherty, 2008, p.167; Astington & Baird, 2005; Harris, 2005]. These findings, which emphasise strong environmental effects, do not mesh well with the theories from modular or nativist accounts of ToM, instead providing more support for those supporting conceptual change, e.g. simulation or theory theory, as language appears to drive the conceptual change.

[10] This is interesting and significant when considered in the context of attachment and epistemic trust, with implications for the development of strategies for cognitive intervention and therapy.

Socio-economic and cultural backgrounds have also been investigated in relation to the role of discourse in ToM, with socio-economic influences seemingly more powerful than cultural [Doherty, 2008]. Studies have found that socio-economic status has a significant impact on performance across language, emotional understanding, and FB measures, with language and FB performance further correlating independently of background [Doherty, 2008: citing Cutting & Dunn, 1999, and Hughes et al., 2005]. Correlations of ToM performance were also identical across 1116 monozygotic and dizygotic pairs of twins in Hughes et al.'s (2005) study, indicating that shared environment contributes more than genetics to performance, with ToM, verbal ability, and socio-economic status all significantly intercorrelated; there was also an equal influence of genetics and shared environment on language and ToM abilities [cited by Doherty, 2008, p.169-170, and Astington & Baird, 2005, p.21-22]. Hughes (2005) also found that environmental factors accounted for a larger proportion of individual differences in ToM, and that only genetic factors which also influenced verbal ability made any contribution [cited in Astington & Baird, 2005, p.21-22]. Hughes concluded that the impacts of genetic and environmental factors could shift across development, with genetics potentially functioning as a limiter in earlier developmental stages, but environment becoming increasingly influential over time [cited in Astington & Baird, 2005, p.21-22]. Doherty argued that the covarying differences of language and ToM across participants, as well as comparable findings of socio-economic differences in pretend play, provided robust evidence that language (e.g. the frequency, content, and form of conversations, in addition to parental presence and parenting styles) mediates ToM development [2008, p.170]. Contrastingly, findings from work on cultural differences have not been analogous, tending, instead, to favour the universal timescale account of ToM development, although some East Asian research has apparently indicated measurable culturo-linguistic variations [in Doherty, 2008]. The former (socio-economic) findings are more in line with linguistic, simulation, and theory accounts although the latter (cultural) findings generally appear more supportive of modular, nativist, and EF accounts. That said, while culture may affect the specific linguistic inputs to ToM development which children receive, it is also true that culture is mediated by general language capability, and that without language, more complex cross-cultural and -societal differences in thinking or behaviour might not have arisen.

Of course, ToM skills also impact on children's ability to engage successfully in discourse: children have to consider the mental states of their conversational partner in order to appropriately introduce and refer to new information, e.g. knowledge that they may (or may not) have or want, or to recognise topics that will be relevant and engaging to them [O'Neill, 2005; Astington & Baird, 2005]. Adult proficiency with mentalising has also been shown to boost shared understanding (and thereby efficiency) within discourse [Sodian & Kristen, 2010, p.195: citing Krych-Appelbaum et al., 2007]. It still remains plausible that early language precedes and facilitates the development of these skills. It makes sense logically that a greater demand for ToM competence would spur development to meet this demand. To summarise, so far reviewed are some robust findings from the research that language is fundamentally necessary for the development and acquisition of ToM, but findings also exist, which seem to indicate that ToM development follows a cross-culturally universal timescale. Section 3.2 will cover evidence from the research more strongly in favour of a critical role of language.

3.2 The Case for a Critical Role of Language in ToM Development

So far, the evidence from the research indicates that language plays a complex role in children's acquisition of ToM. However, is language essential for or causal of ToM development? Is ToM crucial for linguistic development? Or is there some other underlying factor vital for both linguistic and ToM development? Research with atypical populations may provide the necessary evidence to answer these questions. The following subsections will examine the interaction between language and ToM development in two atypically developing populations: SLI and deafness with late acquired sign language.

3.2.1 Specific Language Impairment (SLI)

SLI is the delayed or disordered development and acquisition of language in populations with hearing and non-verbal ability in the normal range, and without any apparent neurological impairment or sensory dysfunction [Cummings, 2014; Farrant et al., 2006, p.1843-1844: citing Bishop, 1997; Nilsson & De López, 2016]. Research has shown that children with SLI present with social difficulties that extend beyond those of direct communicative or pragmatic issues to general interaction with peers even in non-verbal cooperative tasks [Happé & Frith, 2014, p.565]. It remains unclear whether linguistic impairment is the direct cause of difficulties with social interactions and establishing peer relationships, or whether underlying socio-cognitive deficits are responsible for both linguistic and mentalising deficits [Happé & Frith, 2014, p.565]. Nevertheless, it has been argued that the presence of systematic language delays or disorder in the face of intact non-verbal ability means that the SLI population could provide important insights into the relationship between language and ToM [Farrant et al., 2006, p.1844; Nilsson & De López, 2016].

Research by Holmes (2002) and Tucker (2004) found delays of approximately 12-18 months in the ToM development and FB understanding of children with SLI, aged between 4-7 years and 5-6.5 years [cited by Farrant et al., 2006]. Previous work in the 1980s and 1990s by Leslie, Frith, Perner, and colleagues had found that ToM was intact in children with SLI, but it has been observed that this earlier research was carried out with participants whose average chronological age was much higher than the age at which explicit FB understanding typically develops [Farrant, 2015; Farrant et al., 2006]. Thus, the earlier findings did not preclude a potential developmental delay of ToM in SLI [Farrant, 2015; Farrant et al., 2006]. Moreover, Farrant suggested that the data from such studies investigating older children with SLI indicated that delayed ToM development could extend up to and beyond 12 years-of-age depending on each child's experiences with education and intervention [2015, p.3]. Farrant et al. discussed mixed data from Miller's (2001, 2004) work, which found some significant between-group differences in the 2001 study and no significant differences in the 2004 study [2006]. They argued that limited statistical power (participants in each study numbered only 10 and 15 respectively) and a focus on only FB tasks significantly weakened Miller's findings [Farrant et al., 2006]. Farrant et al.'s (2006) study sought to extend the ToM research regarding children with SLI to their VPT1 and VPT2 skills, with reference to Harris' simulation theory (2.2.2), which they argued could account for a causal role of language, as well as the relationships between language, VPT (2.3.3), and ToM development. Farrant (2015) framed ToM as cognitive PT, emphasising the additional importance of and need for further investigation of visual PT and emotional or affective PT (2.3.3). Farrant also cited evidence from the research demonstrating a relationship between language and the development of emotional PT skills in both typically developing and profoundly deaf children (3.2.2) [2015, p.3]. Farrant's more holistic view of ToM and PT is worthy of note and potentially a more representative reflection of ToM as it is often treated within the wider literature. Like Holmes (2002) and Tucker (2004), Farrant et al.'s (2006) study also found delayed ToM development in children with SLI across a wide range of ToM and VPT tasks [2006]. The study assessed 20 children with SLI (13 males; mean age 62.9 months) and 20 typically developing children (12 males; mean age 61.2 months), matched for non-verbal IQ [Farrant et al., 2006]. They found that VPT2 and ToM (particularly in terms of the unexpected contents FB) task performance correlated significantly, as well as a particular delay in the development of VPT2 and ToM in children with SLI [Farrant et al., 2006, p.1847]. For these reasons, they proposed that the ability to consider another's perspective is rooted in language development, with the perceptual mental state of 'seeing' preceding and enabling the epistemic mental state of 'thinking' [Farrant et al., 2006, p.1849]. Notably, this proposed process would mirror children's acquisition of mental-state terms (discussed in 3.1.1), which begins with perceptual verbs ('see', 'look'), progressing to terms of emotion and desire, and finally reaching cognitive verbs ('know', 'think', 'remember'). It also arguably parallels the proposals (discussed in 3.1.2) that complementation employing the verb 'say' (which conveys propositional contents that are more overtly perceptible or objectively verifiable) bootstraps the complementation of mental or cognitive verbs such as 'think' (of which the propositional contents are less overtly perceptible and objectively verifiable).

Farrant's (2015) work further extended these findings, providing evidence that children with SLI performed significantly less successfully than typically developing children across a range of explicit PT skills, including diverse ToM, VPT, and emotional PT tasks [2015]. Farrant recruited 30 (26 boys) children with SLI aged from 48 to 74 months and 30 typically developing children matched for non-verbal ability, gender, and age [2015]. Participants also ranged in socio-economic status, and, importantly, a diagnosis of autism was an exclusion criteria [Farrant, 2015]; this is important because a diagnosis of autism bears other implications, as will be discussed in §4. The matched group of typically developing children performed significantly better than the children with SLI across tasks such as: diverse desires, knowledge access, unexpected contents FB, low verbal FB, emotional PT, and VPT1, VPT2, and a new 'VPT3' task; there was also a marginally significant difference in performance on a diverse beliefs task [Farrant, 2015, p.5]. In order to achieve greater clarification on the relationship between language and ToM development in SLI, Nilsson & De López conducted a (2016) meta-analysis across 17 studies involving 745 children with SLI between 4-12 years-of-age [2016]. They found that children with SLI had significantly lower performance on ToM tasks relative to age-matched typically developing children, and that the statistical effect was not moderated by the factors of age and gender [Nilsson & De López, 2016]. These findings are important, as the subject of gender has been raised as problematic, given the emerging gender differences in the research in terms of brain activation patterns [Frank et al., 2015] and the higher prevalence of a diagnosis of SLI and autism among boys than among girls [Doherty, 2008; Lai et al., 2012; Nilsson & De López, 2016; Parish-Morris et al., 2017]. The results of Nilsson & De López's meta-analysis indicate the presence of an interface between language and ToM development, and provide further support for a language account of ToM [2016]. Finally, the implications of the recent findings considered in this section are also very pertinent to the literature on delayed ToM and associated linguistic deficits in autism; this will be examined further in §4. Before turning to the domain of autism, however, the relationship between language and ToM should also be considered within the context of the late-signing deaf population.

3.2.2 Deafness and the Late Acquisition of Sign Language

Evidence from the research has also demonstrated pragmatic deficits and a delay in the ToM development of children who are profoundly deaf, and whose access to sign language is delayed, such as deaf children with non-signing parents [Astington & Baird, 2005; de Villiers, 2005; Doherty, 2008; Farrant et al., 2006; Farrant, 2015; Peterson & Siegal, 2000]. De Villiers argued that deaf children presented a valuable population within which to examine the causal role of language in ToM development, because even with significantly delayed language acquisition, many demonstrate EF skills, non-verbal intelligence, and sociability that are comparable to typically developing controls [2005; Astington & Baird, 2005; Doherty, 2008; Farrant et al., 2006].

Across two studies, de Villiers found that language-delayed deaf children (of non-signing hearing parents) were significantly worse at mental-state reasoning than their early-signing counterparts (children of deaf signing parents), who had ToM and language skills on a par with typically developing hearing controls [2005; Astington & Baird, 2005; Nilsson & De López, 2016; Peterson & Siegal, 2000]. These findings were applicable to both a delay in sign-language acquisition, and oral-language acquisition (e.g. after a later fitting of cochlear implants) [de Villiers, 2005; Astington & Baird, 2005; Peterson & Siegal, 2000]. De Villiers also found that deaf children's FB reasoning was predicted independently by both general verbal ability in terms of vocabulary and specific aspects of syntax (e.g. false complements), indicating that both semantics and syntax contribute to ToM development [2005; Astington & Baird, 2005] (3.1.1, 3.1.2). A number of similar findings can be seen across the literature. One (1999) study by Peterson & Siegal examined children from 5-13 (mean 9) years-of-age, with the children of signing parents performing comparably to typically developing 4-year-olds but the children of hearing parents performing poorly and slightly worse than a group of children with autism, whose mean VMA was 8-years (§4) [2000; Doherty, 2008, p.164; Farrant et al., 2006, p.1843]. In their (1998) study, Russell et al. found that only the 15-year-old age group (as opposed to the 6-year-old and 10-year-old groups) of their deaf participants passed the unexpected transfer FB task [cited in Doherty, 2008]. These findings also bear up within low-verbal or non-verbal measures, and the participants in the studies by Russell, de Villiers, Siegal, and colleagues, had no known learning difficulties or other impairments [Doherty, 2008]. The most probable explanation for the delay in ToM development is therefore an early lack of cognitive and language input [Doherty, 2008; Peterson & Siegal, 2000].

The absence of exposure to mental-state terms and syntax results in children lacking the means through which to have conversations about mental states, which they cannot otherwise observe, and limits their opportunities for social interaction and participation in activities such as pretend play [Doherty, 2008, p.164; Peterson & Siegal, 2000]. Following Perner's (1999) logic, it makes sense that an absence of language could hinder ToM development, because it is language (rather than just behavioural observation) that provides the most accurate source of information about the contents of other people's minds. Taken together, the findings discussed above suggest that the acquisition of language is necessary in order to facilitate or trigger ToM development.

§4 Domain of Interest: Autism

4.1.1 Why Autism?

§3 examined a wide range of evidence from across the literature as well as from work with typically developing children, children with SLI, and late-signing deaf children, from which it has consistently emerged that language plays an important role in facilitating or spurring the development of a ToM. Another population important for the investigation of the relationship between language and ToM is the autistic population. Deficits in both language and ToM present across the autistic spectrum, and appear to be related. With reference to the conclusions drawn in §3, the focus of this next section will be to consider these deficits within the context of a language account of ToM, as well as to try to establish the nature of their relationship.

4.1.2 What is Autism?

Although the term 'autism' was first coined by Bleuler in 1908 in his work with severely withdrawn schizophrenic patients, its modern meaning derives from concurrent work in the 1940s by Kanner and Asperger [Baron-Cohen et al., 1985; Doherty, 2008; Frith & Happé, 1994; Happé & Frith, 2006; Rogers et al., 2006]. Autism is a neurodevelopmental disorder which presents across a spectrum, and is therefore often referred to in the literature as Autism Spectrum Disorder (ASD) or Autism Spectrum Condition (ASC) [Baron-Cohen, 2017; Chahboun et al., 2016a,b]. A second subtype is recognised diagnostically in the ICD-10 (but not DSM-V) as Asperger's syndrome [Baron-Cohen, 2017], which overlaps with the category of high-functioning autism (HFA). Most children with the 'core' or 'classic' form of autism present with delayed and/or impaired language development, as well as moderate to severe learning difficulties, whereas children with Asperger's syndrome (or HFA) appear to reach language milestones at around the same time as typically developing children and have an IQ within the normal range (higher than 70) or sometimes even higher [Doherty, 2008; Norbury, 2005]. A small percentage of individuals with autism have been termed 'savants' for possessing an extreme and specific talent, for example in mathematics, art, or music, etc. [Doherty, 2008]. Historically, diagnosis of the two types was subject to their differences in language acquisition. However, more recent research has demonstrated that problems with language exist across the autistic spectrum, although the extent of impairment can vary from a total lack of language acquisition at the lower end to almost completely intact structural language at the higher end [Chahboun et al., 2016a, p.27; Norbury, 2005]. Difficulties with pragmatic language manifest across the entire spectrum, for example, impaired understanding of jokes and figurative expressions, as well as inferencing [Chahboun et al., 2016a,b; Norbury, 2005; Peterson & Siegal, 2000; Vulchanova et al., 2015]. Strikingly, the gender ratio across diagnosed individuals is heavily skewed towards male, particularly at the higher IQ end of the spectrum [Doherty, 2008]. This could be due to genetic heritability, biochemical factors, such as exposure to high foetal testosterone levels (which differentially affect males and females), or socio-biological factors masking detection, such as a tendency towards stronger language ability, societal compliance, and empathy in females [Anwar et al., 2018; Doherty, 2008; Lai et al., 2012; Parish-Morris et al., 2017]. This is also pertinent to Frank et al.'s findings that men and women comprehend and process pragmatic language and ToM differently [2015]. Wing & Gould's (1979) triad of impairments for autistic diagnosis specified impairment in social interaction, social communication, and imagination, although the imagination specification has now been replaced by restricted repetitive interests, behaviours, or activities (RRBIs) [Doherty, 2008; Happé, 1995b; Peterson & Siegal, 2000]. Impairments in social interaction and communication are more easily explained by a ToM account than RRBIs, although these may be a coping mechanism for anxiety, which may arise from difficulties with comprehension and/or a lack of awareness of the social inappropriateness of such behaviours [Doherty, 2008; Hollocks et al., 2016; Jones et al., 2017; Joyce et al., 2017]. Typically children with autism do not spontaneously engage in pretend play, seeming to prefer functional play, but they are capable of doing so if prompted [Doherty, 2008; Pickles et al., 2016]. Autistic individuals are often better at isolating and attending to details as well as collating and systemising information than typically developing individuals [Doherty, 2008; Frith & Happé, 1994; Happé & Frith, 2006]. A diagnosis of autism also overlaps and can co-occur with a diagnosis of other disorders such as ADHD, anxiety, dyslexia, dyspraxia, dyscalculia, and psychopathy, which should be borne in mind when investigating ToM (see §5) [Baron-Cohen, 2017; Doherty, 2008; Rogers et al., 2006; Tick et al., 2016].

4.2 Autism and ToM

Baron-Cohen et al. proposed the first ToM account of autism in their (1985) publication, hypothesising that autistic socio-cognitive deficits resulted from an absent or impaired ToM. To argue their case, they compared the performance of autistic children to two control groups (typically developing children and children with Down's syndrome), finding that only the children with autism were significantly impaired in their ability to attribute (false) beliefs to others and predict their behaviour [Baron-Cohen et al., 1985]. Baron-Cohen et al.'s (1985) ToM account has since met with a number of challenges (including those posed by EF accounts, discussed in 4.4) but it does remain clear that there are some deficits and delays in autistic ToM development. Aside from the unexpected transfer FB task (2.3.1), autistic children have also demonstrated impaired performance on the unexpected contents task (2.3.1) for ‘self’ and ‘other’, as well as the appearance-reality task (2.3.1) [cited in Doherty, 2008: Perner et al., 1989; Baron-Cohen, 1989]. Interestingly, research appears to have shown that autistic participants make egocentrically mediated errors in ToM tasks similar to those of pre-4-year-old, typically developing children, with the literature divided upon whether the deficits exist in self-, other-, or both self- and other-oriented mental-state attribution [Bradford et al., 2015]. Importantly, in order to attribute mental states to oneself or others, one must first have conceptualisations and representations of ‘self’ and ‘other’ [Barsalou, 2016]; it may be that the process of forming these situated conceptualisations or representations is somehow disrupted in autism. Autistic participants typically reach a chronological and verbal mental age of at least 9-10 years-old before half are able to pass ToM tasks like the unexpected transfer FB task, a much higher age than the 4 years required by typically developing children [Frith & Frith, 2005; Happé, 1995a; Perner, 1999]. In Happé’s (1995) study, autistic participants did not begin to chance passing ToM tasks until a VMA of 5-6 years, whereas typically developing children had a chance of passing at a VMA of 2-3 years [1995a]. Happé found that her data most closely fit within a two-threshold model, where all participants failed ToM tasks below a lower threshold and passed them above an upper threshold [1995a]. Happé also discussed previous work which had found verbal ability necessary but not sufficient for successful performance on ToM tasks, speculating that an absent ToM would hamper word learning processes by disrupting processes of joint attention, reference, and ostension [1995a, p.851; Bates & Dick, 2002]. As previously mentioned, Happé further proposed that autistic participants might ‘hack out’ a solution to ToM tasks using compensatory cognitive strategies mediated by language; these language skills would take time to develop, which might account for their delayed success on ToM tasks [1995a]. Alternatively, returning to the concept of a staged interactional development of language and ToM, it may be the case that early delays in language acquisition have a knock-on effect on early ToM development, which then push back subsequent co-developing linguistic and ToM processes. Whichever it may be, language ability is clearly involved in the development of ToM in autism; the examination of this relationship will be continued in 4.3.

Relevant to the concept of staged co-development, and the importance of VMA in ToM performance, Happé’s (1994) strange stories task (previously discussed in 2.3.1) revealed a ranged performance among autistic participants [1994]. This range paralleled their performance across first- and second-order FB tasks, where the more able participants who passed second-order FB tasks were nearly able to reach the performance range of the typically developing controls [Doherty, 2008; Happé, 1994; Perner, 1999]. This work indicated that the level of ToM ability in some HFA participants, while still impaired, might eventually be similar to that of some typically developing controls [Doherty, 2008; Happé, 1994]. Some of Happé's strange stories such as the double-bluff condition, might even require a third-order attribution on the behalf of the participant and, interestingly, these types were consistently among the most difficult for Happé's autistic participants, including the more able second-order FB passers [1994]. This is pertinent to Farrant's recent work on a third level of VPT (VPT3) [2015]. Like Farrant and colleagues, Baron-Cohen et al. posed ToM, FB, and higher-order mental-state attribution as forms of conceptual or cognitive PT but the latter differentiated it from perceptual (visual) PT, arguing that the two abilities were not supported by the same processes, because autistic children had been shown to perform successfully on VPT tasks [1985, p.43-44]. More recent research, however, has challenged Baron-Cohen et al.'s (1985) view, suggesting instead that ToM and VPT share common cognitive processes, and function within a wider PT framework [Farrant, 2015; Farrant et al., 2006; Hamilton et al., 2009; Pearson et al., 2013]. Some studies have also demonstrated more recently that VPT is not intact in autism [Hamilton et al., 2009; Pearson et al., 2013]. This is highly relevant to Farrant and colleagues' (2006, 2015) work with children with SLI, because if: (a) VPT and ToM are related processes, and (b) language is to be established as critical in facilitating ToM and higher-order VPT development; then language, ToM and higher-order VPT must all be shown to be impaired not only within the SLI and late-signing populations (3.2.1 and 3.2.2) but also within the autistic population. Linguistic impairment and its relationship to ToM development in autism will be discussed further in 4.3, but first it is important to examine the relationship between VPT and ToM in the context of autism.

As discussed in 2.3.3, a distinction exists between VPT1 (perceptual PT) and VPT2 (conceptual PT), with success on VPT1 tasks emerging at around 2 years-of-age (similar to implicit ToM understanding, see 2.3.1) and on VPT2 tasks at around the same time children start to pass FB or 'first-order' ToM tasks (4-5 years-of-age).[11] Pearson et al. reviewed a number of VPT studies carried out with autistic participants in an attempt to settle the controversy surrounding the status

[11] It would be interesting in future research to make connections between 'second-order' ToM development and that of the proposed VPT3. It may be the case that implicit early ToM parallels VPT1 and might be a better candidate for the label of first-order ToM (or ToM1), shifting 'first-order' FB tasks to second-order ToM (ToM2) paralleling VPT2, and finally 'second-order' ice-cream van tasks to third-order ToM (ToM3) mirroring 'VPT3'. However, these ideas are beyond the present scope of this thesis and require further investigation.

of VPT in autism [2013]. They argued that a variety of methodologies had been used in the past, conflating and lacking the necessary distinction between VPT1 and VPT2 tasks, in addition to comparing older groups of autistic participants to younger groups of neuro-typical controls, which could mask developmental delays [Pearson et al., 2013]. As 'first-order' ToM (e.g. with FB tasks) is impaired/delayed in autism, then it should follow that only performance on VPT2 tasks is similarly impaired/delayed. Pearson et al. also noted that a number of allegedly 'VPT2' tasks in the literature held design flaws, allowing them to be carried out successfully using simple VPT1 (line of sight) strategies, and vice versa with VPT1 tasks [2013]. Under their reanalysis of previously conducted VPT studies, they concluded that the studies claiming no VPT impairment in autistic participants relative to typically developing controls may have only assessed VPT1 abilities, with which no impairment would be expected, while the majority of studies which successfully tapped VPT2 ability revealed autistic impairment [Pearson et al., 2013]. Hamilton et al.'s findings similarly demonstrated impaired performance on VPT2 tasks in autism [2009].

Interestingly, like the (2006, 2015) work by Farrant and colleagues, Pearson et al. also noted that a number of studies had found significant effects of verbal ability within the groups, implicating verbal ability as an important predictor of VPT [2013, p.4]. Conversely, Hamilton et al. found that mentalising ability but not VPT2 ability increased with verbal ability [2009, p.40-41], although this might be accounted for by their limited choice of language measure, as they only employed the British Picture Vocabulary Scale (BPVS). On the whole, the evidence reviewed so far (and in 3.2.1) suggests that VPT and ToM are both part of a broader PT system, the entirety of which, as well as its individual components, have been associated with linguistic development. Reducing autistic socio-cognitive deficits to a purely ToM account may be too narrow a perspective, as autistic individuals also exhibit deficits in language. These early linguistic deficits may be the cause of ToM impairment, and therefore the autistic population may present an ideal group within which to test a language account of ToM.

4.3 Autism, ToM, and Language

If a language account of ToM is to be established, then it must be evidenced and found to be viable within ToM-impaired populations such as autism. As in the SLI (3.2.1) and late-signing deaf (3.2.2) populations, ToM impairments and delays have also been associated with linguistic impairment in the autistic population. Around half of the people with classic autism produce no language, while those who can use language display characteristic abnormalities in their communication [Doherty, 2008]. Some of these atypicalities, such as speaking too quickly, loudly, quietly, repetitively, or in a disjointed manner, are consistent with the concept of an impaired ToM because they indicate that the autistic speaker is unable to take into account the conversational needs of their interlocutor [Doherty, 2008; Peterson & Siegal, 2000]. Another phenomenon, the reversal of pronouns and social-role verbs (e.g. 'lend' and 'borrow', 'learn' and 'teach'), implies further difficulties with PT and understanding social interactions [Doherty, 2008; Peterson & Siegal, 2000]. Other problems have been observed with a lack in or the inappropriate use of introductory tag-phrases (e.g. 'by the way'), and the distinction between new and old information, suggesting an impaired understanding of the purpose of using these phrases and of their listener's informational needs [Doherty, 2008; Peterson & Siegal, 2000]. Research has also shown that autistic children's command of mental-state vocabulary is below that of their typically developing peers, and that they produce fewer references to mental states (especially belief- or knowledge terms) in everyday speech [Tager-Flusberg, 1992, cited in Doherty, 2008, p.188-189; Peterson & Siegal, 2000]. Additional atypicalities such as a reduced motivation to engage in early verbal or gestural communication, and the exhibiting of echolalia (immediate or later word-for-word repetition of the speech of others) indicate that language might be detached from its communicative purpose in autism [Doherty, 2008, p.191; Peterson & Siegal, 2000].

Autistic individuals typically also struggle in a receptive role in conversation, experiencing many difficulties with pragmatic and non-literal language comprehension, and drawing inferences about the speaker's intended meaning [Chahboun et al., 2016a,b; Happé, 1993, 1994, 1995b; Norbury, 2005; Vulchanova et al., 2015] (previously discussed in 3.1.3). It is also important to bear in mind that many studies which have been conducted with the autistic population have often been heavily biased towards male gendered participants because the majority of autistic individuals diagnosed are male, and females may present differently, leading to under-diagnosis [Doherty, 2008; Lai et al., 2012; Parish-Morris et al., 2017]. Very recent research by Parish-Morris et al. examined linguistic 'camouflaging' in the female autistic population, where the social behaviour of girls with autism appears superficially typical [2017]. They compared use of pragmatic language markers such as conversation pause fillers (e.g. 'um', 'uh') across genders in both typically developing and autistic school-age populations, finding parallel sex differences across both populations with greater female production of 'um' vs. greater male production of 'uh', as well as that filled pause rates were higher overall for typically developing children than for children with autism [Parish-Morris et al., 2017]. They also found that higher 'um' ratios correlated with better socialisation in boys with autism, the effect driven by increased use of 'uh' by boys with stronger symptom severity [Parish-Morris et al., 2017, p.1]. Looking at their findings, Parish-Morris et al.

proposed that disfluency patterns which appear typical might normalise how girls with autism sound when compared with the general population, in contrast to autistic boys, providing autistic girls with a form of linguistic camouflage [Parish-Morris, 2017, p.1]. These findings are striking and warrant further investigation, urging caution with the interpretation and evaluation of data collected from different genders in studies on pragmatic language in autism.

As discussed across previous sections, and particularly in 4.2, Happé found that children's VMA and vocabulary levels correlated with their ToM ability [1995a]. She proposed that high-functioning autistic individuals might 'hack out' solutions to ToM tasks through the use of compensatory, linguistically mediated strategies, especially given that autistic second-order ToM 'passers' were more likely than typically developing controls to be able to give some form of justification for their answers [Happé, 1995a, p.852]. Happé also argued that verbal ability was necessary but not sufficient for successful ToM performance, suggesting that the absence of a ToM would hinder the process of vocabulary acquisition by derailing naturalistic learning opportunities which would occur within typical development, such as the establishment of joint attention, promotion of reference and ostension, etc. [1995a, p.851; 1995b; Bates & Dick, 2002]. In Happé's study, children with autism required a much higher VMA to pass FB tasks than other populations [1995a].[12] Happé also made a striking comment about the influence of higher motivation or better teaching on language and ToM task performance [1995a, p.852], which is interesting when considered in the context of Pickles et al.'s (2016) work on parent-mediated social communication therapy for young children (2-4 years-of-age) with autism (PACT). Pickles et al. conducted a follow-up study with participants from the PACT early intervention programme, finding that randomised controlled trials showed long-term (nearly 6 years post-treatment) reduction in symptom severity across various measures, such as the Autism Diagnostic Observation Schedule (ADOS), the Dyadic Communication Assessment Measure (DCMA) of child-initiated interactions with the parent, and a battery of expressive-receptive language tasks [2016]. These findings suggest that language-mediated intervention can stimulate ToM

[12] If this finding is considered with regards to an earlier conjecture made in 4.2: that there is a staged interactional and bidirectional relationship between the development of language and ToM {language → early ToM(1) → language ↔ ToM(2)(3)}, with early impairments or delays causing knock-on effects for subsequent co-developmental stages and processes; then this might imply that early 'corruptions' or variations in language and social cognition are only 'corrected' or compensated for (up to a point) in autism after a long period of language-mediated statistical learning, hence the delays. This is of course pure speculation and would require investigation within a research context.

development in autism, potentially reducing the length of socio-cognitive developmental delays, which provides further support for a language account of ToM.[13]

Relevant to the concept of a language-mediated training effect, Durrleman & Franck conducted a study on the relationship between complement clause sentences and ToM in children with autism relative to younger typically developing controls [2015]. They found significant correlations between a verbal FB task and complement tasks involving verbs of cognition and communication, arguing that their findings provided support for the perspective that knowledge of particular types of complementation served as the means through which autistic individuals could linguistically ‘hack out’ a solution to verbal FB tasks [Durrleman & Franck, 2015]. These types of complements (3.1.2) are those with a truth-value independent to that of the matrix clause, which occur with verbs of cognition and communication, but not of perception, and can embed a false proposition, allowing for the linguistic (and possibly cognitive) representation of FB [Durrleman & Franck, 2015, p.15-16; Tager-Flusberg & Joseph, 2005; Tager-Flusberg, 2007]. While Durrleman & Franck found that their autistic group performed similarly to their typically developing group on complementation syntax when measured independently of other cognitive functions, the typically developing group were much younger, supporting the notion of delayed development in autistic comprehension of complement sentences, which is progressively caught up with advancing cognitive maturation [2015, p.25].

It may be the case that an interaction between processes of language and social cognition experiences an underlying (or possibly even a surface) disruption at a key stage, which is then responsible for delays in ToM development in autism and other similar disorders. Interestingly, Griffin & Baron-Cohen asserted that semantic-pragmatic disorder appeared to share right hemispheric dysfunction with autism and thus might be a part of the spectrum [2002, p.98]. As Doherty discussed, a successful explanation depends upon whether or not autism can be accounted for by a single psychological or physical cause, as well as whether it is a unitary disorder with a spectrum of symptom severity, or comprises genuine subtypes, each with different underlying physical causes and manifesting different psychological effects [2008, p.194-195; Tager-Flusberg, 2007]. Before drawing any final conclusions about the role of language in ToM accounts of autism, it is important to consider EF and cognitive style accounts.

[13] It would be meaningful and informative to follow up this work with research into the FB performance, VMA, and later pragmatic, non-literal, and figurative language competence of these children who have undergone early intervention.

4.4 The Alternative Account: Executive Function

Some studies have indicated that autistic children approach ToM tasks as if they were logical-reasoning problems, using language and other non-social cognitive processes as their primary tools for solving the task instead of social insight [Tager-Flusberg, 2007, p.312; Brent et al., 2004]. Typically developing children, on the other hand, appear to use social intuition and insight about or conceptual knowledge of mental states, as well as general EF skills supporting verbal processing, memory for key events in a narrative, and inhibition of spontaneous responses [Tager-Flusberg, 2007, p.312]. Potential support for this view has also come from the observation that autistic participants who passed second-order ToM tasks were more able to justify their answers than typically developing controls [Happé, 1995a]. However, research with autistic children has also demonstrated existing EF deficits in planning, flexibility, and working memory combined with inhibitory control [Doherty, 2008; Ozonoff et al., 2004; Tager-Flusberg, 2007]. Difficulties with inhibitory control of prepotent responses have also been proposed to explain the third element of the previously mentioned triad of impairments: RRBIs, because developing new behaviours and interests might require inhibiting one's attention in one area in order to switch to another [Doherty, 2008; Frith & Happé, 1994]. These theories and findings are significant, because a number of EF accounts have emerged as alternatives to the ToM explanation of autism. These range from perspectives opposing ToM accounts, to those proposing some form of interaction between EF and ToM, to those presenting diverging cognitive styles. As discussed in 2.2.4, the development of ToM and EF appear to be intricately entwined, and EF may be necessary for ToM to emerge. It also remains possible that they share a staged, bidirectional, and co-developmental relationship, similar to the one proposed here between language and ToM. It makes sense logically that EF would support developments in language, with these linguistic developments potentially enhancing EF, and vice versa in a similarly associative interaction to language and ToM (§5). Thus, it may be more appropriate to conceive of a tripartite, multi-directional relationship between EF, language, and ToM (see §5), where each capability develops and is enhanced through shared benefits stemming from advancements across its counterparts.

Research has found significant correlations between the success of autistic children on ToM and EF tasks independent of age, and both verbal and non-verbal ability [Durrleman & Franck, 2015: citing Pellicano, 2007]. Common brain regions such as the prefrontal cortex and right hemisphere have been implicated in both EF and ToM processes (2.2.1-2.2.4) [Durrleman & Franck, 2015: for more citations; Mahy et al., 2014]. Similar to those in autism, EF deficits have also been found in patients with frontal lobe damage [Doherty, 2008: for more citations; Tsuchida

& Fellows, 2013]. For these reasons, some authors have proposed an EF account of autism and thereby ToM. Durrleman & Franck incorporated EF conditions into their (2015) study in order to examine this purported relationship, but found no deficits in the EF abilities (switching and inhibition) of their autistic group when compared with controls matched on non-verbal IQ [2015]. They did, however, highlight that the autistic group were on average two years older than the controls, which suggested a delay in the maturation of autistic EF [Durrleman & Franck, 2015], and could be consistent with language (or tripartite) accounts of ToM. Durrleman & Franck found no correlations between verbal FB and EF, but did find a significant correlation between one EF index (flexible rule-switching) and non-verbal FB, which they proposed was an indication that rigidity caused autistic participants to perform well on the non-verbal FB task [2015]. This might reflect that autistic participants were biased during the screening phrase instead of genuinely representing FB [Durrleman & Franck, 2015]. Jones et al. also conducted a (2017) study on the relationships between ToM, EF, and autism, finding that while EF was a correlate of ToM, EF had no direct association with autistic symptomatic expression, and that ToM ability was associated with both symptoms in social communication and RRBIs [2017]. The concept of an indirect association between EF and autism is also presented by White, whose 'Triple I Hypothesis' theorised that the pattern of autistic impairment on EF tasks was driven by difficulties relevant to mentalising [2013; Jones et al., 2017]. The above findings indicate that language is a more influential facilitator of ToM than EF, but do not resolve the question concerning the role of EF in the emergence or development of language and ToM. Theoretically, it may be possible that a developmental disruption or atypicality occurs at an intermediary point between an interaction of language and ToM development rather than during a stage of development in EF, which then impacts on later EF development, but this remains an area for investigation.

4.4.1 Executive Dysfunction: Weak Central Coherence (WCC) Account

Weak central coherence (WCC), proposed by Frith & Happé [1994; Happé & Frith, 2006], which shares similarities with Baron-Cohen's hyper-systemising theory [Barnes & Baron-Cohen, 2012], is related to EF accounts of autism. The WCC account frames autism in terms of a divergent cognitive style, where, relative to the general population, there are difficulties in perceiving something as a whole but strengths in focusing on its parts [Doherty, 2008; Frith & Happé, 1994; Happé & Frith, 2006]. Frith & Happé asserted that information processing in typically developing individuals was characterised by a tendency to pull together various, wide-ranging information to construct higher level meaning within context, or a tendency towards 'central coherence' [1994, p.121]. For example, with strong central coherence, the gist of a story is more easily recalled than its individual details and wording [Frith & Happé, 1994]; this pattern appears to be reversed in autism [Barnes & Baron-Cohen, 2012]. Strong central coherence might also explain how typically developing people easily infer the contextually appropriate meanings of ambiguous, non-literal language [Frith & Happé, 1994]. In these ways, WCC could explain the unusual profile and opposite patterns of performance on IQ and recall tests by autistic people relative to typically developing individuals, where autistic people demonstrate superior performance with pattern segmentation on the Wechsler scale Block Design subtest and are not assisted (and are even disadvantaged) by set relatedness in item recall tasks [Doherty, 2008; Frith & Happé, 1994]. Happé also found that autistic children were impaired in their reading of homographs relative to typically developing (7-year-old) controls, generally choosing the most frequent pronunciation without taking the rest of the sentence into account [1997: cited in Doherty, 2008; Frith & Happé, 1994]. Eberhardt & Nadig found a similar impairment in autistic interpretation of homonyms [2016]. A (2006) study by Snyder et al. claimed to have temporarily induced detail-focused 'savant-like' ability in typically developing participants through the use of repetitive transcranial magnetic stimulation (rTMS) to temporarily inhibit ability in the left anterior temporal lobe associated with meaning and conceptual processing [cited in Doherty, 2008, p.201]. These findings could provide support for a WCC account regarding savant abilities and hypersensitivity (and potentially also sensory integration dysfunction) in autism, although they do not necessarily explain other autistic socio-cognitive deficits. Savant abilities do not occur in every individual with autism and as discussed in 4.3, it may be the case that autism comprises several subtypes, with each having different underlying causes and symptomatic expressions.

Early theories of WCC proposed that it could account for ToM deficits in autism, given that effective social cognition involves the large-scale integration of information [Doherty, 2008, p.201: citing Frith, 1989]. Frith & Happé also argued that neither the ToM accounts nor EF accounts of autism fully explained areas of non-social deficits in autism, which they believed WCC could account for [1994]. However, WCC theories have since been revised to present WCC and ToM deficits as separate but interacting [Doherty, 2008, p.201; Frith & Happé, 1994; Happé & Frith, 2006], which, combined with issues of comorbidity, might fit within the aforementioned concept of a staged, multi-directional relationship between EF, ToM, and language (§5). Happé & Frith have also more recently framed WCC as a local processing bias or 'detail-focused cognitive style', independent from both executive functioning and ToM impairment [2006]. It does remain possible that WCC could work as some form of 'schematic' for directing EF. The revisions of WCC theory are a consequence of the general lack of evidence across the research in favour of an association between WCC and ToM, for example in terms of task performance on ToM and Block Design tasks [Happé, 1994: cited in Doherty, 2008]. Evidence from the research has also shown that people with autism do integrate information to a certain extent, and are subject to certain priming effects [Happé & Frith, 2006]. Happé also found that both autistic passers of first- and second-order FB tasks performed similarly on reading homographs in context and were better than those who did not pass, but were not as successful as controls [1997: cited in Doherty, 2008]. This might further implicate verbal ability as a stronger correlate of ToM. Similarly, Eberhardt & Nadig's (2016) findings contradicted WCC predictions, finding only structural language competence to be a significant predictor of performance on homonym tasks across autistic and SLI groups [2016]. These findings also reflect those of Chahboun et al. (2016a), considered in 3.1.3. From another angle, Happé & Frith discussed Just et al.'s 'underconnectivity' theory, which hypothesises that autistic symptomatology results from atypical brain connectivity, and has been demonstrated in some of the research literature, although this remains an area for investigation [2006; Frank et al., 2015].

4.4.2 Executive Function, ToM, and Language

ToM does not emerge independently: it co-occurs with and develops alongside EF and an understanding of non-mental representation (e.g. language), as well as meta-linguistic awareness (mLA) [Doherty, 2008, p.207]. For this reason, Doherty argued that EF and ToM developments may be different facets of the same phenomenon, with mLA, ToM, and EF all related to one another in complex, interdependent ways [2008, p.207]. This thesis expands upon that idea, proposing a tripartite (EF, ToM, and language) model (§5), which incorporates staged, multi-directional interactions between each component and its counterparts across development. Taking this perspective, EF and WCC accounts of autism do not adequately function as counterarguments to the main argument of this thesis, which is in favour of a language account of ToM, because all three could be accommodated within this model. Work on autism and other atypical populations has indicated that ToM is most likely an independent ability to EF, more closely associated with language capacity, although all three appear to interact, and the relationship between early EF, ToM, and language requires further investigation.

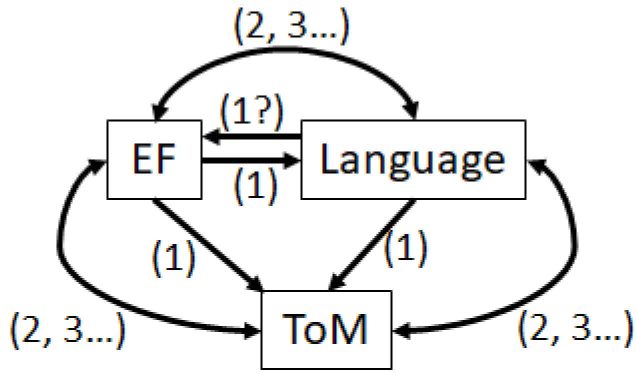
§5 Critical Discussion

The aim of this thesis was to establish a critical role of language in the development and acquisition of a ToM. To this end, a number of **common themes** have emerged throughout the consideration of the theoretical and research literature: **(1)** ToM development appears to be supported by prior developments in EF but also, importantly, in language; **(2)** ToM is part of a broader system of related processes comprising visual, spatial, affective, and cognitive PT; **(3)** ToM development seems to engender further development of EF and language, which then reciprocally enhance ToM development, and so on so forth; **(4)** There are delays in the ToM development of language-impaired populations, such as children with SLI, and late-signing deaf children, who demonstrate typical levels of EF; **(5)** In autism, ToM and linguistic deficits co-occur and appear to be related; and **(6)** EF deficits also present in autism, and they may or may not be related to the deficits in ToM and language development.

Another important question, which is relevant to theme **(1)**, has emerged within the literature and concerns the debate around the roles of EF and language in the 'emergence' and 'expression' of ToM. As discussed in 2.2.4 and throughout 2.3, a certain level of EF appears necessary (but not sufficient) for early language and early ToM to emerge, although it remains unclear whether EF is also necessary for their expression. Another aspect worth considering is that neural maturation appears to occur earlier in areas responsible for handling auditory and visual stimuli than those responsible for EF [Petersen, 2017; Toga et al., 2006]. The formation of concepts (critical for representation) is also rooted in audio-visual and sensorimotor stimuli [Barsalou, 2016]. It may be the case that receiving linguistic input in the form of auditory and visual stimuli promotes early developments in EF, which feed back into further developments in language capacity (represented by the '1?' and '1' arrows between EF and language in the model on p.75). The research debate surrounding the influence of bilingualism on cognitive processing and executive functioning could be relevant here [Kroll & Bialystok, 2013].

Section 3.1 considered the multipurpose functions of language: as a tool for both intra-individual representation and inter-individual communication [Astington & Baird, 2005]. This multidimensional functionality is also reflected in the distinction between recognising the basic referential nature of language, and understanding its representational nature, or possessing meta-linguistic awareness (mLA). Like EF, some basic level of linguistic competence is necessary in order for ToM to emerge, as without it, one cannot access opportunities to interact with others and make connections between their observed behaviour and (expressed) mental states. This idea of a base threshold of language capacity seems to encompass nicely the theories advancing the role in ToM development of specific aspects of language such as semantics (3.1.1) and syntax (3.1.2), or general language development (3.1.4). This linguistic interface between words for concepts and a system of combinatorial rules is clearly an important facilitator of ToM development. However, it seems that the later-acquired meta-representational level of language is also crucial for ToM to develop; this is reflected in the theories espoused with regards to pragmatics (3.1.3) and mLA (3.1) as well as to some extent by those advocating the role of discourse (3.1.4). Considering the associations implied by work on SLI (3.2.1) and late-signing deafness (3.2.2), autism (§4) represents an ideal focus domain within which to explore the relationship between the meta-representational level of language and ToM, as both abilities appear to be delayed and impaired. As discussed in 3.1.3, it seems that at some point in typical development, the meta-level of language stops exerting such an influence on ToM-related processes, possibly because ceiling levels of ability have been reached [Chahboun et al., 2016a], or because these processes have become largely automatised (refer to discussions of automaticity: 2.2.2, 2.2.4, 2.3.3; and automatisation: 3.1.3) [Happé & Frith, 2014]. Contrastingly, it appears that language still plays a role in the autistic population's acquisition and processing of ToM, which may indicate that their abilities are still undergoing development. In this sense, the 'hacking out' theory of language-mediated ToM in autism may simply reflect delayed ToM ability emerging after associated delays in the development and emergence of the underlying linguistic skills. Here, it is interesting that work on figurative language has discovered that priming of the literal meaning facilitates metaphorical comprehension [Weiland et al., 2014]; this may further indicate the importance of language skills for ToM/pragmatic competence. Moreover, while the concept of a base level of language ability supporting the emergence of ToM may also be linked to EF-related developments, the relationship between ToM and the meta-representational level of language cannot be easily explained by EF and WCC accounts (4.4), or by simulation and MNS theories of ToM (2.2.2), because mental states and representations can exist separately from the processes involved in their employment, and it is these distinct representational deficits which seem to manifest in autism.

For the reasons discussed above, and with reference to the themes highlighted here and within the context of the wider literature, this thesis has conceived of the associations between EF, ToM, and language in the form of a holistic tripartite model, integrating both uni- and bi-directional interactions between its three components over a period of staged, co-dependent development. This model of systemic interactions might be represented by something like the figure below, expressing the critical role of language in ToM development, but also allowing for both direct and indirect underlying effects of EF.



[Figure 1: This figure was created by the author of this thesis. The numbers refer to potential causal primacies and stages of developmental interaction between components (EF, ToM, and Language) during the typically developing process of ToM acquisition. Unidirectional arrows represent the initial interactions, whereas bidirectional arrows represent later developments that may be reciprocal, and influence development across the other components.]

It is also important to consider a number of **gaps** in the theoretical and research literature in relation to the conclusions drawn within this thesis. **Firstly**, the diagnosis and origins of disorders such as autism and SLI are not clear-cut. They are still diagnosed behaviourally so it is crucial to bear this lack of clarity in mind when attempting to make assertions about their causes. There are a number of studies emerging within the research that have proposed potential genetic, biological, or neurological markers and causes of autism, but these have not yet been established, and questions remain about gender differences in symptomatic presentation and expression. Another issue is the matter of whether autism is a unitary disorder or comprises various subtypes with differing underlying causes and symptomatic manifestations (previously raised in 4.3). All of these issues require further investigation. **Secondly**, as discussed in 3.1.1, 3.1.2, and 3.1.4, issues of cross-linguistic and cross-cultural differences remain and must be explored. The relationship between language and ToM in autism and other atypically developing populations must also be examined in a cross-linguistic context.

Thirdly, there are problems with the term and conceptualisation of EF, as it has not been established as a clear and coherent construct. However, of the proposed EFs, attentional control appears to be the most concretely substantiated within the literature in relation to language acquisition [Cuevas & Bell, 2014; Loy et al., 2018; Miller & Marcovitch, 2015; Morales et al., 2000]. While thought to reflect the speed of information processing, attentional control is also related to later abilities in memory, language, and intelligence, and correlates with performance on other proposed measures of EF such as working memory, inhibitory control, and cognitive flexibility, independently of verbal ability [Cuevas & Bell, 2014: citing Colombo, Kapa, & Curtindale, 2010]. Miller & Marcovitch's (2015) study found that performance on a battery of EF tasks at 18 months-of-age was predicted by 14-month representational abilities, such as language comprehension and episodes of initiating joint attention [2015]. From their data, they drew the conclusion that transition from early infant behavioural control to more complex preschool EF is supported by representational abilities around 2 years-of-age [Miller & Marcovitch, 2015, p.101]. Research has also found that responding to episodes of joint attention across 6-18 months-of-age is positively correlated with individual differences in vocabulary development [Bates & Dick, 2002; Morales et al., 2000, p.283; Loy et al., 2018]. It may be that a basic, implicit form of attentional control emerges, which is then refined with the advent of language and representational capacity into an explicit form necessary for mentalising [Doherty, 2008]. These ideas would fit nicely within the tripartite model of staged co-development discussed above. Deficits in joint attention have been noted in autism, which have been thought to reflect deficits in ToM [Frith & Happé, 1994; Happé, 1993; Happé & Frith, 2014; Jones et al., 2017; Lombardo et al., 2007; Tager-Flusberg, 2007]. Interestingly, attentional deficits may also explain the overlaps in or comorbidity of ADHD, anxiety, depression, and autism [Happé & Frith, 2014; Jones et al., 2017]. Another relevant area for consideration might be work conducted on the failure to orient towards social stimuli, including the atypical reactivity of the autistic brain's reward system to non-social stimuli and RRBIs [Happé, 1993; Kohls et al., 2018; Rogers, 2006; Tager-Flusberg, 2007].

Finally, it is difficult to pinpoint exactly where instances of atypical development in ToM and language acquisition originate from, as well as what causes them: is there a single cause or are they the outcome of multiple factors interacting with one another? This is further obfuscated by the existence of potential knock-on or even compensatory effects stemming from the wider developmental system (e.g. such as the tripartite model presented above). As discussed by Nilsson & De López, there have been arguments against a language account of ToM, which utilise research on preferential looking by Onishi, Baillargeon, and colleagues (2005, 2010) to assert that toddlers show some implicit understanding of ToM prior to acquiring the linguistic ability to explicitly express these concepts [2016]. There are also those who make such arguments with reference to cases of adults with late-acquired aphasia (discussed in 3.1.2) [Nilsson & De López, 2016; Sodian & Kristen, 2010]. However, both of these counter-arguments to language accounts of ToM could be accommodated within the tripartite staged model considered above. It could be that an 'implicit' ToM actually reflects other abilities such as early EF, like attentional control or children's grasp of secondary representations such as desire and 'prelief' (2.3.1), which enable meta-representational language abilities to develop, which then support further EF and explicit ToM development, and that these ToM abilities become automatised over time and no longer rely on meta-representational linguistic capacity for their expression [Chahboun et al., 2016a; Happé & Frith, 2014]. It may also be the case that specific atypicalities have been more conspicuous in their expression, and thus, have been easier to identify and discern than others, gaining disproportionate focus within the research literature. Moreover, there may be a shared underlying cause of linguistic and ToM deficits, or an underlying cause of linguistic or socio-communicative deficits, which then affects ToM development, causing ToM deficits; so it may be a question of nuance. Here, a discussion of the developmental precursors of ToM is relevant, for example: gaze-following (2.3.2). In autism there is a marked lack of eye contact (one of the DSM-V diagnostic criteria); this is also the case in social anxiety disorder [Schneier et al., 2011]. As discussed in 2.3.3, experiencing anxiety causes impairment of VPT ability [Todd & Simpson, 2016], and thereby, potentially ToM ability as they appear to be part of a shared, broader system. It has also been proposed that autistic RRBIs are a coping mechanism for anxiety [Hollocks et al., 2016; Jones et al., 2017; Joyce et al., 2017] (discussed in 4.1). With this in mind, do intervention strategies such as PACT (introduced in 4.3) merely correct or compensate for deficits in the development of ToM precursors, which then trigger ToM and linguistic development, or do they purely engender linguistic development, which then facilitates ToM development? To some degree, it may be a combination of both. However, it does seem likely that ToM and linguistic impairments in autism result from disrupted development at some later stage, as studies have found no differences in the social behaviour of genetically at-risk infants at 3-6 months-of-age [Happé & Frith, 2014, citing: Ozonoff et al., 2010]; this may provide support for a language account of ToM. In a (2014) review of the literature, Happé & Frith compiled a developmental timeline, highlighting key milestones of social cognition as well as developmental disorders which diverge from the typical developmental pattern at each stage of development [2014]. Their map of social cognition included components such as: agent identification, emotion processing and empathy, mental state attribution, self-processing, and social hierarchy mapping involving social 'policing' and 'in' vs. 'out' group categorisation [Happé & Frith, 2014, p.553]. Happé & Frith asserted that developmental disorders demonstrated dissociable deficits across distinct components of their map, arguing that there was a need for a cross-disorder perspective on atypical development within social neuroscience [2014]. The conceptualisation of a developmental timeline with reference to the findings on autism, late-signing deafness, and SLI, does provide some support for the notion that language impairments drive developmental deficits or delays in ToM, and would fit within the tripartite model discussed above.

While this thesis was intended to be as thorough and comprehensive as possible, there were also some inherent **limitations** in the writing of it. Due to the hugely broad scope of the topic and its interrelated areas, it was not possible to conduct an exhaustive review or analysis of the literature with reference to each specific aspect. Consequently, some (hopefully minor) details may have been overlooked. As was considered above, accounts of the role of language in the emergence, development, acquisition, and expression of ToM are very diverse but also overlap with respect to how directly they consider ToM processes to be facilitated or mediated by language. For these reasons, it is difficult to disentangle and separate them. Since language is a multipurpose tool utilised for both communication and representation and involved across different stages of ToM development, it might even beg the question: can it actually be separated out? It is also important to consider here: what is the fundamental purpose of language? Beyond this, a great many other wider debates within the literature could be relevant to the concepts investigated in this thesis. For example, how are language and concepts represented, stored, and retrieved? Do the linguistic representations overlap with those of ToM concepts? With more time and space, one could also consider: interactions between EF and language; the relationship between EF, ToM, and language deficits in other atypically developing or clinical populations, e.g. schizophrenia; as well as an investigation of emergence vs. expression accounts within these contexts. The finer-grained differences between 'self' vs. 'other' referential cognition, as well as the notion of the absent self in autism could also be explored [Lombardo et al., 2007; van Veluw & Chance, 2014].

The proposed relationship between language and ToM has important **implications** for the treatment of and development of clinical intervention strategies for autism and other language-impaired populations, such as PACT [Pickles et al., 2016; discussed in 4.3]. Intervention strategies could work on how autistic individuals represent language and ToM concepts, improving access to opportunities to engage in linguistic and socio-cognitive development, and employing a language focus to enhance mentalising abilities. In **future research**, it may be prudent to extend investigation of the relationship between language and ToM to the exploration of whether specific language training spurs enhanced ToM development in autism. It would also be interesting to study whether there is language impairment in clinical populations found to have ToM deficits, as well as to consider whether clinical interventions aimed at developing patient competence with mentalising language could facilitate or aid treatment. In this regard, another avenue for research could be the study of the interaction between language and ToM in the context of the development of epistemic trust and parent-child attachment. This would be relevant to clinical practice, for example: the development of the patient-therapist relationship. Greater cross-field collaboration with a holistic focus is also needed, as well as more longitudinal and follow-up studies employing diverse behavioural and neurological measures. In work on ToM, it is also important for future research to embrace the holistic interpretation of ToM as part of a broader PT system, incorporating measures of visual, spatial, and affective PT within these studies.

§6 Conclusions

This thesis has explored the relationship between language and theory of mind (ToM) by conducting a qualitative, critical review of the theoretical and research literature. In §2, this review began by considering the theoretical background of, and developmental perspective on, the concept of ToM, as well as its instantiation within the research. This section also examined a number of accounts that have been proposed since the concept of ToM first appeared within the literature. These have presented varying degrees of perspectives on the extent to which they overlap or conflict with language accounts of ToM. §3 investigated a broad range of language accounts of ToM, considering the roles played by semantics, syntax, pragmatics, and general language development, before turning to the associations between language and ToM ability in children with specific language impairment (SLI) and the late-signing deaf population. Autism is another disorder, which, as yet, is only behaviourally diagnosed, but demonstrates significant, relevant deficits in ToM and language. §4 explored the association of language and ToM within autism as a focus domain, considering ToM accounts of autism, and probing the fit of a language account of ToM in the context of autism, with reference to the executive function (EF) and weak central coherence (WCC) accounts of autism.

A number of important conclusions can be drawn from the undertaking of this critical review. Firstly, meta-representational ToM development appears to be supported by prior developments in EF and language, but also seems to feed back into the further development of EF and language within a multidirectional, co-developmental process of staged interactions. Taking these interactions into account, §5 proposed a tripartite model to express these relationships within a more holistic framework, which could account for the findings across the wider literature and with respect to the delayed and impaired development of language and ToM in atypically developing populations such as SLI and autism. Secondly, regarding the specific influence of language on ToM development, it appears that language plays a critical role in the emergence, development, acquisition, and expression of a ToM, both in terms of its basic function as a tool for communication, and, more importantly, its facilitation of developing meta-representational capabilities, supported by emerging meta-linguistic awareness. Thirdly, delays in the ToM development of language-impaired populations, such as children with SLI, and late-signing deaf children, have ramifications for theoretical accounts of autism, providing support not only for ToM accounts of autism, but also language accounts of ToM deficits in autism. The findings of this thesis bear implications for future research on ToM, autism, and other language-impaired populations, as well as for the treatment of ToM-impaired disorders, and the development of clinical intervention strategies. Future work could investigate the effect of early and specific language training on the ToM development of children diagnosed with autism, as well as the effect of intervention programmes, designed to develop mentalising language competence in clinical populations, on aiding and facilitating psychological or psychiatric treatment.

§7 References

Anwar, A., Abruzzo, P.M., Pasha, S., Rajpoot, K., Bolotta, A., Ghezzo, A., Marini, M., Posar, A., Visconti, P., Thornalley, P.J., & Rabbani, N. (2018). Advanced glycation endproducts, dityrosine and arginine transporter dysfunction in autism - a source of biomarkers for clinical diagnosis. *Molecular Autism,* 9(3).

Astington, J. & Baird, J. (2005). Introduction: Why Language Matters. In Astington, J., & Baird, J. (Eds.), *Why Language Matters for Theory of Mind*, Oxford, UK: Oxford University Press, pp.3-25.

Baillargeon, R., Scott, R.M., & He, Z. (2010). False-belief understanding in infants. *Trends in cognitive sciences*, *14*(3), 110-118.

Bambini, V., Arcara, G., Bechi, M., Buonocore, M., Cavallaro, R., & Bosia, M. (2016). The communicative impairment as a core feature of schizophrenia: frequency of pragmatic deficit, cognitive substrates, and relation with quality of life. *Comprehensive psychiatry*, *71*, 106-120.

Barnes, J.L. & Baron-Cohen, S. (2012). The big picture: Storytelling ability in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *42*(8), 1557-1565.

Baron-Cohen, S. (2017). Editorial Perspective: Neurodiversity–a revolutionary concept for autism and psychiatry. *Journal of Child Psychology and Psychiatry*, *58*(6), 744-747.

Baron-Cohen, S., Leslie, A.M., & Frith, U. (1985). Does the autistic child have a “theory of mind”?. *Cognition*, *21*(1), 37-46.

Barsalou, L.W. (2016). Situated Conceptualization: Theory and Application. In Coello, Y., & Fischer, M.H. (Eds.), *Foundations of embodied cognition: Volume 1. Perceptual and Emotional Embodiment*, East Sussex, UK: Psychology Press, pp.11-37.

Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental psychobiology*, *40*(3), 293-310.

Bird, G., & Viding, E. (2014). The self to other model of empathy: providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience & Biobehavioral Reviews*, *47*, 520-532.

Bradford, E.E., Jentzsch, I., & Gomez, J.C. (2015). From self to social cognition: theory of mind mechanisms and their relation to executive functioning. *Cognition*, *138*, 21-34.

Brent, E., Rios, P., Happé, F., & Charman, T. (2004). Performance of children with autism spectrum disorder on advanced theory of mind tasks. *Autism*, *8*(3), 283-299.

Bretherton, I. (1984). Representing the social world in symbolic play: Reality and fantasy. In Bretherton, I. (Ed.), *Symbolic play*, New York: Academic Press, pp.3-41.

Byom, L.J., & Mutlu, B. (2013). Theory of mind: Mechanisms, methods, and new directions. *Frontiers in human neuroscience*, *7*.

Carlson, S.M., Claxton, L.J., & Moses, L.J. (2015). The relation between executive function and theory of mind is more than skin deep. *Journal of Cognition and Development*, *16*(1), 186-197.

Carlson, S.M., & Moses, L.J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child development*, *72*(4), 1032-1053.

Chahboun, S., Vulchanov, V., Saldaña, D., Eshuis, H., & Vulchanova, M. (2016a). Predictors of metaphorical understanding in high functioning autism. *Lingue e linguaggio*, *15*(1), 29-58.

Chahboun, S., Vulchanov, V., Saldaña, D., Eshuis, H., & Vulchanova, M. (2016b). Can you play with fire and not hurt yourself? A comparative study in figurative language comprehension between individuals with and without autism spectrum disorder. *PloS one*, *11*(12), e0168571.

Corcoran, R., Mercer, G., & Frith, C.D. (1995). Schizophrenia, symptomatology and social inference: investigating "theory of mind" in people with schizophrenia. *Schizophrenia research*, *17*(1), 5-13.

Craig, F., Margari, F., Legrottaglie, A.R., Palumbi, R., de Giambattista, C., & Margari, L. (2016). A review of executive function deficits in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Neuropsychiatric disease and treatment*, *12*, 1191-1202.

Cuevas, K., & Bell, M.A. (2014). Infant attention and early childhood executive function. *Child Development*, *85*(2), 397-404.

Cummings, L. (2014). *Pragmatic Disorders.* Amsterdam: Springer.

DeLoache, J.S., & Burns, N.M. (1994). Early understanding of the representational function of pictures. *Cognition*, *52*(2), 83-110.

Dennett, D.C. (1978). Beliefs about beliefs. Behavioral and Brain Sciences, 1, 568–570.

de Villiers, P.A. (2005). The Role of Language in Theory-of-Mind Development: What Deaf Children Tell Us. In Astington, J., & Baird, J. (Eds.), *Why Language Matters for Theory of Mind*, Oxford, UK: Oxford University Press, pp.266-297.

Doherty, M. (2008). *Theory of mind: How children understand others' thoughts and feelings*. Psychology Press.

Dunn, J., & Brophy, M. (2005). Communication, Relationships, and Individual Differences in Children's Understanding of Mind. In Astington, J., & Baird, J. (Eds.), *Why Language Matters for Theory of Mind*, Oxford, UK: Oxford University Press, pp.50-69.

Dunphy-Lelii, S., & Wellman, H.M. (2004). Infants' understanding of occlusion of others' line-of-sight: Implications for an emerging theory of mind. *European Journal of Developmental Psychology*, *1*(1), 49-66.

Durrleman, S. & Franck, J. (2015). Exploring links between language and cognition in autism spectrum disorders: Complement sentences, false belief, and executive functioning. *Journal of communication disorders*, *54*, 15-31.

Eberhardt, M., & Nadig, A. (2016). Reduced sensitivity to context in language comprehension: A characteristic of Autism Spectrum Disorders or of poor structural language ability?. *Research in developmental disabilities*.

Farrant, B. (2015). Specific Language Impairment and Perspective Taking: Delayed Development of Theory of Mind, Visual and Emotional Perspective Taking. *Journal of Childhood & Developmental Disorders*, *1*(1), 8.

Farrant, B.M., Fletcher, J., & Maybery, M.T. (2006). Specific language impairment, theory of mind, and visual perspective taking: Evidence for simulation theory and the developmental role of language. *Child development*, *77*(6), 1842-1853.

Fine, C., Lumsden, J., & Blair, R.J.R. (2001). Dissociation between 'theory of mind' and executive functions in a patient with early left amygdala damage. *Brain*, *124*(2), 287-298.

Flavell, J.H., Flavell, E.R., & Green, F.L. (1983). Development of the appearance-reality distinction. *Cognitive psychology*, *15*(1), 95-120.

Fodor, J.A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Frank, C.K., Baron-Cohen, S., & Ganzel, B.L. (2015). Sex differences in the neural basis of false-belief and pragmatic language comprehension. *NeuroImage*, *105*, 300-311.

Frith, C. & Frith, U. (2005). Theory of mind. *Current Biology*, *15*(17), R644-R645.

Frith, U. & Happé, F. (1994). Autism: beyond “theory of mind”. *Cognition*, *50*(1), 115-132.

Gadow, K.D. (2012). Schizophrenia spectrum and attention-deficit/hyperactivity disorder symptoms in autism spectrum disorder and controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(10), 1076-1084.

Garnham, W.A., & Ruffman, T. (2001). Doesn’t see, doesn’t know: is anticipatory looking really related to understanding or belief?. *Developmental Science*, *4*(1), 94-100.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child development*, 26-37.

Griffin, R. & Baron-Cohen, S. (2002). The Intentional Stance: Developmental and Neurocognitive Perspectives. In Brook, A., & Ross, D. (Eds.), *Daniel Dennett*, Cambridge, UK: Cambridge University Press, pp.83-116.

Hamilton, A.F.D.C., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition*, *113*(1), 37-44.

Happé, F. & Frith, U. (2014). Annual research review: towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry*, *55*(6), 553-577.

Happé, F. & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of autism and developmental disorders*, *36*(1), 5-25.

Happé, F. (1995a). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child development*, *66*(3), 843-855.

Happé, F. (1995b). Understanding minds and metaphors: Insights from the study of figurative language in autism. *Metaphor and symbol*, *10*(4), 275-295.

Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders*, *24*(2), 129-154.

Happé, F. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, *48*(2), 101-119.

Harris, P. (2005). Conversation, Pretense, and Theory of Mind. In Astington, J., & Baird, J. (Eds.), *Why Language Matters for Theory of Mind*, Oxford, UK: Oxford University Press, pp.70-83.

Harris, P. (1992). From simulation to folk psychology: The case for development. *Mind & Language*, *7*(1-2), 120-144.

Hogrefe, G.J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child development*, 567-582.

Hollocks, M.J., Jones, C.R., Pickles, A., Baird, G., Happé, F., Charman, T., & Simonoff, E. (2014). The association between social cognition and executive functioning and symptoms of anxiety and depression in adolescents with autism spectrum disorders. *Autism Research*, *7*(2), 216-228.

Howes, C., & Matheson, C.C. (1992). Sequences in the development of competent play with peers: Social and social pretend play. *Developmental Psychology*, *28*(5), 961-974.

Jacques, S., & Zelazo, P.D. (2005). Language and the Development of Cognitive Flexibility: Implications for Theory of Mind. In Astington, J., & Baird, J. (Eds.), *Why Language Matters for Theory of Mind*, Oxford, UK: Oxford University Press, pp.144-162.

Jones, C.R., Simonoff, E., Baird, G., Pickles, A., Marsden, A.J., Tregay, J., Happé, F., & Charman, T. (2017). The association between theory of mind, executive function, and the symptoms of autism spectrum disorder. *Autism Research*.

Joyce, C., Honey, E., Leekam, S. R., Barrett, S. L., & Rodgers, J. (2017). Anxiety, Intolerance of Uncertainty and Restricted and Repetitive Behaviour: Insights Directly from Young People with ASD. *Journal of Autism and Developmental Disorders*, 1-14.

Kavanaugh, R.D., & Harris, P.L. (1994). Imagining the outcome of pretend transformations: Assessing the competence of normal children and children with autism. *Developmental Psychology*, *30*(6), 847.

Kohls, G., Antezana, L., Mosner, M.G., Schultz, R.T., & Yerys, B.E. (2018). Altered reward system reactivity for personalized circumscribed interests in autism. *Molecular Autism*, *9*(1), 9.

Kroll, J.F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497-514.

Lai, M.C., Lombardo, M.V., Ruigrok, A.N., Chakrabarti, B., Wheelwright, S.J., Auyeung, B., Allison, C., Baron-Cohen, S., & MRC AIMS Consortium. (2012). Cognition in males and females with autism: similarities and differences. *PLoS One*, *7*(10), e47198, 1-15.

Lamm, C., & Majdandžić, J. (2015). The role of shared neural activations, mirror neurons, and morality in empathy–a critical comment. *Neuroscience Research*, *90*, 15-24.

Le Barton, E.S., & Iverson, J.M. (2016). Gesture development in toddlers with an older sibling with autism. *International journal of language & communication disorders*, *51*(1), 18-30.

Leslie, A.M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, *50*(1), 211-238.

Leslie, A.M., Friedman, O., & German, T.P. (2004). Core mechanisms in 'theory of mind'. *Trends in cognitive sciences*, *8*(12), 528-533.

Lillard, A.S. (1993). Pretend play skills and the child's theory of mind. *Child development*, *64*(2), 348-371.

Lohmann, H., Tomasello, M., & Meyer, S. (2005). Linguistic Communication and Social Understanding. In Astington, J., & Baird, J. (Eds.), *Why Language Matters for Theory of Mind*, Oxford, UK: Oxford University Press, pp.245-265.

Lombardo, M.V., Barnes, J.L., Wheelwright, S.J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PLoS One*, *2*(9), e883.

Loy, M., Masur, E.F., & Olson, J. (2018). Developmental changes in infants' and mothers' pathways to achieving joint attention episodes. *Infant Behavior and Development*, *50*, 264-273.

Mahy, C.E., Moses, L.J., & Pfeifer, J.H. (2014). How and where: Theory-of-mind in the brain. *Developmental cognitive neuroscience*, *9*, 68-81.

Meert, G., Wang, J., & Samson, D. (2017). Efficient belief tracking in adults: The role of task instruction, low-level associative processes and dispositional social functioning. *Cognition*, *168*, 91-98.

Meinhardt, J., Kühn-Popp, N., Sommer, M., & Sodian, B. (2012). Distinct neural correlates underlying pretense and false belief reasoning: Evidence from ERPs. *NeuroImage*, *63*(2), 623-631.

Miller, S.E., & Marcovitch, S. (2015). Examining executive function in the second year of life: Coherence, stability, and relations to joint attention and language. *Developmental Psychology*, *51*(1), 101.

Milligan, K., Astington, J.W., & Dack, L.A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child development*, *78*(2), 622-646.

Morales, M., Mundy, P., Delgado, C.E., Yale, M., Messinger, D., Neal, R., & Schwartz, H.K. (2000). Responding to joint attention across the 6-through 24-month age period and early language acquisition. *Journal of applied developmental psychology*, *21*(3), 283-298.

Nelson, K. (2005). Language Pathways into the Community of Minds. In Astington, J., & Baird, J. (Eds.), *Why Language Matters for Theory of Mind*, Oxford, UK: Oxford University Press, pp.26-49.

Nilsson, K.K., & De López, K.M.J. (2016). Theory of Mind in children with Specific Language Impairment: A systematic review and meta-analysis. *Child development*, *87*(1), 143-153.

Norbury, C.F. (2005). The relationship between theory of mind and metaphor: Evidence from children with language impairment and autistic spectrum disorder. *British Journal of Developmental Psychology*, *23*(3), 383-399.

Norman, D.A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In Davidson, R.J., Schwartz, G.E., & Shapiro, D. (Eds.), *Consciousness and Self-Regulation (Vol. 4): Advances in Research and Theory*, New York: Plenum Press, pp.1-18.

O'Neill, D.K. (2005). Talking about "New" Information: The Given/New Distinction and Children's Developing Theory of Mind. In Astington, J., & Baird, J. (Eds.), *Why Language Matters for Theory of Mind*, Oxford, UK: Oxford University Press, pp.84-105.

Ozonoff, S., Cook, I., Coon, H., Dawson, G., Joseph, R.M., Klin, A., McMahon, W.M., Minshew, N., Munson, J.A., Pennington, B.F. and Rogers, S.J., (2004). Performance on Cambridge Neuropsychological Test Automated Battery subtests sensitive to frontal lobe function in people with autistic disorder: evidence from the Collaborative Programs of Excellence in Autism network. *Journal of autism and developmental disorders*, *34*(2), 139-150.

Parish-Morris, J., Liberman, M.Y., Cieri, C., Herrington, J.D., Yerys, B.E., Bateman, L., Donaher, J., Ferguson, E., Pandey, J., & Schultz, R.T. (2017). Linguistic camouflage in girls with autism spectrum disorder. *Molecular autism*, *8*(1), 48.

Pearson, A., Ropar, D., & Hamilton, A.F.D.C. (2013). A review of visual perspective taking in autism spectrum disorder. *Frontiers in human neuroscience*, *7*.

Perner, J. (1999). Theory of mind. In M. Bennett (Ed.), *Developmental psychology: Achievements and prospects*, New York, NY, US: Psychology Press, pp.205-230.

Perner, J. (1991). *Understanding the representational mind*. The MIT Press.

Perner, J., Baker, S., & Hutton, D. (1994). Prelief: The conceptual origins of belief and pretence. *Children's early understanding of mind: Origins and development*, 261-286.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…" attribution of second-order beliefs by 5- to 10-year-old children. *Journal of experimental child psychology*, *39*(3), 437-471.

Petersen, A.C. (2017). *Brain maturation and cognitive development: Comparative and cross-cultural perspectives*. New York: Routledge.

Peterson, C.C., & Siegal, M. (2000). Insights into theory of mind from deafness and autism. *Mind & Language*, *15*(1), 123-145.

Pickles, A., Le Couteur, A., Leadbitter, K., Salomone, E., Cole-Fletcher, R., Tobin, H., Gammer, I., Lowry, J., Vamvakas, G., Byford, S., & Aldred, C. (2016). Parent-mediated social communication therapy for young children with autism (PACT): long-term follow-up of a randomised controlled trial. *The Lancet*, *388*(10059), 2501-2509.

Povinelli, D.J., & DeBlois, S. (1992). Young children's (Homo sapiens) understanding of knowledge formation in themselves and others. *Journal of Comparative Psychology*, *106*(3), 228.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and brain sciences*, *1*(4), 515-526.

Repacholi, B.M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14-and 18-month-olds. *Developmental psychology*, *33*(1), 12.

Rogers, J., Viding, E., Blair, R.J., Frith, U., & Happe, F. (2006). Autism spectrum disorder and psychopathy: shared cognitive underpinnings or double hit?. *Psychological medicine*, *36*(12), 1789-1798.

Samson, D., Apperly, I.A., Braithwaite, J.J., Andrews, B.J., & Scott, S.E.B. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. Journal of Experimental Psychology - Human Perception and Performance, 36, 1255–1266.

Schaafsma, S.M., Pfaff, D.W., Spunt, R.P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in cognitive sciences*, *19*(2), 65-72.

Schneier, F.R., Rodebaugh, T.L., Blanco, C., Lewin, H., & Liebowitz, M.R. (2011). Fear and avoidance of eye contact in social anxiety disorder. *Comprehensive Psychiatry*, *52*(1), 81-87.

Schurz, M., Kronbichler, M., Weissengruber, S., Surtees, A., Samson, D., & Perner, J. (2015). Clarifying the role of theory of mind areas during visual perspective taking: Issues of spontaneity and domain-specificity. *NeuroImage*, *117*, 386-396.

Sodian, B., & Kristen, S. (2010). Theory of mind. In B.M. Glatzeder et al. (Eds.), *Towards a Theory of Thinking*, Springer-Verlag Berlin Heidelberg, pp.189-201.

Stanovich, K.E., & Toplak, M.E. (2012). Defining features versus incidental correlates of Type 1 and Type 2 processing. *Mind & Society*, *11*(1), 3-13.

Surtees, A., Samson, D., & Apperly, I. (2016). Unintentional perspective-taking calculates whether something is seen, but not how it is seen. *Cognition*, *148*, 97-105.

Tager-Flusberg, H. (2007). Evaluating the theory-of-mind hypothesis of autism. *Current directions in psychological science*, *16*(6), 311-315.

Tager-Flusberg, H., & Joseph, R.M. (2005). How Language Facilitates the Acquisition of False-Belief Understanding in Children with Autism. In Astington, J., & Baird, J. (Eds.), *Why Language Matters for Theory of Mind*, Oxford, UK: Oxford University Press, pp.298-318.

Tick, B., Colvert, E., McEwen, F., Stewart, C., Woodhouse, E., Gillan, N., Hallett, V., Lietz, S., Garnett, T., Simonoff, E., & Ronald, A. (2016). Autism Spectrum Disorders and other mental health problems: Exploring etiological overlaps and phenotypic causal associations. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*(2), 106-113.

Todd, A.R., & Simpson, A.J. (2016). Anxiety impairs spontaneous perspective calculation: Evidence from a level-1 visual perspective-taking task. *Cognition*, *156*, 88-94.

Toga, A.W., Thompson, P.M., & Sowell, E.R. (2006). Mapping brain maturation. *Trends Neurosci., 29(3),* 148–159.

Tsuchida, A., & Fellows, L.K. (2013). Are core component processes of executive function dissociable within the frontal lobes? Evidence from humans with focal prefrontal damage. *Cortex*, *49*(7), 1790-1800.

van Veluw, S.J., & Chance, S.A. (2014). Differentiating Between Self and Others: an ALE Meta-Analysis of fMRI Studies of Self-Recognition and Theory of Mind. *Brain imaging and behavior*, *8*(1), 24-38.

Vulchanova, M., Saldaña, D., Chahboun, S., & Vulchanov, V. (2015). Figurative language processing in atypical populations: the ASD perspective. *Frontiers in human neuroscience*, *9*, 24.

Vulchanova, M., Talcott, J.B., Vulchanov, V., & Stankova, M. (2012a). Language against the odds, or rather not: The weak central coherence hypothesis and language. *Journal of Neurolinguistics*, *25*(1), 13-30.

Vulchanova, M., Talcott, J.B., Vulchanov, V., Stankova, M., & Eshuis, H. (2012b). Morphology in autism spectrum disorders: local processing bias and language. *Cognitive neuropsychology*, *29*(7-8), 584-600.

Weiland, H., Bambini, V., & Schumacher, P.B. (2014). The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. *Frontiers in human neuroscience*, *8*, 583.

Wellman, H.M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child development*, *72*(3), 655-684.

White, S.J. (2013). The Triple I Hypothesis: Taking Another('s) Perspective on Executive Dysfunction in Autism. *Journal of Autism and Developmental Disorders*, *43*(1), 114-121.

Whiten, A., & Suddendorf, T. (2001). Meta-representation and secondary representation. *Trends in Cognitive Sciences*, *5*(9), 378.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103-128.