Psychology Lancaster University

Investigating the role of attentional biases on language development in autistic children and neurotypical children

Charlotte Rothwell

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Psychology

Lancaster University

July 2023

Table of Contents

[submission in alternative format form; removed as per Lancaster University guidelines]

Declaration

I declare that this thesis is entirely my own work completed under the supervision of Calum Hartley and Gert Westermann (author contributions are listed at the start of chapters). None of the work in this thesis has been submitted elsewhere in support of application for another degree at this or any other institution.

The parts of this thesis that are presented in publishable paper format, or that have been accepted for publication, are indicated at the beginning of the relevant chapter along with author contributions for each.

This thesis is 65,992 words including footnotes, tables, figures, and citations, thus does not exceed the permitted maximum. The word count does not include items that do not comprise the main text of the thesis, such as the table of contents, title page, declaration, list of tables, list of figures, acknowledgements, abstract, appendices, or bibliography.

Charlotte Rothwell (2023)

List of tables

Chapter 2

Chapter 3

Chapter 4

Table 15. Summaries of the fixed effects in the final generalised linear mixed-effects models (log odds) for 24-hour retention accuracy, predicted by proportion of time spent looking at the target stimuli

Table 16. Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) for number of looks towards the target stimuli during 24-hour retention

Table 17. Summary of the fixed effects in the final generalised linear mixed-effects model (log odds) of 24-hour retention accuracy, predicted by the longest look to the target/foil stimuli during 24-hour retention

List of figures

Chapter 2

Chapter 3

Chapter 4

Acknowledgements

Amid a global pandemic, undertaking this PhD has, at times, been challenging. However, I am very lucky to have an amazing support network and am grateful for all the people who helped me to achieve this PhD. Without you, it would not have been possible.

Of course, I have been incredibly lucky to have two very knowledgeable and supportive supervisors, Calum Hartley and Gert Westermann. Thank you for guiding me through this journey. Calum, thank you for your endless patience and reassurance, for always keeping your door open, and for answering my many questions, even when the pandemic meant I could no longer pop my head into your office. In the face of challenges, your continued dedication and enthusiasm inspired me to continue this journey. It has been a privilege to be your PhD student, and I truly cannot thank you enough. Gert, without you I would not have embarked on my research journey. As an undergraduate research assistant, you welcomed me into your research group where I was inspired and amazed by your knowledge and expertise. Your encouragement and initiatives afforded me the confidence and motivation to undertake this PhD, and for that I will be forever grateful.

I am very grateful for Lancaster BabyLab, especially for the wonderful managers Kat Kaduk and Uschi Mason who have supported me on my research journey, and always provided a friendly and welcoming face. Thank you to our incredible technicians Peter Tovee and Barrie Usherwood for designing and creating our bespoke experimental set up. Without your ideas and incredible knowledge these studies genuinely would not have been possible!

Of course, I must thank all the schools, teachers, caregivers, and children who made this research possible. Your dedication to advancing science and willingness to support my PhD work will forever be appreciated. I will treasure all the hand-drawn pictures and cards I received during my data collection. Thank you also to my research assistants Ana Ferreras

Polo, Becky Warne, Emily Duckett, Fi Catherall, Joanna Meacham, Lydia Brooks, Lucy Dixey, Maisie Wong, Paisleigh Wilson-Thomas, Rachel Ivory, and Sarah Fowler. Your contribution to data coding and processing was invaluable. Whilst it was always sad to see my research assistants move onto their new journeys, I am grateful that I had the opportunity to supervise so many students and am proud of the destinations you achieved.

I am grateful for my fellow PhD students who joined me on this journey, you are all incredible. Thank you, Rebecca Kvisler, for showing me the ropes and for making the beginning of my PhD a much smoother transition! Kim Hudspeth, thank you for always being by my side – although our PhD journeys did not progress as anticipated, your support and lovely notes left on my desk were always appreciated. Elise Oosterhuis, thank you for keeping me company in the office, and for sharing this rollercoaster ride. Thank you, Leigh Keating, for your encouragement and conference companionship. Sophie Lund, thank you not only for working with me on our meta-analysis project, but for your support and commitment. You made doing a PhD during a pandemic less lonely, and your attitude to life is inspirational.

Thank you to my family and friends for everything. Your support and motivation allowed me to achieve my goals and I am very fortunate to be surrounded by so many incredible people. To my brothers, Alex and Will, thank you for being my work from home companions, and for always bringing a sense of connection and encouragement. Thank you to my grandparents Carol and Stan for always believing in me. To my grandparents watching down on me, Jim and Joyce, I hope I can always make you proud. Of course, I cannot conclude this section without thanking my incredible Mum and Dad. From the beginning, you shaped my academic journey. Thank you for showing me the importance of hard work and dedication, and for your unwavering support. I am so lucky to have you as my parents.

Investigating the role of attentional biases on language development in autistic children and neurotypical children

Charlotte Rothwell

Thesis abstract

From their first year of life, neurotypical children learn words rapidly and with apparent ease. However, many autistic children experience delays in language development and struggle to acquire vocabulary. To learn a word, children must identify the intended referent (referent selection) and then store the word-referent association in memory for later retrieval (retention). In light of recent evidence that referent selection and retention mechanisms are unimpaired in autism, difficulties acquiring vocabulary may be attributed to differences in attention to stimuli. Autistic children often focus intensely on preferred stimulus categories, potentially disregarding stimuli that do not align with their preferential interests. Therefore, the purpose of this thesis was to examine how attentional salience of stimuli influences autistic and neurotypical children's novel word learning.

The three studies that comprise this thesis investigated how children's stimulus preferences and interests influence their referent selection and retention after delays of 5 minutes and 24 hours. Participants were neurotypical and autistic children with delayed language development matched on receptive vocabulary (age equivalent ~5 years). All studies utilised touch-screen technology to capture children's response times and accuracy. Studies 1 and 2 examined how children's word learning is impacted by the presence of target and distractor stimuli related to a preferred stimulus category. 'High interest' stimuli in these studies were novel animals, as they are often favoured by both neurotypical and autistic children. Participants' interest in animals was confirmed via a questionnaire. Study 3 investigated how children's individual preferences for novel objects (i.e. generic experimental stimuli) affect novel word learning. Studies 2 and 3 employed a novel combination of

accuracy and looking measures, exploring children’s in-trial visual attention and how this predicted children’s response accuracy.

In Study 1, autistic and neurotypical children were taught novel words associated with unfamiliar animals (high interest stimuli) and objects (neutral interest stimuli) in the presence of high interest animal distractors during referent selection. This manipulation enabled us to investigate how salient non-target stimuli affect novel word learning. The presence of interesting distractor stimuli enhanced autistic children’s encoding of novel word-referent associations by increasing attention across the whole visual scene. Study 2 studied how children’s categorical interests influenced their word learning. Children learnt names for novel animals and novel objects, but the distractor stimuli presented at referent selection were neutral interest familiar objects, allowing us to identify how perceptual distinctions between stimulus categories influences novel word-referent encoding. Categorically interesting novel stimuli, that stood out from less interesting distractors, enhanced autistic children’s retention of novel words after 5 minutes. However, autistic children demonstrated a substantial increase in their retention of novel object labels between 5 minutes and 24 hours, outperforming neurotypical peers in both conditions following sleep consolidation. Study 3 examined how children’s individual preferences for novel objects impacts word learning (labels were taught for liked and disliked objects, identified by children in a preceding task). The findings showed that idiosyncratic preferences for novel experimental stimuli, an object category that is unlikely to align with children’s pre-existing interests, do not affect autistic or neurotypical children’s attention allocation to an extent that is sufficient to impact retention after 5 minutes or 24 hours. Our looking data in Studies 2 and 3 indicated that greater visual attention towards stimuli predicted novel word retention. Moreover, observed population differences in visual attention did not directly translate to diminished behavioural accuracy.

This thesis advances understanding of relationships between word learning processes and attentional mechanisms in autism and neurotypical development. Our data demonstrate that categorically interesting stimuli aid novel word learning in autistic children. This finding, coupled with observed relationships between visual input at referent selection and novel word retention, confirm that increased attentional allocation to stimuli during encoding leads to more robust novel word-referent representations. Contrasts between population differences in visual attention and response accuracy outcomes suggest that autistic children potentially achieve similar results through a different route than neurotypical peers. Our novel combination of visual attention and accuracy measures highlight the importance of choosing dependent variables that appropriately target mechanisms of interest and the risk of drawing inaccurate conclusions based on looking behaviour alone. Given that stimulus categories influenced autistic and neurotypical children's learning to different extents, researchers should carefully select experimental stimuli that appeal to different populations comparably. Overall, our findings have the potential to inform practice in clinical and educational settings, providing insight into how learning contexts can be optimised for maximally efficient language acquisition.

Chapter 1: Introduction and Literature Review

1.1. Thesis Introduction

Language development is one of the most significant milestones in a young child's cognitive and social development (Carpenter et al., 1998), and word learning is a crucial part of this process. Acquiring language is necessary for children to be able to understand and engage with the world around them (Nelson, 2007). Neurotypical children acquire language rapidly in the early years, uttering their first words at around 12 months (Samuelson & McMurray, 2017), and learning the meanings of over two hundred words before two years of age (Dale & Fenson, 1996). By school-age, neurotypical children are thought to learn as many as ten new words per day (Bloom, 2002a), enriching their social development as well as academic achievement (Durkin et al., 2012). However, this process can present challenges for some children, including those with neurodevelopmental delays (Luyster et al., 2008; Weismer et al., 2010).

Autism is a pervasive neurodevelopmental disorder that affects around 1% of the population (Lai et al., 2014). It is characterised by a broad range of symptoms which manifest in various ways, including social and communicative difficulties, restricted and repetitive interests and behaviours, and language delay (Kjelgaard & Tager-Flusberg, 2001). Autism is up to four times more likely to be diagnosed in males than females (Baron-Cohen et al., 2011). This may be partially attributed to the fact that autistic females are more likely to mask certain diagnosis-defining characteristics, leading to a decreased likelihood of receiving a diagnosis (e.g. Goldman, 2013; Hull et al., 2017). Some research also suggests that females may be genetically less likely to develop autism (Robinson et al., 2013; Skuse, 2000). While a single identifiable cause of autism is yet unknown, genetic factors such as differences in chromosomal makeup and brain structures have been identified in autistic people (De Fossé et al., 2004; Hyde et al., 2010; Volkmar et al., 2005).

The vastly heterogenous nature of autism means that autistic individuals experience symptoms to varying degrees (Abrahams & Geschwind, 2008). One significant area of symptomatology in autism impacts attentional mechanisms. Restricted, Repetitive Patterns of Behaviour, Interests, or Activities (RRBIs) have been included as a core feature of autism since its first conceptualisation (Kanner, 1943), and remain a key diagnostic domain in the most recent revision of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V; American Psychiatric Association, 2013). These behaviours are high in frequency, and often pertain to very narrow topics or focal points (Richler et al., 2007). They can manifest in several ways, but some examples include insistency on sameness in the environment or routines, or intense perseverative focus on specific topics, interests, and activities (Kanner, 1943). These behaviours influence autistic children's daily lives in numerous ways and can relate to poorer social cognition and language development (Larkin et al., 2017; Ray-Subramanian & Weismer, 2012). Studies suggest that language acquisition may be affected by RRBIs due to children not flexibly attending to the correct audio-visual stimuli, and instead retaining focus on their specific interests (Venker et al., 2018). This potentially creates limited or spurious associations between novel words and referents (Elsabbagh et al., 2009; Tenenbaum et al., 2017), a point that will be revisited later in this chapter.

Whereas neurotypical children commonly begin to speak around their first birthday, this is often delayed until around preschool age in autism (Ellis Weismer & Kover, 2015; Howlin, 2003). Consequently, autistic children demonstrate a varied range of language proficiencies, with language delay defined as one of three categories: non-verbal, minimally verbal, and pre-verbal (DiStefano & Kasari, 2016; Norrelgen et al., 2015). Language use differs across these categories, with minimally verbal children using five or fewer functional words in speech (excluding echolalia, repetitions of recently heard or scripted speech; Kasari et al., 2008; Tager-Flusberg & Kasari, 2013; Tek et al., 2014). Although many autistic children do acquire functional spoken language eventually (Lord et al., 2004), a significant number remain minimally verbal throughout their lifetime, presenting profound difficulties in the acquisition and usage of spoken words (Anderson et al., 2007; Pickles et al., 2014). Delays in receptive vocabulary are often more profound than expressive language difficulties, suggesting that some autistic children have problems acquiring novel words (Charman et al., 2003; Ellis Weismer et al., 2020; Volden et al., 2011). However, language impairments in autism depend largely upon age and developmental level, with receptive vocabulary impairments often appearing within the first two years of life (Mody & Belliveau, 2013).

Promisingly, the overall proportion of minimally verbal children in the autistic population appears to be reducing. Earlier studies suggest approximately 50% of autistic children are minimally verbal (Prizant, 1983; Sigman & McGovern, 2005), but more recent studies highlight proportions of around 25-30% (Tager-Flusberg & Kasari, 2013). This could be due to earlier diagnosis and intervention (Goldstein, 2002), as well as more effective intervention programmes due to advancements in knowledge and healthcare (Rogers & Vismara, 2008; Smith & Eikeseth, 2011). In the DSM-IV, autism could be diagnosed in the absence of language difficulties (e.g. Asperger's; American Psychiatric Association, 1994), and in the DSM-V, language difficulties were removed from the diagnostic criterion altogether (American Psychiatric Association, 2013). This change was made due to the broad trajectory language acquisition can follow for both autistic and non-autistic individuals (Matson & Neal, 2010; Solomon et al., 2011). Despite this, while verbally able individuals are now more likely to receive an autism diagnosis (Centers for Disease Control and Prevention, 2012), language delay remains evident in many autistic children, noted both by caregivers and teachers, as well as through assessment by professionals (Coonrod & Stone, 2004; Eigsti et al., 2011; Tager-Flusberg et al., 2001). As such, we must continue to strive to understand how word learning processes compare across neurotypical and autistic children, to discover how we can advance children's learning further.

When investigated experimentally, some studies show that autistic children present word learning deficits under controlled conditions (e.g. McDuffie et al., 2006a; Parish-Morris et al., 2007), whereas other studies suggest that autistic children can learn words with comparable accuracy to neurotypical children (e.g. Haebig et al., 2017; Hartley et al., 2019). Indeed, recent studies consistently find that fundamental word learning mechanisms of autistic children are unimpaired (e.g. Hartley et al., 2019, 2020). Several studies highlight that language delays may result from atypical attention to information in the social environment (Arunachalam & Luyster, 2015). For instance, some autistic children demonstrate difficulties in attending to external social cues, such as joint attention and intention, and applying these to word learning contexts (Baron-Cohen et al., 1997; Fletcher-Watson et al., 2009; Leekam & Ramsden, 2006, Parish-Morris et al., 2007).

The course of language development for autistic children is diverse. Therefore, it remains unclear exactly why many autistic children present delayed receptive language skills. If it is a consequence of diminished attention, then perhaps discovering ways in which children's attention can be increased or maintained in learning contexts would highlight strengths that can be utilised to scaffold word learning. It is important to address these language difficulties early in childhood as language proficiency can predict longer term outcomes, so examining how we can increase language acquisition is imperative for promoting the developmental trajectory of autistic children (Lord & Paul, 1997; Stone & Yoder, 2001).

Indeed, the RRBIs that autistic children demonstrate can present in a variety of forms, including focus on a circumscribed range of topics of interest, insistence on sameness, and stereotyped behaviours and motor movements (Honey et al., 2012). This attunes autistic children's attention to specific, narrowed focal points. Prior research emphasises that RRBIs cause autistic children to focus on a constrained range of specific stimuli, restricting intake of information from the environment (e.g. Akechi et al., 2011, 2013; Arunachalam & Luyster, 2018). This can be detrimental for word learning, as it may mean that children lose important informants from the environment or focus on incorrect target referents during novel word learning, limiting novel word acquisition or causing incorrect information to be encoded (Elsabbagh et al., 2009; Tenenbaum et al., 2014, 2017; Venker et al., 2018). Despite the clear relationship between attention and word learning, and autistic children's atypical attentional mechanisms, little research has directly investigated how attention allocation influences novel word learning in autism.

Whilst previous studies suggest that autistic children focus on stimuli, or aspects of stimuli, that appeal to their circumscribed interests (Sasson et al., 2008, 2011), it remains unclear how children's specific interests impact their word learning. This thesis therefore aims to investigate how children's interests affect their attentional allocation, and how this impacts word learning. Given the importance of language acquisition for children's development, investigating under what conditions autistic children can learn words, and how individual characteristics and interests can influence word learning, allows us to decipher exactly *how* we can facilitate their language learning. These discoveries have the potential to impact interventions in clinical and educational settings.

This literature review begins with an overview of word learning in neurotypical development, followed by a review of how attentional allocation influences word learning in this population. Once these areas are addressed in neurotypical development, an overview of how word learning manifests in autism is provided, and how attention affects autistic children's word learning. Finally, an outline of the studies that comprise the thesis are presented. Overall, this literature review will highlight gaps within the literature that will be explored within the studies that comprise the thesis.

1.2. Word Learning

Words are the fundamental building blocks of language (Pinker & Jackendoff, 2005). Undoubtedly, word learning is a crucial developmental achievement that is necessary for children to understand the world around them (Carpenter et al., 1998). Neurotypical children acquire a new word every day between the ages of one and two years, with this increasing to over ten new words per day by eight years of age (Bloom & Markson, 1998; Fenson et al., 1994). This process appears effortless in neurotypical development; however, it is deceptively complex. Word learning is a multi-stage process and requires children to form a lasting relationship between the phonological word and its meaning. To learn a new word, children must 1. Segment the novel word from a speech stream, 2. Identify the meaning of the novel word (*referent selection*), and 3. Store the word-meaning mapping in memory for later retrieval (*retention*; Gleitman, 1990). This thesis targets the second and third stages of the word learning process (referent selection and retention).

1.2.1. Referent Selection

Determining the meanings of novel words is one of the biggest challenges for young language learners. Quine's (1960) seminal research originally highlighted the complexity of novel word disambiguation. Quine likened infants to 'naïve linguists' studying a community who converse in a foreign language. To learn this foreign language, we must draw conclusions about the meaning of the spoken words in the moment. For example, if someone from the community shouts 'gavagai' as a rabbit hops past, how do we decipher what 'gavagai' means? It could refer to the rabbit, but it could also refer to properties describing the rabbit or its activities (e.g. hopping, fluffy), amongst a host of alternative meanings. How then, does the naïve linguist overcome this referential uncertainty and determine the correct meaning of this word?

Word learning requires children to understand that during a naming event, numerous potential novel words can refer to a plethora of potential objects, actions, or categories of objects (Monaghan & Mattock, 2012). Thus, to learn a word, children must overcome referential ambiguity, narrowing their attention to a single target and label, correctly identifying the novel word-referent mapping (Markman & Wachtel, 1988). Research demonstrates that neurotypical children can perform referent selection from incredibly early on in development, identifying and encoding novel word-referent mappings from as young as six months of age by utilising statistical learning mechanisms (Bergelson & Swingley, 2012).

To successfully identify a correct word-referent mapping, children's attentional focus to both auditory and visual input must be narrowed through either intrinsic or extrinsic sources of information (Hollich et al., 2000). This process is facilitated by children applying heuristics and lexical assumptions, utilising many internal and external sources of information to drive their learning (e.g. Baldwin 1991, 1993; Markman, 1989; Yu & Smith, 2007). These information sources support a process of 'fast mapping' which refers to rapid 'in the moment' mapping of novel words to novel referents (Carey and Bartlett, 1978). Markman (1994) initially posited three constraints that the child places upon potential word meanings: the taxonomic assumption, the whole object assumption, and the mutual exclusivity assumption.

The taxonomic assumption suggests that words should be extended to objects of the same taxonomic category, such as vehicles, foods, or animals, rather than objects that are thematically related (Markman & Hutchinson, 1984). Children therefore quickly learn that 'pig' not only refers to a pig on a farm, but also to Peppa Pig on the television, rather than thematically related aspects like a curly tail (Waxman & Kosowski, 1990). This assumption allows children to quickly expand their knowledge in order to make sense of the dynamic world around them. More recently, children have been found to apply a 'shape bias' alongside this principle. The shape bias assumes that a novel word applies to other items of the same shape, rather than grouping items based on texture, size, or colour (Landau et al., 1988; Smith, 2000). This bias emerges at approximately two years of age in neurotypical children, or after children have acquired at least fifty count nouns in their productive vocabularies (Jones et al., 1991; Samuelson & Smith, 1999, 2000; Tek et al., 2008).

The whole object assumption posits that children assume each novel label must refer to one whole object, rather than its constituent parts or properties (Macnamara, 1972; Markman & Wachtel, 1988). For example, the word 'car' refers to the car as a whole, rather than the wheel or door of the car. This principle also emerges at around two years in neurotypical toddlers (Samuelson & Smith, 1999; Smith et al., 2002), and guides children to generalise novel words to other contexts and exemplars of the same category, allowing their language development to advance rapidly (Gelman, 2003; Kalashnikova et al., 2014; Markman, 1991).

1.2.1.1. Mutual Exclusivity. The mutual exclusivity principle (ME; also known as disjunctive syllogism) is a lexical heuristic that informs correct word-referent mappings. To utilise ME, children assume that only one word relates to each referent (a one-to-one mapping), thus knowing the name of one object drives the mapping of an alternative label to another unfamiliar object (MacWhinney, 1991; Monaghan & Mattock, 2012). This assumption is based on a process of elimination and is driven by prior knowledge. Crucially, children can successfully decipher the correct novel referent from an array of known competitors by narrowing down the one unfamiliar object, and correctly creating a novel word-referent mapping. For example, if a child has a pre-existing label for all but one animal in an array of toy animals, they can assume that on first hearing the novel label 'Okapi', this must refer to the label-less animal (rather than animals with known labels 'cat' and 'dog', for example). Experimentally, this principle is commonly tested by presenting an unfamiliar object amongst familiar objects and asking children to identify the referent of a novel word (e.g. Dautriche et al., 2015). Neurotypical children have been shown to successfully apply the ME principle and spontaneously create novel word-object associations by two years of age (Carey & Bartlett, 1978; Markman & Wachtel, 1998).

Fast mapping was initially coined by Carey and Bartlett (1978), and relates to children's ability to immediately identify a novel word, using the linguistic and non-linguistic context to acquire information about its meaning and apply the principle of ME. In Carey and Bartlett's (1978) study, children were asked to retrieve a tray "the chromium one, not the red one, the chromium one" (p. 18). Children used their prior knowledge of the colour red to infer that the teacher was requesting the olive-green tray, determining in that moment that chromium must refer to olive green. This was deciphered to be 'fast mapping.' Together, the lexical heuristics operate in tandem to allow children to acquire information from their environment, both immediately and over time, generating robust novel word-referent mappings which can be embedded within the lexicon.

1.2.1.2. Social Cues. Children's developing social cognitive skills also help them to acquire language. By the age of 9 to 12 months, infants develop gaze following and social referencing skills (Brooks & Meltzoff, 2002; Phillips et al., 2002; Woodward, 2003). The social pragmatic account of word learning suggests that children decipher word meanings by inferring the intentions and referential focus of others (Baldwin et al., 1996; Bloom 1997). Children's social environments contain a variety of social cues from other people, including eye-gaze and gestures (e.g. Bannard & Tomasello, 2012; Booth et al., 2008; Tomasello, 2000). Studies experimentally investigating the influence of social cues on language development demonstrate that infants as young as 18-months-old follow an adult's gaze to infer correct novel word-referent mappings (Baldwin, 1993). Moreover, children's understanding of social intention to inform their word learning can be even more sophisticated, with the ability to infer correct word-referent associations even when the novel word was uttered prior to the introduction of the object (Tomasello, 2000). Children can follow multiple social cues (e.g. gaze and point concurrently) but research establishes that novel word learning is optimal when a single informative cue is utilised (Booth et al., 2008). Clearly, in the correct context social cues can be successfully utilised by neurotypical children from an early age to inform their word learning.

1.2.1.3. Statistical Learning Mechanisms. Although lexical constraints guide learning, sometimes accurate fast mapping is impossible due to the number of potential referents or lack of cues present in the environment. In this case, statistical cross-situational learning can occur (Gleitman, 1990; Pinker, 1989). This method of learning suggests that on hearing a novel word, a child can hypothesise a set of potential meanings for that word from the non-linguistic context. Initially, multiple ambiguous naming events occur (Yurovsky & Frank, 2015). However, over time the child will continue to hear the word in several different contexts. They will therefore identify which of the potential meanings remain consistent across multiple cooccurrences, rejecting previously incorrect pairings (Yu & Ballard, 2007). For example, on hearing an unfamiliar word ‘bird’ whilst viewing two novel animals, a bird and a horse, children have no way of deciphering the correct referent without external cues. However, on hearing ‘bird’ again in a different context, this time in the presence of a dog, they can use cross-situational occurrences to accurately identify what, in fact, ‘bird’ refers to (Yurovsky et al., 2013a, 2013b). Neurotypical children have been demonstrated to successfully utilise cross-situational statistics to inform their word learning from as young as one year of age (Smith & Yu, 2008). Cross-situational learning is thought to be particularly successful, as when children learn word-object associations incrementally, the connections become strengthened due to numerous exposures to the mapping (McMurray et al., 2012; Plunkett, 1997).

1.2.2. Retention

Whilst identification of novel word meanings through referent selection is a crucial step towards word learning, these novel word mappings need to be integrated into the lexicon. In a pivotal study, Horst and Samuelson (2008) identified the requirement to study word learning as a multi-stage process involving both referent selection and retention. They highlighted that although many studies demonstrated children's abilities to map a novel label to a novel stimulus during referent selection (e.g. Mervis & Bertrand, 1994; Halberda, 2003; Wilkinson et al., 2003), research investigating children's ability to retain these novel word-referent mappings was limited. Indeed, since retention represents 'true' word learning, investigating this process is vital. As such, Horst and Samuelson (2008) taught 24-month-olds novel names via a fast-mapping task. When toddlers were required to identify novel targets after merely 5 minutes, they failed to recall this information with above-chance accuracy. However, when the salience of the word-referent mapping was enhanced following referent selection through ostensive naming accompanied by a pointing cue, children's retention increased to above chance levels. This suggests that accurate referent selection alone is not enough to facilitate novel word retention, and other factors alongside accurate fast mapping are required for successful retention (Horst & Samuelson, 2008).

The dynamic associative account is a theory of word learning which aims to explain how referent selection and retention inter-relate. The theory posits that word learning comprises separate 'fast mapping' and 'slow associative learning' processes that operate on different timescales (Kucker et al., 2015; McMurray et al., 2012; Samuelson & McMurray, 2017). Whilst fast mapping refers to the ability to identify a referent in the moment, associative learning involves acquiring knowledge through repeated occurrences of novel words and referents over multiple learning instances (McMurray et al., 2012; Smith et al., 2011). Associative learning means that the child can apply the new word-object mapping across multiple contexts, and thus slow learning is thought to afford more accurate retention than a single disambiguation event like fast mapping (Axelsson et al., 2012). The dynamic associative theory posits that accurate fast mapping alone is not sufficient to afford successful retention. In accord with empirical evidence (e.g. Horst & Samuelson, 2008; Horst et al., 2010, 2011; Kucker & Samuelson, 2012; Kucker et al., 2016), this framework proposes that multiple factors (e.g. prior knowledge and external cues) must interact to afford successful retention (McMurray et al., 2012). As such, research continues to investigate under what conditions novel word retention can be optimised.

Crucially, effective retention is constrained by external inputs and cues, which support the integration of novel words into the lexicon (Arunachalam & Luyster, 2016; Dumay & Gaskell, 2007). Kucker and Samuelson (2012) taught 24-month-olds novel labels in a paradigm similar to Horst and Samuelson (2008). However, prior to referent selection, in one condition children had two minutes to explore the novel objects that would be presented during referent selection and retention. In the other condition, children were pre-familiarised with the novel words (but not objects). Their results indicated that this brief object pre-familiarisation period was enough to boost retention of novel words learnt during referent selection. An impressive 73% of the 24-month-olds who experienced the object pre-familiarisation demonstrated accurate retention. However, children who experienced word pre-familiarisation did not demonstrate retention at levels significantly above chance. The researchers also investigated the impact of novel object pre-familiarisation with 18-month-old children and determined that their retention performance did not increase significantly. They concluded that increasing 24-month-old children's represented knowledge of objects is sufficient to create a significant change in novel word retention. This is because Kucker and Samuelson's (2012) study targets both fast mapping processes at referent selection, and slow-mapping processes due to repeated exposure (including initial familiarisation). Thus, the relationship between these mechanisms affords more robust learning.

Moreover, Kucker et al. (2020) determined that the strength of prior knowledge of familiar items at referent selection influences retention of new word-object mappings. Specifically, they used referent selection to teach novel words to 18- and 24-month-old children in the presence of competitor stimuli that parents identified as being either ‘well known’ (e.g. shoe, dog) or ‘weakly known’ (e.g. whisk, slinky). Perhaps surprisingly, they found that both 18- and 24-month-olds demonstrated better retention of the mappings taught in the context of weakly known foils, although this effect was greater for the 18-month-olds (83% correct) compared to the 24-month-olds (52% correct). Clearly, children could map novel words even when knowledge was not robust. The authors concluded that challenging referent selection contexts may result in deeper processing at encoding and thus better long-term outcomes for novel word retention (McMurray et al., 2012; Vlach & Sandhofer, 2014).

Further, some studies demonstrate that knowledge can attune learners to similarities between novel and known items, and therefore increase learning when these features are shared (Smith et al., 2002). For example, Twomey et al. (2014) found that in comparison to consistently presenting a novel object of the same colour, children retained novel labels only when the novel object varied in colour across exposures. It is suggested here that moderate variability in the novel objects allowed children to recognise the commonalities between objects. Thus, children could compare each new exemplar to their stored memory representation for that object category, drawing upon these commonalities to aid retrieval of key category-defining information after a short delay. Twomey et al. (2018) then examined whether this effect remains when the background colour of the learning environment differs, but stimuli properties stay the same. Two-year-olds learnt names for novel objects presented on either a uniform white background, or on backgrounds that differed in colour. Overall, only children who learnt in the presence of varying background colours retained novel word-object pairings. Potentially, this variation in background context allowed children to learn more about stimuli than when a uniform background was present. In line with adult research (e.g. Smith & Handy, 2014), it is possible that when children learn with a uniform background, they learn about the whole context and consequently encode more information than required. However, when learning in a varied background, children decipher that they only need to encode information about the object, not the background, encoding more information about the novel object for successful retrieval. Overall, these studies reveal the importance of stimuli for novel word retention.

1.2.2.1. Sleep and Retention. A plethora of studies have investigated the effect of sleep on neurotypical children's word retention both after napping and overnight sleep (e.g. Brown et al., 2012; Friedrich et al., 2015). Axelsson et al. (2018) demonstrated that sleep can enhance declarative memories for novel words in children. They taught 2-year-olds novel words via referent selection, and tested their retention immediately after referent selection, 4-hours later, and the following morning. Half of the toddlers napped prior to the 4-hour test, and half remained awake. They discovered that for the children who napped, retention scores remained steady at each of the three retention stages. However, for the wake group, retention scores declined, and were significantly lower than those of the nap group. Nap duration also predicted better retention scores within the children who napped. Overall, these findings suggest that sleep protects novel word-object associations from decay.

Williams and Horst (2014) also investigated the effects of sleep on word learning in 3-year-old children using a shared storybook reading task. Novel word retention following storybook reading was tested after delays of 2.5 hours, 24 hours, and 7 days. Half of the participants napped after storybook reading, prior to the 2.5-hour retention test, and the other half remained awake prior to the retention test. They found that napping was beneficial for novel word retention, with effects persisting across the different retention time points. In a similar study, Horváth et al. (2015) taught novel words to neurotypical infants and retested them after a 2-hour delay following either a nap or no nap. Infants who napped demonstrated a greater increase in word learning performance compared to those who remained awake, demonstrating that sleep, even in the form of a nap, helps to consolidate word learning in infancy (see also Horváth et al., 2016; Horváth & Plunkett, 2018 for similar results). Following this, Horváth and Plunkett (2016) studied the role of sleep on neurotypical children's general language development. They revealed a longitudinal relationship between napping and language development – in particular, children who napped more had greater receptive and expressive vocabularies. Taken together, these studies indicate the importance of napping on young children's vocabulary development both within and outside of an experimental context.

Horváth et al. (2016) suggested that greater overnight sleep efficiency, including less fragmented sleep, was associated with larger expressive and receptive vocabulary scores. Empirical research investigating the role of overnight sleep on word learning also indicate a facilitative role of sleep on novel word consolidation. For example, Henderson and colleagues have consistently shown that school-aged neurotypical children's novel word retention improves after a night's sleep (Henderson et al., 2012; see also Brown et al., 2012). In their 2013 study, Henderson et al. establish the robustness of sleep consolidation, with consolidatory effects persisting after one week. They discovered that when children learn semantic knowledge alongside novel words, there is a particularly strong effect of sleep consolidation (Henderson et al., 2013). Overall, they highlight that a period of off-line consolidation is required for children to strengthen new vocabulary knowledge and integrate it within the lexicon, with the presence of semantic information, alongside the consolidatory effects of sleep, being particularly beneficial to develop stable long-term lexical representations for easy retrieval (Henderson et al., 2013). These studies demonstrate a robust positive effect of sleeping, both napping and overnight sleep, on novel word consolidation.

Moreover, Henderson et al. (2012) investigated sleep-associated consolidation effects in the declarative memories of neurotypical 7- to 12-year-olds. Children were exposed to non-word lexical competitors (that sounded like words in their pre-existing lexicons), but their recognition and recall of these words only improved after sleep. This improvement in performance remained stable after one week. Henderson and colleagues explain their findings utilising the dual memory system approach. This approach suggests that novel words must firstly be integrated into the lexicon to reshape existing neocortical networks in order for consolidation to occur. Then, competition during word recognition between novel words and similar sounding well-established lexical neighbours occurs. This acts to protect the developing lexicon from overwriting existing information, and ultimately affords better novel word consolidation.

How sleep benefits word learning can be explained by ‘active system consolidation theory’ (Diekelmann & Born, 2010). This theory posits that sleep reactivates recently encoded declarative memories, such as word-referent representations. Whilst initial novel word representations are fragile, synaptic connections are pruned and strengthened during sleep, implicating them in hippocampal-neocortical consolidation (Diekelmann & Born, 2010; Gais & Born, 2004). This means that novel word-referent associations are reactivated and integrated into memory networks, enabling longer-term retention (Drosopoulos et al., 2007; Gais & Born, 2004). This theory would therefore suggest that memory representations for novel words are less likely to decay following sleep as they have been consolidated into the lexicon through these refining processes.

1.3. Attention and Word Learning

Word learning is moderated by attention, a network of systems that work together across different brain regions to maintain alertness and sustain and shift focus (Posner, 1988; Posner & Dehaene, 1994). Word learning involves paying attention to visual and auditory stimuli, identifying competitor and target objects, and encoding information about targets to inform novel word-referent mappings (Twomey et al., 2016; Zosh et al., 2013). Consequently, the strength of associations and likelihood of retention is affected by attention (Samuelson & Smith, 2000). A growing body of word learning research has investigated how word learning is mediated by attention. Studies have examined the extent to which attention is important for word learning by intentionally manipulating children's attention towards, or away from, target stimuli. Initially, Horst and Samuelson (2008) report poor retention across their experiments (as described in section 1.2.2). However, they discovered that when the saliency of the novel object and novel label was increased immediately following referent selection through ostensive naming accompanied by a pointing gesture, retention improved significantly to above-chance levels. Here, increased attention to the novel object as it is labelled determined the likelihood that a novel word-referent association is retained. Imperatively, this study also demonstrates that neurotypical children were able to utilise external social cues to identify novel words and improve their learning.

Further, Axelsson et al. (2012) sought to increase neurotypical 2-year-old children's attention during word learning to improve novel word retention. Here, they either illuminated the target stimuli, covered the competitor stimuli, or both. These conditions were compared to a control group who were given a social cue (pointing). Overall, children successfully retained novel word-referent mappings if the target object was illuminated, and when the target was illuminated, and competitors simultaneously dampened. By comparison, children who received the social pragmatic cue did not demonstrate retention. This contrasts with

Horst and Samuelson’s (2008) findings, who found that a similar social pragmatic cue enhanced retention in 2-year-olds. The potential disparity here could be due to a slight difference in methodology – Horst and Samuelson (2008) lifted the object away from competitors whilst pointing to it, whereas Axelsson et al. (2012) left the object in place. This act of changing the spatial location in Horst and Samuelson (2008) may have drawn children’s attention towards the target, and away from competitors, akin to the illumination condition in Axelsson et al. (2012) which yielded comparable findings. Axelsson and colleagues concluded that manipulating children’s attention to increase focus on a target aids novel word retention. Directing children’s attention away from competitors and *simultaneously* directing attention toward a target also assisted novel word retention. Given that ME requires children to exclude competitor objects and attend to the target object to encode the correct novel word-referent mapping, the mechanisms used in this study help children to apply these key components of ME. These attentional manipulations reduce the cognitive demands required to utilise ME, thus enabling children to allocate their cognitive resources to encoding stronger novel word-referent associations.

A study by Bion et al. (2013) examined how children’s visual attention influences novel word retention. They demonstrated that between the ages of 18 and 30 months, children’s abilities to disambiguate and retain novel word meanings increased. Interestingly, children who looked relatively more at the novel object after hearing the novel label achieved more accurate retention (defined as proportionally greater looking towards the target at retention). They suggested that neurotypical toddlers’ greater looking towards a novel object afforded more robust encoding of information about the word-referent mapping and thus more successful retention.

1.3.1. Manipulating Attention Through Target Stimuli

Additionally, the actual properties of to-be-learned stimuli influence word-learning. Brandone et al. (2007) manipulated stimuli saliency by altering the outcome of an action during children's learning. Two- and three-year-old children were shown two actions which were accompanied by a label. In one condition, the labelled action led to no result, and in another condition, the labelled action caused a light to illuminate. Their findings revealed that 2-year-olds learned a novel verb when the experimenter labelled the action that had a salient result, but not when the action produced no result. However, 3-year-olds performed well in both conditions. Overall, this study suggests that perceptual salience is a valuable cue that is particularly important to aid early word learning.

To investigate salience of stimuli, Hollich et al. (2000) studied word learning in neurotypical 12-, 19-, and 24-month-olds, examining how perceptual and social cues interact. In their studies, children were required to follow social eye gaze to learn novel labels via referent selection. For 19- and 24-month-olds, attention to social information was sufficient to afford word learning irrespective of whether the speaker labelled the object that was more, or less, perceptually salient. However, 12-month-olds only succeeded when both social and perceptual cues aligned, but not when the speaker labelled the boring object. Clearly, perceptual salience is crucial for word acquisition, particularly to younger learners. These findings were corroborated by Pruden et al. (2006), who suggested that children as young as ten months rely on the perceptual salience of an object in the service of word learning. They discovered that these infants learnt more novel labels associated with brightly coloured objects with noises/moving parts, than dull objects or objects highlighted by a speaker's social cues. Overall, these studies consistently present evidence for children's sensitivity to object salience, and the positive impact that increased salience has on word learning due to increased focus on stimuli.

Additionally, children’s own interests can influence their attention allocation and acquisition of novel information (Begus & Southgate, 2012). Neurotypical children have been shown to more easily retain words that refer to individual items, or members of specific categories of interest, presenting evidence that increased attention to stimuli aids word learning. For example, Ackermann et al. (2020) investigated how neurotypical children’s stimulus and category interests impacts word learning. They used both questionnaires and pupillometry to measure the extent to which 2-year-olds preferred specific stimuli and categories of stimuli, whereby increased pupillary change following exposure to stimuli was interpreted as increased interest. Children then learnt four novel word-referent pairings via ME-based referent selection, one from each category (animals, clothes, vehicles, drinks). Following training on the first two novel words, word recognition was tested by presenting the two novel images side by side, and children were directed to look at one of the images. This was repeated for the second block of the remaining two words. Children’s proportionally longer looking was inferred as their response selection. They found that children showed more robust learning of word–object associations from categories they were more interested in (e.g. animals), compared to categories they were less interested in (e.g. clothes). Interest in specific novel stimuli also moderated word learning. Overall, these findings suggest that neurotypical children's preferences for particular categories and category members shapes their word learning.

However, Kidd and colleagues argue that children’s interests may be subject to a ‘Goldilocks effect’ (Kidd et al., 2012, 2014). They highlight that infants cleverly allocate their auditory and visual attention to stimuli that are neither too simple nor too complex to avoid wasting cognitive resources. Children may consider their knowledge of simple stimuli to be saturated and therefore believe they do not need to attend to such stimuli. Alternatively, stimuli that are too complex may be beyond children’s remit of learning and thus they do not attempt to intake information from them. This idea originates early in developmental psychology. Piaget (1970) suggested that children attempt to integrate new knowledge into existing knowledge structures by a process of assimilation. When this is not possible, children create new knowledge structures - a process called accommodation. On the one hand, stimuli must be interesting and relevant enough to be applied to existing structures. On the other hand, knowledge acquisition of complex stimuli must be achievable to warrant creating new structures. Piaget's theory is supported by empirical evidence that confirms children prefer to look at familiar stimuli that still require some encoding into memory, or novel stimuli that can be readily encoded into memory (e.g. Hunter & Ames, 1988; Roder et al., 2000).

Ackermann et al. (2020) highlight that it is possible that category interests may too be subject to the 'Goldilocks effect.' Interest in certain categories may only be acquired once category knowledge reaches a certain level, resulting in better learning of new category members and thus an increase in category size. Potentially, if the child is familiar with many members of a category, they may not be interested in this category because they feel their knowledge is saturated. Similarly, if they have prior knowledge of very few members of a category, they may have no existing category exemplars to link the new members to, and therefore feel learning about this category is impossible. However, a category size that is 'just right' might make children most curious – they have existing representations to link the new category members to, but do not feel they have learnt all they need to know about the category, resulting in better learning. Ackermann et al.'s (2020) findings suggest that pre-existing knowledge influences children's learning both in terms of linguistic and category knowledge (Oudeyer & Smith, 2016).

Naturally, children demonstrate individual variation regarding stimuli that they prefer to learn about. Stimuli that are salient and interesting to one child may not be quite so striking to another. As such, self-directed learning has been suggested as a successful form of active, curiosity-based learning whereby children can choose what to learn about (Gureckis & Markant, 2012; Montessori, 1912; Partridge et al., 2015). Children display self-directed learning in several ways, including self-questioning, which affords independent control and direction of learning (Gavelek & Raphael, 1985; Palincsar & Brown, 1987). It is thought that self-directed learning is successful because, as the ‘Goldilocks effect’ suggests, children can cleverly identify best-suited materials to optimise learning (Kidd et al., 2012, 2014). Multiple studies have determined self-directed learning to be more effective at helping children to learn across many domains (Yang & Li, 2013). For instance, Begus et al. (2014) allowed 16-month-olds to choose one of four objects to learn about. The experimenter modelled an action for either children’s chosen, or non-chosen object. After a 10-minute delay, infants replicated significantly more of the functions of chosen objects than those of non-chosen objects. This confirms that like adults (e.g. Markant et al., 2014), self-directed learning is beneficial to children as young as 16 months of age.

The advantage of self-directed learning has also been extended to word learning. A study by Partridge et al. (2015) investigated self-directed touchscreen-based learning in neurotypical 3- to 5-year-olds. In one condition, learning was self-directed and thus children learnt about stimuli they selected by pressing on the novel toy characters. In the other condition, children learnt about novel toy characters that were randomly labelled (children did not choose specific characters). Children learnt the four novel labels via ostensive naming, with the character expanding in size during labelling. Labels were taught in the presence of zero, one, two, or three competitors. Following training, fifteen toys were presented on the screen at once and children were asked to identify the correct novel label referent. Overall, they found that children retained significantly more novel word-object pairings in the choice condition compared to the no-choice condition. This suggests that children's active engagement in self-directed learning can enhance short-term retention for novel word-object mappings.

1.3.2. Manipulating Attention Through Distractor Stimuli

In addition to the range of studies demonstrating the influence of attention towards target stimuli on word learning, it is important to consider that learning contexts often present an array of competing stimuli. Indeed, real world learning environments are rife with extraneous information and require children to exclude myriad distractors to learn about the correct intended referent. However, most word learning studies do not consider how children may divide their attention across the whole visual scene. Importantly, to use ME, children must select the correct to-be-learned referent from an array of familiar distractor stimuli. As such, stimuli within an array could potentially affect children's learning success. If to-be-learned stimuli appeal to children's interests, they may allocate greater attentional resources to these stimuli, more easily encoding novel information about the targets (Ackermann et al., 2020). However, if competitor stimuli capture children's attention, focus may be drawn away from target stimuli, resulting in the formation of weak or incorrect representations (Venker et al., 2018).

To investigate the effect of distractor stimuli saliency in neurotypical development, Pomper and Saffran (2018) paired novel objects with familiar objects that varied in their visual salience. In this study, high salience items were brightly coloured and engaging (e.g. animals, vehicles, foods), whereas low salience items were dull in colour and less engaging (household objects). They discovered that children were slower and less accurate to identify novel objects in the presence of highly salient familiar objects than in the presence of less salient familiar objects. Moreover, children retained fewer novel word-object mappings when learning occurred in the presence of highly salient familiar objects. This study highlights the influence of the whole learning context, including the properties of the distractor stimuli.

Highly salient distractors have the potential to increase attentional division between distractor and target stimuli, thus affecting strength of encoding novel word-referent associations.

Additionally, Horst et al. (2010) investigated how quantity of distractors influences learning. They presented children with referent selection trials which had varying numbers of competitor stimuli presented alongside a single target. Whilst the number of competitors did not affect children's accuracy at referent selection, increasing the number of competitors decreased the likelihood that novel words were retained. This suggests that, when attentional division across the visual scene was greater, and less attention was allocated to the novel object, the association between the word and the object became more susceptible to decay.

Axelsson and Horst (2014) built upon this by investigating how either consistently presenting the same familiar competitors alongside a novel object, or varying competitors across exposures, influenced word learning. They found that children retained more novel word-object mappings when the novel object had been presented with consistent familiar competitors. It was suggested here that children could better focus their attention on the novel target in the consistent competitor condition because they did not always need to rule out the competitors as potential referents. By contrast, the varied competitor condition required an additional step of children attending to and ruling out familiar referents on every occasion. Similarly, Horst et al. (2019) demonstrated that 30-month-old children could retain novel word-referent mappings both when taught in the presence of competitors that were the same colour as the target or varied in colour. However, they could only generalise the novel names to new exemplars of the target categories when the objects were taught in the presence of same-coloured targets, not if the targets and competitors varied in colour. Overall, these data highlight that children retain more novel words when learning through referent selection if distractor stimuli are consistent and require minimal attentional demands. This means that greater attention can be allocated to novel stimuli which affords more successful novel word retention and generalisation.

1.4. Word Learning in Autism

Many autistic children experience delayed receptive vocabulary development, which may be due to difficulties learning words (e.g. Tek et al., 2008). Explanations for why autistic children experience atypical language development are varied. Studies often attribute autistic children's word learning difficulties to obstacles interpreting social pragmatic cues to inform word-referent mappings (e.g. Baron-Cohen et al., 1997; McDuffie et al., 2006a; Parish-Morris et al., 2007). However, social pragmatics are not the only way in which children acquire novel words – children also utilise non-social cues and cognitive constraints to inform their learning (Preissler & Carey, 2005). Research has therefore explored whether autistic children's word learning mechanisms are qualitatively different to those of neurotypical children (e.g. Carter & Hartley, 2021; Hartley et al., 2019, 2020; Venker, 2019).

1.4.1. Referent Selection in Autism

In neurotypical word learning, children use certain constraints to inform word learning (Markman, 1990). It has long been debated whether these constraints are innate or learned in neurotypical development (e.g. Hollich et al., 2000). Autistic children also appear to possess word learning biases that are present in typical development and are used to inform referent selection and thus subsequent retention (e.g. de Marchena et al., 2011; Hartley et al., 2019, 2020). A growing body of literature aims to identify under what conditions autistic children can successfully utilise lexical heuristics and external cues in the service of word learning (McGregor et al., 2013; Parish-Morris et al., 2007; Preissler & Carey, 2005; Tenenbaum et al., 2014, 2017).

1.4.1.1. Mutual Exclusivity in Autism. A plethora of studies state that autistic children can utilise the ME lexical heuristic to inform accurate referent selection (de

Marchena et al., 2011; Preissler & Carey, 2005). Initially, Preissler and Carey (2005) demonstrated that language delayed minimally verbal autistic children can use ME to decipher correct novel referents with comparable accuracy to neurotypical 24-month-olds. This finding has been corroborated by more recent research with lexically able autistic children. For example, de Marchena et al. (2011) presented an ME-based referent selection task to neurotypical and autistic children matched on age and vocabulary abilities. Children either learnt a label for a novel object or a fact via ME. Overall, children in both groups could accurately utilise ME to learn both words and facts, although this tendency was stronger for words than facts across populations. These studies indicate that autistic children's learning mechanisms are not qualitatively atypical.

Mathée-Scott and colleagues (2021) investigated the use of ME in autistic and neurotypical toddlers matched on nonverbal cognition. In their study, autistic children varied in vocabulary abilities to examine how vocabulary size influenced children's use of the ME heuristic. They found that found that autistic children's use of ME was dependent on vocabulary size. Of course, given that ME requires use of pre-existing knowledge to determine the correct novel referent, this finding is unsurprising (de Marchena et al., 2011; Mathée-Scott et al., 2021). However, this highlights that initial use of ME may not always be an innate heuristic, but a skill that is refined as vocabulary acquisition advances.

In another recent study, Carter and Hartley (2021) explored whether neurotypical and language delayed autistic children matched on receptive vocabulary abilities were more likely to retain words when learning from colour photographs than black and white cartoons. Here, they taught neurotypical and autistic children novel words through ME-based referent selection. Following a 5-minute delay, autistic children achieved significantly greater retention accuracy when learning from photographs than cartoons. Autistic children also retained novel labels more accurately when learning from photographs than neurotypical children. Overall, these findings highlight that autistic children's abilities to use ME in the service of novel word acquisition are commensurate to neurotypical peers when expectations are based on receptive language abilities.

Hartley et al. (2019) demonstrated that autistic children were less accurate than neurotypical peers to identify words via ME, but this did not negatively influence their novel word retention. Importantly, they too discovered that greater ME accuracy predicted larger vocabulary sizes for both autistic and neurotypical children (de Marchena et al., 2011; Mathée-Scott et al., 2021). This suggests that ME is an important strategy for language acquisition in both populations. Given autistic children's lower referent selection accuracy and difficulties learning words overall, ME may be a particularly important strategy for this population. Crucially, where the use of ME is impossible, word learning may likely be impaired due to inability to use this important heuristic.

1.4.1.2. Use of Social Cues in Autism. Attending to social cues to decipher novel referents can be difficult for children who have poorer social cues, such as shy children (Hilton et al., 2023). Hence**,** language difficulties in autism are often attributed to problems using extrinsic social cues. A concurrent relationship between attention following and vocabulary comprehension has been well-documented for both neurotypical and autistic development (e.g. Carpenter et al., 1998; Charman et al., 2003; Landry & Loveland, 1988). Generally, social behaviours help children attend to and engage with the environment. Many of the social cues that promote language development involving drawing attention to specific stimuli externally, such as attention-following (McDuffie et al., 2006a; Rowland, 2014). Evidence for how social cues such as gaze, gesture, posture, and facial expressions affect word learning in autistic children is mixed (e.g. Baldwin & Moses, 2001; Tomasello, 2003; Watt et al., 2006).

Whilst neurotypical children can use social cues to disambiguate the meanings of novel words (Baldwin, 1991; Bloom, 2002b), some research suggests that autistic children may struggle to use these cues, and therefore have impairments in understanding and inferring word meanings (Baron-Cohen et al., 1997). To utilise social cues in informing lexical acquisition, autistic children must understand the referential intent of the cue. In Baron-Cohen et al.'s (1997) seminal study, language delayed autistic children were required to map a novel word to a novel object that was either the focus of the speaker, or the focus of the child (the child held the object in this condition). They demonstrated that autistic children were insensitive to the speaker's direction of gaze as a cue to assign word meaning, and instead mapped the novel word to an object of their own focus due to their own preferences. Similar findings were reported by Preissler and Carey (2005) – minimally verbal autistic children were more likely to map a novel word to an item they were attending to, irrespective of the experimenter's gaze. If autistic children do focus on an incorrect referent due to difficulties utilising social cues, they may experience an increased number of ambiguous learning events whereby correct novel word-referent mappings are either not encoded or are incorrectly identified and encoded (Baron-Cohen et al., 1997; Tenenbaum et al., 2017; Venker et al., 2018).

However, autistic children have been shown to successfully utilise social cues to learn words in some contexts. For example, Luyster and Lord (2009) utilised a similar experimental design to Baron-Cohen et al. (1997) and Preissler and Carey (2005). Luyster and Lord's sample of autistic children were much younger, but more lexically able, than the groups of participants in these earlier studies. They found that autistic children did not differ in their performance from expressive vocabulary matched neurotypical children. This suggests that autistic children with more advanced vocabularies can use social cues to acquire novel words. Since social skills have been posited as an important precursor to language development (Brooks & Meltzoff, 2015), this finding is not unsurprising. It may be that autistic children with greater social skills develop language more easily. If so, social proficiency may predict successful language development (Luyster et al., 2008; Thurm et al., 2007). If autistic children can use social cues to direct their attention to the correct to-be-learned stimuli, their chances of successful word learning are improved. However, given the differences in findings across similar studies, further research is required to disentangle the relationship between lexicon size and the ability to utilise social pragmatics in autistic children.

Additionally, Adamson et al. (2009) demonstrated that language delayed autistic children could understand social cues in some word learning contexts. They investigated supported joint attention, whereby the child and adult attend simultaneously to the same shared object and label. Specifically, they found that supported joint attention is unimpaired in autism. Furthermore, supported joint attention has been shown to contribute to receptive and expressive language outcomes in autism, and can be utilised to scaffold novel word acquisition (Adamson et al., 2004, 2009; Mundy et al., 1986). This highlights that the understanding of some social cues is unimpaired in autism, thus use of such cues can benefit language development.

A study from Parish-Morris et al. (2007) investigated the use of social cues to inform novel word learning in language delayed autistic children in comparison to language and mental age matched neurotypical children. They found that autistic children were able to learn the meaning of a novel noun by attending to a speaker’s perceptually salient social cues, like eye gaze and pointing, when target objects were also visually salient (Parish-Morris et al., 2007). However, autistic children’s word learning was diminished when target objects were not interesting to them, despite the presence of social cues. Autistic children also struggled to utilise speaker intent, a more subtle social cue which has been described as beneficial, but not essential, to word learning in neurotypical development (Hollich et al., 2000). Nevertheless, Parish-Morris et al. (2007) discovered that understanding of social intention predicted vocabulary size in autistic children, but not neurotypical children (see also McDuffie et al., 2006a). Overall, the findings of most of the studies reviewed here (e.g. Adamson et al., 2009; Luyster & Lord, 2009; Parish-Morris et al., 2007) suggest that autistic children can use salient social cues in a similar manner to neurotypical to support novel word learning. However, proficiency in understanding the intention of these cues may relate to vocabulary abilities (Hani et al., 2013; Norbury et al., 2010). The findings show that social cues are valuable to inform accurate novel word learning. However, it remains uncertain as to how other factors such as stimuli salience and more subtle social nuances may influence autistic and neurotypical children's word learning, and under what context which word learning mechanisms are more beneficial.

1.4.1.3. Statistical Learning Mechanisms in Autism. Research investigating cross-situational word learning in autism is limited. An early account of cross-situational learning in autistic children comes from McGregor et al. (2013), who demonstrated that 11-year-olds without language delay could utilise cross-situational statistics to learn novel words. Here, success at cross-situational learning was moderated by vocabulary skills, suggesting that cross-situational learning is crucial for autistic children's vocabulary development. However, it is important to consider that in McGregor's study, ostensive naming trials were interleaved with referentially ambiguous trials, meaning participants may have already been cued to the meaning of some words before experiencing the referentially ambiguous trials. As such, caution should be taken when drawing conclusions from these findings, due to the difficulty determining whether children mapped words based on cross-situational statistics alone, or utilised ostensive cues.

Moreover, Venker (2019) went on to investigate cross-situational word learning in a younger age group. They compared the word learning abilities of a sample of 4- to 7-year-old autistic children to vocabulary matched neurotypical peers. Children participated in both a cross-situational and ostensive word learning task. Overall, both groups performed cross-situational learning with comparable accuracy. Similarly, Hartley et al. (2020) found that language delayed autistic children could utilise cross-situational statistics to determine novel word meanings as accurately as neurotypical children matched on receptive vocabulary. These studies add to a growing body of literature that autistic children's word learning mechanisms are qualitatively similar to those of neurotypical children (Arunachalam & Luyster, 2016; de Marchena et al., 2011; Luyster & Lord, 2009; McGregor et al., 2013).

1.4.2. Retention in Autism

It is also important to recognise that identification of meaning is just one component of word learning (McMurray et al., 2012). It is currently unclear how effectively autistic children can retain novel words. Whilst some studies suggest that autistic children or children at risk of autism are less able to retain novel word information than their neurotypical peers (Bedford et al., 2013; Norbury et al., 2010), more recent research posits that autistic children can retain novel word meanings as accurately as their neurotypical peers (Carter & Hartley, 2021; Hartley et al., 2019, 2020).

In a series of studies, Hartley et al. (2019, 2020) investigated how language delayed autistic children and neurotypical children matched on receptive vocabulary identify and retain novel words. In Hartley et al.'s (2019) first study, both populations utilised mutual exclusivity to identify referents of unfamiliar words. However, both groups of children showed significantly reduced accuracy on delayed retention and generalisation trials. In the second experiment of Hartley et al. (2019), autistic children were less accurate than neurotypical peers to identify novel words through ME-based referent selection. However, autistic children still demonstrated significantly above-chance responding, demonstrating that they were adept in their use of ME. Autistic children who received social feedback in the form of a head turn and gaze shift towards the target responded more accurately on delayed retention and generalisation trials than neurotypical controls, and more accurately than autistic children who received either non-social feedback or no feedback. Overall, these findings highlight that word learning mechanisms subserving identification and retention of meaning appear to be intact in autism.

In Hartley and colleagues' (2020) study, similar samples learnt novel words through cross-situational learning that was not cued, or that was supported by social or non-social attentional cues. Across these studies, both groups retrieved and generalised word-referent representations with comparable accuracy. However, autistic children were significantly slower to identify correct referents under both cued and non-cued conditions. Overall, the authors indicate that although word learning mechanisms are intact in autism, the increased time required to generate comparable responses suggest that autistic children may be less efficient at processing language input (see also Arunachalam & Luyster, 2016; Venker et al., 2018).

Given that consolidating novel word-referent mappings into long term memory requires coordination of attention and memory processes, it is likely that the diverse attentional allocation present in autism may account for individual differences in novel word retention (Arunachalam & Luyster, 2018; Omaki & Lidz, 2015). Where both autistic and neurotypical children were comparably successful to retain novel words after a delay under experimental conditions in Hartley and colleagues' studies, they highlight that autistic children may not retain novel word information as successfully in more naturalistic environments. This is because these dynamic contexts likely contain extraneous distractions alongside more complex and diverse arrays of stimuli which are presented at a faster pace.

1.4.2.1. Retention and Sleep in Autism.

Research suggests that sleep disorders are one of the most common symptoms in autism, with autistic children being 50-80% more likely to experience sleep issues than neurotypical peers of a similar age (Kotagal et al., 2012; Ming et al., 2008; Souders et al., 2009). Such sleep problems include parasomnias, poor sleep efficiency, irregular sleep-wake patterns, and longer sleep-onset latency (Díaz-Román et al., 2018; Fletcher et al., 2017; Polimeni et al., 2005; Souders et al., 2009). Significant relationships between sleep and memory consolidation have been demonstrated from infancy to adulthood – sleep consistently helps to consolidate declarative memories for longer term retrieval (Dumay & Gaskell, 2007; Kurdziel et al., 2013; Tamminen et al., 2010).

There are also studies demonstrating that sleep influences longer term learning outcomes in neurotypical development. Touchette and colleagues (2007) aimed to investigate whether sleep duration is an independent risk factor for behavioural and cognitive functioning. They found that a reduction of just 1 hour of sleep per night can decrease cognitive performance. In particular, short sleep duration related to lower performance on the Peabody Picture Vocabulary Test (PPVT), a test of receptive vocabulary. This finding proposes that novel word acquisition and vocabulary development could be significantly impeded by chronically shortened sleep duration throughout childhood. Similarly, Seegers et al. (2016) assessed children's sleep duration annually from age 2.5 to 10 years, and receptive vocabulary using the PPVT at four and ten years. They categorised children based on their sleep trajectories – short persistent sleepers, short increasing sleepers, 10-hour sleepers and 11-hour sleepers. They demonstrated that persistent short sleep duration was associated with poorer receptive vocabulary of 10-year-olds in comparison to 11-hour sleepers, highlighting the importance of sleep for lexical development.

Whilst research demonstrates that sleep, both in the form of napping and overnight sleep, benefits novel word consolidation in neurotypical development (e.g. Friedrich et al., 2015; Henderson et al., 2012, 2013; Horváth et al., 2015, 2016; Williams & Horst, 2014), this research area is extremely underdeveloped in autism. Given the prevalence of sleep disorders in autism (Souders et al., 2009), and the importance of sleep to children's long-term lexical consolidation, it is somewhat surprising that the majority of studies investigating the role of sleep on autistic children's language development focus on intellectually able autistic children with age-expected language abilities (e.g. Fletcher et al., 2020; Henderson et al., 2014).

Henderson and colleagues (2014) investigated the influence of sleep on novel word retention in autistic and neurotypical children. Their participants were 8- to 13-year-olds without language delay matched on age and vocabulary knowledge. Children learnt stimulus triplets of lexical neighbours, as used in Henderson et al. (2012), via a listening paradigm in two phonological tasks. Firstly, participants were asked to identify phonemes present in the novel word. Then, they were asked to repeat either the first or the last sound within the novel word. Measures of cued recall and forced choice followed this to assess children's explicit memory of the novel words. Overall, both groups showed similar improvements in novel word retention after a 24-hour delay. However, whilst neurotypical children showed lexical competition effects after 24 hours for phonologically similar words (e.g. biscuit, biscal), autistic children demonstrated these effects immediately. These findings suggest a qualitative difference in how words are integrated into the lexicon in autistic children in comparison to neurotypical development.

Similarly, Fletcher et al. (2020) taught nine rare animal names via ostensive naming to autistic and neurotypical children matched on age, nonverbal ability, and vocabulary. Children were asked to complete numerous learning tasks. Participants were exposed to the novel animal words by repeating an auditory label and reading the written name aloud. Children then completed forced choice trials where participants were presented with two options and asked to select which image matched the written word for image-matching trials, and which word matched the image for orthography-matching trials. Participants also performed a size judgement task, where they deciphered which animal image was larger. In addition, children completed a definitions task, where they were asked to describe the animal to the experimenter. A naming speed task was also undertaken, where children's verbal response pace when naming the animals was recorded. Polysomnography recordings were acquired to assess overnight sleep quality. Overall, the findings demonstrated that autistic and neurotypical children retained novel animal labels with comparable accuracy after overnight sleep, despite autistic children demonstrating differences in sleep properties. However, autistic children forgot significantly more of the unique features of the novel animals. Whilst these studies suggest that autistic children's novel word learning may benefit from overnight sleep in the same way as their neurotypical peers, their autistic participants had developmentally expected language skills. As such, it remains unclear whether these findings would generalise to autistic children with delayed language development (Fletcher et al., 2020; Henderson et al., 2014).

1.5. Attention in Autistic Children's Word Learning

Attention substantially influences word learning processes, as children must correctly identify the intended referent for a novel word and exclude non-target competitors (Leung & Rheingold, 1981). Previous research with neurotypical children has demonstrated the benefit of manipulating children's attention to facilitate word learning processes and language acquisition (e.g. Ackermann et al., 2020; Axelsson et al., 2012; Pomper & Saffran, 2018). Autistic children, however, often display atypical attentional behaviours, which may have downstream consequences for their receptive vocabulary development due to a narrowing of attention as a result of their own preferences (e.g. Baron-Cohen et al., 1997).

1.5.1. Exogenous Attentional Manipulations

As in the neurotypical literature, some studies have investigated how manipulating autistic children's attention through experimental conditions influences word learning. Gliga et al. (2012) studied gaze-following abilities as a prerequisite for word learning in children at familial risk for autism. In their study, examination of looking behaviour demonstrated that gaze following did not directly relate to successful word learning performance. Rather, the distribution of looking between the target and distractor object was a stronger predictor of word learning success than the use of social pragmatic cues. These findings highlight that increased attention towards the target stimuli was important for word learning. Whilst gaze following may not have enhanced children's attention to targets sufficiently, other mechanisms that may increase children's looking distribution could be used to improve word learning. However, it is important to consider that not all of the participants in Gliga et al.'s (2012) study would receive an autism diagnosis.

In Akechi et al. (2011), neurotypical and autistic children matched on age, nonverbal intelligence and verbal mental age were required to map a novel word to an object that was in the speaker's focus. They found that, in contrast to neurotypical children, autistic children were more likely to map the novel word to an object of their own focus rather than the speaker's focus. However, when the target object that was in the speaker's focus was increased in salience by incorporating movement, both groups mapped the novel label to the correct target object. This increased saliency enhanced autistic children's attention towards the novel object and thus afforded more accurate word-referent mappings. Akechi et al. (2013) built on their previous study with the addition of pointing as a social cue, alongside speaker gaze and object movement. They discovered that similar samples of autistic children could successfully use multiple cues, following the point to inform their referential word learning. However, when examining both response accuracy and visual attention, Akechi and colleagues found that these measures were not always congruent across conditions and experiments. For example, some autistic children attended to the speaker's face and followed gaze as frequently as neurotypical children, but autistic children were still more likely to make incorrect selections for the novel word referent.

Tenenbaum et al. (2017) directed children's attention to the intended referent in a word learning task by holding the object close to the speaker's mouth. Although this action decreased word learning for neurotypical children, it facilitated word learning amongst their sample of language delayed autistic children. Holding the object away from the mouth and pointing towards the mouth, however, hindered word learning for autistic children, as attention was divided between two distinct areas of the visual scene. While neurotypical children can utilise social cues such as joint attention or gaze following to facilitate word learning, this research suggests that autistic children struggle to do so (e.g. Luyster & Lord, 2009; Parish-Morris et al., 2007). However, similarly to Akechi and colleagues' findings, autistic children could successfully utilise social cues to direct their attention and thus learn a novel word when accompanied by other cues that draw attention to the target. Tenenbaum et al. (2017) state that word learning can proceed effectively if we can increase autistic children's intake of available information from the environment by limiting distractions (Akechi et al., 2011, 2013; Arunachalam & Luyster, 2018; Parish-Morris et al., 2007). It is important to consider, however, that Tenenbaum utilised gaze measures alone as their dependent variables to infer word learning. Given findings that visual attention alone may not sufficiently influence word learning (e.g. Gliga et al., 2012), it is possible that attentional differences, which may be due to visual preferences, could be concluded as less accurate for learning.

Incorporating explicit forced choice methodology, Hartley et al. (2019) highlighted autistic children's accurate use of attentional cues. In their study, autistic children who received social feedback demonstrated more accurate novel word retention and generalisation than language matched neurotypical children (Hartley et al., 2019). Social feedback also increased autistic children's performance to a greater extent than non-social feedback (a flashing light). Moreover, Hartley et al. (2020) found that autistic children spontaneously utilised social cues to correctly disambiguate novel word meanings, and these cues significantly improved their training accuracy in comparison to when no cues were provided. These studies state that autistic children can use social cues to inform novel word learning as accurately as neurotypical peers when expectations are based on receptive vocabulary abilities.

Like neurotypical children (Pruden et al., 2006), autistic children attend to perceptual salience when learning the names for objects (Parish-Morris et al., 2007). Parish-Morris et al. (2007) investigated the role of perceptual salience, amongst other cues, in neurotypical children in comparison to autistic children with language delays. Both groups of children selected their favourite of two salient objects, which was determined to be the 'interesting object,' compared to boring non-chosen and non-salient objects. When the interesting object was labelled, autistic children reliably demonstrated word learning. However, when boring objects were labelled, unlike neurotypical children, autistic children did not reliably demonstrate word learning even when attempts to redirect their attention were made. These studies highlight potential differences in autistic children's attentional mechanisms and reinforce the importance of appropriate allocation of attention for successful word learning.

Several studies have tried to elucidate how labelling influences visual attention and word learning in autism (Benjamin et al., 2015; McDuffie et al., 2006b; Vivanti et al., 2016). McDuffie et al. (2006b) revealed that whilst autistic children demonstrated less attention following than neurotypical children, verbal labelling facilitated attention in both neurotypical and autistic children. However, more recent studies have found contrasting findings. For example, Vivanti et al. (2016) determined that hearing a label increased visual attention to unfamiliar objects in neurotypical children and children with Williams Syndrome, but not autistic children. Vivanti's findings with autistic children were also corroborated by Benjamin et al. (2015), who found that labelling did not increase attention allocation in either autistic or neurotypical children. Prior studies have shown that increased attention allocation through social cues, such as attention following, are concurrently related to word learning outcomes for autistic and neurotypical children (Baron-Cohen et al., 1997; Charman et al., 2003). However, studies by Benjamin et al. (2015) and Vivanti et al. (2016) suggest that proficiency following more subtle pedagogical cues may also be affected in autism. Thus, if autistic children cannot accurately follow attentional cues such as labelling, word learning will be substantially impacted in naturalistic contexts.

Venker et al. (2021) investigated how visual allocation differs when perceptual salience competes with linguistic information. They compared familiar word recognition of neurotypical and language delayed autistic children matched on receptive language skills. On neutral trials, both the target and distractor images were high in salience, with bright colours and geometric patterns. On competing trials, the distractor image was high in salience, but the target image was low in salience. This elicited competition between bottom-up salience driven processes and top-down language driven processes. Both groups of children showed word recognition but competing perceptual salience significantly decreased attention to the target only in autistic children. These findings indicate that, unlike neurotypical development, attention allocation in autism is more strongly driven by bottom-up processes than top-down processes. Moreover, perceptual properties of objects can disrupt attention to relevant information in autistic children. This has implications for word learning, as decreased attention to relevant visual stimuli may inhibit or delay word learning or result in incorrect mappings.

In a similar study, Venker et al. (2022) explored how autistic children's visual attention influences word learning. They examined how the perceptual salience of objects affected novel referent selection in lexically able autistic children in comparison to age matched neurotypical children. High-salience objects were brightly coloured and visually complex with multiple parts, whereas low-salience objects were less colourful and did not have moving parts. Test trials were divided into low-difficulty and high-difficulty trials. Low-difficulty trials occurred immediately after each teaching sequence and only required children to differentiate a labelled object from an object that was not presented in the previous trials. High-difficulty trials required children to differentiate two previously labelled objects. They discovered that high perceptual salience disrupted novel referent selection in autistic children but facilitated attention to the target object in neurotypical peers. Attention of autistic children was disrupted so significantly by the perceptual salience of the object that they failed the referent selection task on high-difficulty trials. Autistic children were also slower to disengage from high-salience distractor images, indicating that they had stickier attentional mechanisms. Overall, these findings highlight that high perceptual salience can disrupt novel referent selection in autistic children. This has implications for subsequent retention - if children do not attend to target stimuli during learning events, retention will likely be hindered. These results have broader implications and suggest that learning contexts should take into account low-level perceptual features of stimuli, such as visual salience, to maximise attention and thus learning.

Another study from Venker (2019) utilised eye-tracking to examine whether autistic children can use cross-situational statistics to acquire novel words. Autistic and neurotypical children with similar vocabulary knowledge performed comparably well during cross-situational learning, with performance mediated by familiar word processing in both groups. Venker (2019) discovered that children who looked longer at named novel objects also looked longer at named familiar words. This demonstrates that children with more sustained visual attention mechanisms likely had greater skills aligning auditory and visual input. As such, these children were able to consolidate their representations of familiar words and learn new words due to their more advanced word learning abilities (Kucker et al., 2015). On the other hand, Venker (2019) observed that visual inattention detrimentally impacted learning - autistic children who looked away from the images in the cross-situational task also had the weakest language skills. Overall, it can be concluded that attention affords learning opportunities and inattention is therefore a risk factor for delayed language acquisition.

1.5.2. Endogenous Attentional Manipulations

Autistic children also demonstrate intrinsic preferences that drive their attentional allocation and influence their input during learning. Research has shown that, unlike neurotypical children, autistic children and children at risk for autism prefer to attend to objects rather than people and spend more time shifting attention between objects and non-social stimuli (Bhat et al., 2010; Ibanez et al., 2008; Swettenham et al., 1998). This is reflected in RRBIs, a common characteristic in autism which involves intense and perseverative attentional allocation to specific topics of interest (Honey et al., 2012; Kanner, 1943; Richler et al., 2007). RRBIs influence what autistic children attend to and learn about. If children focus on specific, narrowed topics of interest, they may lose out on important information within the environment. For example, if a child is particularly interested in dinosaurs, on hearing a novel word, they may not be able to disengage from a dinosaur toy to follow the speaker's point and thus infer the correct novel referent. This can result in weak or incorrect novel word-referent representations.

Research investigating the visual attention of autistic children and children at risk of autism suggest that these populations are slower and less able to disengage attention from specific stimuli (Elsabbagh et al., 2013; Landry & Bryson, 2004). This inflexibility has the potential to lead to atypical visual processing, meaning that children may not take a top-down approach when examining the visual environment, and may instead focus on a limited subset of stimuli (Elsabbagh et al., 2009). Unsurprisingly, how autistic children process stimuli is often influenced by their interests. Studies reveal that autistic children tend to process high interest stimuli with greater focus and intensity. For example, Elison and colleagues (2012) asked autistic and neurotypical children of varying ages to explore a visual scene which included high and low interest stimuli, as well as social and non-social stimuli. Overall, results indicated that autistic children had an attentional bias to certain non-social stimuli from early to late childhood. They also consistently demonstrated a bias towards high interest objects across the vast age range of 2 to 18 years, with high exploration of these stimuli persisting throughout childhood. These findings corroborate previous evidence that autistic children possess a bias towards non-social and high interest stimuli (Klin et al., 2009; Pierce et al., 2011; Sasson et al., 2011). This atypical attentional distribution in comparison to neurotypical children has the potential to restrict learning towards specific subsets of stimuli. On the one hand, learning may be improved if children show an attentional bias towards stimuli of interest – this could be an advantageous mechanism to scaffold learning in such populations. On the other hand, learning may be disrupted when situated in high interest environments with myriad distractors that decrease children's attention to target stimuli. Currently, limited research investigates how these mechanisms directly impact novel word learning in autistic and neurotypical populations.

Additionally, Sasson and colleagues (2008, 2011) investigated how children's specific interests influence their visual attention. They suggested that autistic children's attention to social stimuli was diminished in the presence of more salient stimuli that appealed to children's interests, such as trains. In Sasson et al. (2008), autistic children's visual attention was described to be more circumscribed, perseverative, and detail oriented across both social and object arrays. Moreover, severity of RRBIs correlated positively with exploration of object pictures and negatively with perseveration on social pictures. In Sasson et al. (2011), autistic children exhibited greater exploration and perseverative attention to objects related to circumscribed interests than their neurotypical peers. Overall, these studies reveal that autistic children may perseverate on images of interest and explore them in a more detail-oriented manner. Consequently, autistic children may be more able to learn words associated with categories that appeal to their interests, in comparison to categories they are less interested in.

Autistic children often focus on specific perceptual features of stimuli to an atypical degree and may prefer specific geometric shapes and patterns (Pierce et al., 2011, 2016). Pierce and colleagues (2011) found that autistic toddlers spent significantly more time fixating on dynamic geometric images in comparison to neurotypical children or children with other developmental delays. In a follow up study, Pierce et al. (2016) corroborated their original findings by examining a wider range of participant groups and individual characteristics. They found that autistic toddlers fixated on geometric images significantly more than neurotypical toddlers, toddlers with developmental delay, and toddlers with other conditions thought to affect development (e.g. premature birth, or those at genetic risk of ASD; Pierce et al., 2016). They also discovered that a particularly strong preference for geometric images predicted cognitive, language, and social skills in the autistic group - children with stronger preferences demonstrated more profound deficits in all domains. The authors conclude that enhanced visual preference for geometric patterns may be an early developmental biomarker of autism symptom severity. These studies suggest that high perceptual salience can influence novel word learning in autistic children through atypical allocation of attention. If stimuli are the focus of learning, then stronger representations may be formed due to enhanced focus. However, if competitor stimuli are more salient than target stimuli, autistic children may allocate their attention away from novel targets and towards the distractors, diminishing appropriate learning.

1.5.3. Theoretical Accounts of Attentional Differences in Autism

Autistic children's relatively restricted visual attention has important implications for word learning. If children cannot flexibly attend to an array of items, they may lose informants from the environment as they cannot easily shift their attention away from distractor stimuli towards informative cues and target stimuli. This theoretical perspective has been aptly named auditory-visual misalignment (Venker et al., 2018). Formulating this theory, Venker et al. (2018) reviewed an array of literature examining how autistic children's attentional allocation affects their learning. They suggested that children's attentional focus determines statistical input, meaning that auditory-visual co-occurrences perceived by the child determines what can be learned from the input. The authors conclude that attending to the right thing at the right time facilitates vocabulary development, and looking at the wrong thing likely inhibits accurate referent selection and potentially generates erroneous or incomplete word-object mappings (e.g. Baron-Cohen et al., 1997; Ellis et al., 2014; Tenenbaum et al., 2017; Venker et al., 2018). Crucially, further empirical research is required to directly test this theory and inspect the influence of looking allocation on word learning accuracy.

Differences in autistic children's visual attention, such as those from Venker et al. (2021, 2022) that demonstrate autistic children's difficulties disengaging from perceptually salient stimuli, can be explained by stimulus 'over-selectivity' theory. This theory proposes that an individual responds only to a subset of stimuli in their environment, thus restricting learning (Lovaas et al., 1971, 1979). Evidence for over-selectivity has been identified in autistic children (Lovaas et al., 1971) as well as neurotypical preschoolers (Reed et al., 2013). This theory suggests that children are less able to attend to multiple cues due to their narrowed attentional focus. For example, in Parish-Morris and colleagues' study (2007), autistic children were unable to ignore perceptual salience even when alternative social cues were utilised to redirect children's attention. However, the age at which stimulus over-selectivity diminishes is unclear. Some research indicates that stimulus over-selectivity disappears at around three years of age in neurotypical children (Reed et al., 2013), although research has shown the phenomenon to be present in some older children, as well as autistic children and children with mental ages older than three years (Rincover & Ducharme, 1987; Schover & Newsom, 1976; Wilhelm & Lovaas, 1976). Consequently, it may be more suitable to regard stimulus over-selectivity as a general developmental cognitive delay rather than a milestone that occurs at a fixed time point (Reed et al., 2013).

Susceptibility to stimulus over-selectivity can be explained by executive function. Executive function encompasses many cognitive abilities, including working memory, inhibitory control, and shifting of attention (Blair, 2016). Research has demonstrated executive function problems in both autistic children and adults (Hughes et al., 1994; Ozonoff et al., 1991). Sáez and colleagues (2012) split executive functions into different attentional processes, defined as: attention-memory, holding and updating information; attention-set shifting, shifting away from inappropriate responses; and attention-inhibitory control, suppressing inappropriate activity. They found that attention-memory, including focusing on and retaining information, which is important for referent selection and retention, was the strongest predictor of reading performance. If children's attentional mechanisms are impaired, they may not focus on correct information at the right times during word learning, thus inhibiting language acquisition.

Moreover, executive function has been shown to account for word learning differences (Kapa & Erikson, 2020) and cognitive inflexibility (Hughes et al., 1994). In particular, more developed executive function skills help children with attentional control -

the ability to focus on a task and ignore irrelevant information (Garon et al., 2008; Posner & Rothbart, 2000). It may be that, given autistic children’s less developed executive function skills, they are less able to ignore irrelevant information and focus on correct stimuli (e.g. Elison et al., 2012; Venker et al., 2021, 2022). Overall, stimulus over-selectivity suggests that autistic children may be less able to attend to multiple cues and stimuli within a learning environment, and thus learning is limited only to the cues and stimuli that they do successfully attend to.

Similarly, the theory of weak central coherence proposes that autistic individuals are more likely to attend to and retain specific details rather than the global form or meaning (Frith, 1989). This tendency to have focused and narrowed attentional constraints are paramount to the phenotype of autism (Behrmann et al., 2006). Weak central coherence is often used to explain circumscribed interests, and the exceptional skills some autistic individuals demonstrate on specific tasks or topics (Happé & Frith, 2006; van der Geest et al., 2002). In Venker and colleagues’ recent studies (2021, 2022), high perceptual salience perhaps disrupted word identification in autistic children, but not neurotypical children, due to differences in processing stimuli. In line with the weak central coherence account, autistic children may have focused on specific salient aspects of the images, such as the patterns (e.g. Pierce et al., 2011), detracting from the image as a whole. As such, word identification could have been limited by reduced ability to encode the whole visual scene. By contrast, autistic children allocate greater attention to stimuli that appeal to their interests, which may be advantageous for learning (e.g. Sasson et al., 2011). However, studies examining how autistic children’s interests influence their stimuli exploration and word learning are extremely limited. Directly investigating how children’s attention can be manipulated by interests and preferences can illuminate how specific stimuli may affect referent selection and novel word retention.

1.6. Thesis Objectives

This literature review has highlighted how neurotypical children can learn words when attentional mechanisms are manipulated in various ways. Although autistic children commonly have language delays in comparison to neurotypical children, we know relatively little about how their word learning is affected by attentional mechanisms and how these effects persist over delays of more than a few minutes (e.g. 24 hours). Some research suggests that when autistic children's attention is directed towards to-be-learned stimuli in a manner that appeals to their strengths, they can achieve accurate word learning which is comparable to neurotypical peers matched on receptive vocabulary (e.g. Akechi et al., 2011, 2013; Parish-Morris et al., 2007). Additionally, research investigating the influence of sleep on novel word consolidation in autism only targets a sub-set of linguistically and intellectually able autistic children (Fletcher et al., 2020; Henderson et al., 2014). It is therefore unclear whether language delayed autistic children also experience beneficial effects of sleep consolidation on word learning in the same manner as neurotypical children and linguistically able autistic children, and how such effects may interact with attentional preferences for stimuli.

This thesis seeks to advance understanding of how autistic children's preferences and interests influence their identification and retention of novel word meanings. The three studies that comprise this thesis each investigate word learning as a system across three distinct stages: referent selection, retention after a 5-minute delay, and retention after a 24-hour delay. Neurotypical and autistic children matched on receptive vocabulary abilities of approximately five years were taught novel word-referent mappings via ME (Bion et al., 2013; Horst & Samuelson, 2008). All studies utilised a within-subjects design, with all participants experiencing conditions involving stimuli with varying levels of attentional salience. The studies draw upon a range of methods used within the extant literature but are some of the first to explore how children's interests affect referent selection alongside short- and long-term retention in neurotypical and autistic populations. By comparing these populations, this project examines the importance of attentional flexibility to word learning by demonstrating the consequences of preferential biases to selective stimuli.

The studies utilised touch-screen technology to measure accuracy and response times. The use of touchscreens, such as tablets, has been highly successful in promoting learning and engagement with both neurotypical and autistic children (El Zein et al., 2016; Fletcher-Watson et al., 2016; Lee et al., 2015). For autistic children, it is possible that touchscreens may reduce environmental stress and the pressures of processing time that would otherwise be present during social interaction, meaning that greater cognitive resources can be allocated to the learning task (Southall, 2013). In terms of response time, there is variable evidence of whether autistic children take longer to generate correct responses during word learning, so our research aims to disentangle under what conditions autistic children may, or may not, respond slower than neurotypical peers (Hartley et al., 2020; Ricketts et al., 2015). Accuracy data from touch-screen responses will reveal how accurately children can identify and retain novel words when target and distractor stimuli are high or neutral interest. Participants' individual characteristics were assessed via a battery of standardised assessments. These characteristics were analysed to examine how the groups differed, and how individual variability within the groups influenced accuracy across the word learning stages.

In addition to recording children's response accuracy and speed, cameras recorded children's looking behaviour during the studies. This innovative approach enables us to explore how children's online attention during word learning influences their identification and retention of meaning, and how gaze patterns may differ across autistic and neurotypical populations. These data are analysed in Studies 2 and 3 and are some of the first to demonstrate how measures of visual attention and behavioural word learning accuracy measures inter-relate.

The studies that comprise this thesis differ in terms of how experimental manipulations were employed to investigate the impact of attentional salience of target and distractor stimuli on referent selection and retention. In Chapter 2 (Study 1), children learnt names for novel animals in the ‘high interest’ condition and novel objects in the ‘neutral interest’ condition, via referent selection. Competitor stimuli during referent selection were familiar animals, as it is well-documented that children generally prefer animal stimuli over non-animal stimuli (Celani, 2002; Prothmann et al., 2009) and many autistic individuals are particularly fond of animals (Martin & Farnum, 2002). In the object condition, targets were low in interest but high in categorical discriminability in comparison to familiar animal distractors. In the animal condition, targets were high in interest but low in categorical discriminability. Investigating how categorical salience influences word learning will emphasise the importance of attentional flexibility and perceptual distinction to word learning and illuminate how interests may help or hinder the word learning of autistic and neurotypical children.

In Chapter 3 (Study 2), high interest stimuli were novel animals and neutral interest stimuli were novel objects, but unlike Study 1, children mapped word-referent associations in the presence of familiar object competitors. In Study 1, overall interest in the word learning context may be greater than in Study 2 due to the presence of animal distractors, although high interest novel animals may not stand out categorically from familiar animals. In Study 2, overall interest in the stimuli presented at each naming event might be lower due to object distractors, but high interest novel animals are categorically salient in comparison to familiar objects. Therefore, across Studies 1 and 2, we examine how overall interest in stimuli vs. categorical salience of targets influences word learning. In Study 2, the unique combination of looking time data and explicit behavioural responses will illuminate how children's preferential biases affect visual attention during learning, and how this in turn influences response accuracy.

Chapter 4 (Study 3) did not incorporate animal stimuli targeting children's pre-existing interests and instead investigated children's unique preferences for experimental stimuli. Children identified their 'liked' and 'disliked' novel objects, and these specific selections were allocated to liked and disliked conditions respectively. As in Study 2, children learnt novel words via referent selection in the presence of familiar objects. This study examines how individual preferential biases to selective stimuli influences attentional allocation, elucidating how these predilections impact word learning in autistic and neurotypical populations.

In terms of looking behaviour, across Studies 2 and 3 we expected autistic children to spend longer looking towards novel items regardless of whether they were intended targets due to difficulties disengaging attention from interesting stimuli, particularly in the high interest conditions (Elsabbagh et al., 2009, 2013; Landry & Bryson, 2004; Sacrey et al., 2014). We expected autistic children to make more frequent looks to target stimuli due to longer processing times required to generate correct responses. We also anticipated that increased visual attention to targets would predict response accuracy across conditions, groups, and task stages. We expected increased attention at referent selection to also predict superior retention due to greater opportunities for novel word-referent encoding at training (e.g. Hilton et al., 2019; Hilton & Westermann, 2017).

Importantly, these studies will advance theoretical understanding of word learning by disentangling the influence of preferential biases to selective stimuli across different word learning stages and populations. These data will also reveal the importance of flexible visual attention for successful word learning. Our studies will highlight effective ways to investigate learning in autistic children experimentally and emphasise the importance of selecting appropriate dependent variables when investigating learning across populations. Furthermore, these studies have the potential to impact educational and clinical contexts by discovering multiple ways in which learning can be enhanced in autistic children. These findings can subsequently be applied in future interventions and within the design and creation of learning materials.

Chapter 2: Is autistic children's word learning facilitated or hindered by high interest distractors?

2.1. Chapter Introduction

Word learning is contingent on children attending to the right information at the right time. Whilst often a topic of debate, autistic children's frequent language delays are commonly attributed to attentional differences that affect children's intake of visual and auditory information (Arunachalam & Luyster, 2018; Venker et al., 2018). Autistic children often have difficulties flexibly allocating their attention across a range of stimuli (Noterdaeme et al., 2002; Ozonoff et al., 1994), and their attentional allocation is influenced by circumscribed interests, which may result in limited environmental informants during word learning (Elsabbagh et al., 2009, 2013; McGregor et al., 2013; Sacrey et al., 2014). Currently, these theories have not been directly tested across word learning stages in autism. Elucidating how children's interests affect word learning is an imperative step to advance understanding of optimal learning contexts to afford effective language acquisition.

In this chapter (Study 1), autistic and neurotypical children learnt novel animal and object names in the presence of familiar animal distractor stimuli. In one condition, children learnt names for object stimuli, which were low in interest but high in categorical discriminability compared to distractors. In the other condition, children learnt names for animal stimuli, which were high in interest but low in categorical discriminability. Investigating these differences in categorical salience will emphasise the importance of attentional flexibility to word learning. It will also enlighten how interests may help or hinder children's word learning, providing insight for practical applications.

Author contribution: *Charlotte Rothwell*: design, data collection, analysis, writing, review. *Gert Westermann*: design, review. *Calum Hartley*: design, analysis, review.

2.2. Abstract

Word learning is influenced by children's attention and interests – determining correct word-referent mappings requires paying attention to the right things at the right times. Whilst autistic children's fundamental word learning mechanisms appear to be intact, their perseverative interests and atypical attention allocation may affect word learning. We investigated whether autistic children with delayed language development and neurotypical children matched on receptive vocabulary differ in accuracy when learning words associated with novel animals (high interest stimuli) and objects (neutral interest stimuli), in the presence of high interest animal distractors during referent selection. In a fast-mapping task, both groups identified meanings of novel words associated with unfamiliar animals and objects with comparable accuracy and retained them after a 5-minute delay. Greater interest in animals predicted superior retention accuracy in autistic children at 24-hour retention, due to heightened interest in the visual scene at encoding. After 24 hours, autistic children retained more novel object names than neurotypical children. This may be due to the perceptual contrast between categorically distinct object stimuli and familiar animal stimuli enhancing encoding at referent selection. Thus, following a period of sleep, these more strongly encoded representations were consolidated for greater retrieval after a 24-hour delay. These findings suggest that heightened attention to interesting non-target competitors during mutual exclusivity, plus increased visual distinctiveness between targets and distractors, both afford more robust encoding of word-target representations in autism.

Keywords: Word learning; Autism; Interests; Attention; Referent Selection; Retention

2.3. Introduction

Language acquisition is crucial for children's cognitive and social development (Carpenter et al., 1998; Tomasello, 2003). Neurotypical children begin to produce language around 12-18 months (Tager-Flusberg et al., 2009; Zubrick et al., 2007), and learn the meanings of over two hundred words before two years of age (Dale & Fenson, 1996), increasing to ten words per day by school age (Bloom, 2002a). This milestone is often delayed in autistic children, who begin to speak at ~38 months on average (Anderson et al., 2007), and demonstrate delayed receptive vocabulary profiles (Artis & Arunachalam, 2023; Saldaña, 2023). Although many autistic individuals develop functional language skills over the school years (Pickles et al., 2014), approximately 25-35% of autistic children remain non-verbal or minimally verbal (Ellis Weismer & Kover, 2015; Tager-Flusberg & Kasari, 2013). Delays in autistic children's receptive and expressive language development are often attributed to difficulties acquiring new words (Ellis Weismer et al., 2020; Volden et al., 2011). Despite their problems learning language, recent evidence shows that fundamental word learning mechanisms – and relationships between them – are not atypical in many autistic children, including those with concomitant language impairments (Foti et al., 2015; Hartley et al., 2019, 2020; Rothwell et al., accepted, see Study 2). In light of this evidence for intact word learning mechanisms, it has been proposed that language learning difficulties may be attributed to atypical attentional behaviours that impair autistic children's intake of visual and auditory stimuli from their environment through distraction (Arunachalam & Luyster, 2015, 2018; Venker et al., 2018). However, few studies to date have empirically studied how children's attentional mechanisms influence word learning in autistic or neurotypical development. Here, we investigate whether autistic and neurotypical children's learning of words associated with high- and neutral-interest stimuli is impacted by the presence of attentionally salient distractor stimuli.

Word learning is a complex, multi-stage process. Initially, children must identify the novel word during speech, and link the novel word to an intended referent (referent selection; Carey & Bartlett, 1978). Then, the child must store the correct word-referent association for later retrieval (retention; Gleitman, 1990). The relationship between these word learning mechanisms is explained by the 'dynamic associative account' (McMurray et al., 2012; Samuelson & McMurray, 2017), which proposes that referent selection and retention utilise separate 'fast mapping' and 'slow associative learning' processes that occur on different timescales (McMurray et al., 2012).

Fast mapping occurs when children form rapid, in-the-moment, associations between novel words and their referents (Kucker et al., 2015; Samuelson & McMurray, 2017). During fast mapping, children must select the correct target from a range of distractors, overcoming the challenge of referential ambiguity – a novel word could have multiple potential referents (Cartmill et al., 2013; Markman, 1989). Referent selection is facilitated by lexical heuristics, one of which is the mutual exclusivity principle (ME; also known as disjunctive syllogism; Carey 1978). ME refers to the assumption that object labels are mutually exclusive, and thus each referent only has a single label. This principle is demonstrated when a single unfamiliar object is presented alongside one or more familiar objects, and children are asked to identify the referent of a novel word (Halberda 2003, 2006). Since children know the names of familiar objects, they use ME to deduce that a novel word must refer to the unfamiliar object (Markman & Wachtel, 1988; Merriman & Bowman, 1989).

However, accurate referent selection alone does not constitute word learning – neurotypical toddlers who perform at ceiling on a fast-mapping task often forget the referents of novel words after a five-minute delay (Horst & Samuelson, 2008; Gurteen et al., 2011). In order to truly learn a word, children must successfully encode and retain a word-referent association in memory for later retrieval (Vlach & DeBrock, 2019). The dynamic associative account posits that retention is driven by associative learning mechanisms that gradually strengthen word-object relationships over multiple exposures across situations and contexts (McMurray et al., 2012). In support of this theory, Hartley et al. (2020) report that children's retention is more accurate following cross-situational learning (involving multiple word-referent exposures) than fast mapping (involving individual naming events).

Evidence from neurotypical development clearly demonstrates that word learning and attention are inter-related. To correctly identify a novel word's intended referent from myriad distractors, children must attend to multiple components of their environment, coordinating their attention to visual and auditory stimuli concurrently (Axelsson et al., 2012; Samuelson et al., 2017). This requires children to shift their attention between stimuli, ignore competing irrelevant stimuli, and attend to correct stimuli during naming events. Research has revealed that neurotypical children can flexibly shift their attention to multiple information sources within their environment to decipher the meanings of words (Hollich et al., 2000; Preissler & Carey, 2005). Moreover, the extent to which neurotypical children attend to stimuli during referent selection also influences their subsequent retention of newly learned words (e.g. Hilton et al., 2019; Hilton & Westermann, 2017, Smith & Yu, 2013). Heightened interest in certain objects increases children's attentional focus, strengthening their encoding of word-referent associations (Spiegel & Halberda, 2011). For example, Ackermann et al. (2020) discovered that children learned more words associated with interesting stimuli (e.g. animals), compared to stimuli they were less interested in (e.g. objects), suggesting that category preferences shape children's word learning.

Word learning environments are often rife with clutter, including numerous familiar and unfamiliar non-target objects (distractors) with known and unknown names. While children may show superior learning for words associated with categorically interesting target stimuli (e.g. animals; Ackermann et al., 2020), the presence of such stimuli as non-target distractors during referent selection could have a disruptive influence on subsequent retention. Only a small number of studies to date have investigated how manipulating children's attention to distractor stimuli affects word learning. For example, Pomper and Saffran (2018) found that neurotypical children could retain the names of novel objects that were labeled in the presence of familiar objects with low salience, but not when familiar objects were highly salient. Similarly, Axelsson et al. (2012) found that neurotypical 2-year-olds' word learning was hindered when attention to distractors was not suppressed, compared to when target objects were highlighted and distractor objects concurrently dampened (Zosh et al., 2013). Axelsson and Horst (2014) also suggested that when the same distractors were repeated across multiple referent selection trials, neurotypical children could more easily retain novel words than when different distractors objects appeared on each trial. This indicates that contextual repetition facilitates novel word encoding due to the reduced demand required to rule out distractors over target stimuli. Overall, these studies reveal that increased attention to distractor objects can limit children's ability to encode sufficient information about novel targets, resulting in fragile word-referent associations that may not be retrievable after a delay.

Like neurotypical children, autistic children with varying language abilities can utilise lexical heuristics such as ME to accurately identify novel word-object pairings (de Marchena et al., 2011; Parish-Morris et al., 2007; Preissler & Carey, 2005). Autistic children can also successfully disambiguate novel word meanings through tracking cross-situational word-object correspondences over multiple naming events (Hartley et al., 2020; Venker, 2019). Interestingly, Hartley et al. (2019) found that receptive vocabulary in autistic children was predicted by their ability to accurately employ ME to identify novel word meanings at referent selection, suggesting that optimising children's environment during naming events could have long-term benefits for vocabulary development. Potentially then, poor environments that hinder autistic children’s accurate referent selection could have detrimental consequences for their long-term vocabulary development.

Whilst it is evident that autistic children can successfully perform referent selection in an array of settings, little research has examined their ability to retain novel words. Recent studies have found that language delayed autistic children can successfully retain the names of novel objects as accurately as neurotypical children matched on receptive vocabulary following ME-based referent selection (e.g. Carter & Hartley, 2021; Hartley et al., 2019). Rothwell et al. (accepted, see Study 2) investigated how attentional biases influence the relationship between fast mapping and retention by comparing autistic and neurotypical children’s ability to learn labels associated with high interest stimuli (novel animals) and neutral interest stimuli (novel objects). Importantly, the distractors in this study were all familiar neutral interest stimuli (e.g. balloon, chair, hat). They found that language delayed autistic children and neurotypical children matched on receptive vocabulary both successfully identified the names of novel animals and objects via ME. When comparing retention of autistic and neurotypical children after a 5-minute delay, both groups retained novel word-object associations with similar accuracy, but autistic children retained word-animal associations with *greater* accuracy. Unexpectedly, autistic children retained names for novel objects and novel animals with significantly greater accuracy than neurotypical children when retention was tested after 24 hours. However, although autistic children were unimpaired on measures of learning accuracy, they were significantly slower to indicate correct referents during referent selection. These findings highlight that autistic children can learn words for different types of stimuli at least as accurately as vocabulary-matched neurotypical children when distractors facilitate ME and do not excessively draw children’s attention away from targets. Nevertheless, autistic children’s slower reaction times at referent selection suggest that their processing of audio-visual input during early word learning processes may be less efficient (also see Hartley et al., 2020).

Whilst neurotypical children can flexibly shift their attention across their environment during word learning, autistic children often have difficulty allocating sustained or selective attention (e.g. Courchesne et al., 1994; Ozonoff et al., 1994, 2004). Due to diagnosis-defining restricted and repetitive behaviours and interests (RRBIs; Kanner, 1943; Honey et al., 2012; Richler et al., 2007), autistic children often fixate intensely on a narrow range of topics or stimuli. Consequently, autistic children may often restrict attentional resources to information pertaining to their specific interests, and struggle to disengage from these stimuli, potentially limiting input and informants for word learning (Ellis et al., 2014; Fitneva & Christiansen, 2011; Oakes, 2011). For example, Walton and Ingersoll (2013) found, that unlike neurotypical children, autistic children may mismap novel words to their own focus of attention rather than utilising cues from a social informant to identify intended referents (also see Baron-Cohen et al., 1997). Similarly, Parsons et al. (2019) found that children at high risk of developing autism engaged less with target objects than neurotypical children, and spent more time distracted by other environmental features such as an experimenter's face. The limited time that autistic children spent focusing on target objects was associated with lower concurrent and longitudinal verbal abilities, demonstrating that atypical distribution of attention could be a limiting factor for language development in infants at familial risk for autism. Together, these studies indicate that restricted interests and attentional mechanisms could impede autistic children's word learning by limiting their ability to intake the right visual-auditory stimuli at the right times. If autistic children are more interested in non-target distractors than the intended referent of a novel word, they may struggle to disengage from these distracting stimuli and be at increased risk of forming inaccurate word-referent mappings.

How autistic children’s performance on word learning tasks is influenced by their attention to distractor stimuli is largely unknown. Research from Venker et al. (2021) demonstrated that when distractor stimuli were more salient than target stimuli, autistic children decreased their visual attention to the target, but neurotypical children were unaffected. Similarly, Venker et al. (2022) highlighted that perceptual salience of target stimuli also disrupted the attentional allocation of autistic children and thus novel word recognition. However, little is known about how categorical salience determines autistic children’s word learning, or how retention is affected by attention. Studying autistic children’s learning in the presence of high-interest distractors may provide important insights into how autistic children’s language acquisition may be hindered by their environment. Crucially, discovering how autistic children’s word learning is influenced by distractors that are more or equally interesting relative to target stimuli will inform understanding of optimal word learning settings. This knowledge will not only advance understanding of *why* many autistic children demonstrate language deficits in natural learning environments that are rife with distractions (e.g. Anderson et al., 2007; Tager-Flusberg et al., 2005), but also indicate *how* practitioners can craft audio-visual environments to scaffold learning.

The objective of the present study was to investigate how autistic and neurotypical children’s word learning is impacted by the presence of non-target distractors that are as salient, or more salient, than to-be-learned target referents. Autistic and neurotypical children matched on receptive vocabulary completed two word learning tasks on different days. In one task, children learned names for four novel animals via ME-based fast mapping. In the other task, children learned names for four novel objects. In both conditions, the distractor stimuli were familiar animals. It is well-documented that children generally prefer animal stimuli over non-animal stimuli (Ackermann et al., 2020; Celani, 2002; Prothmann et al., 2009) and many autistic individuals are particularly fond of animals (Martin & Farnum, 2002). The strength of children’s interest in animals was measured via a caregiver questionnaire and we examined whether this predicted children’s word learning performance. Retention of the novel names was tested after 5 minutes and 24 hours, based on literature stating that neurotypical 2-year-olds often forget new words after just five minutes (Horst & Samuelson, 2008), and a period of sleep is necessary to assimilate novel words (Dumay & Gaskell, 2007). Response time and accuracy were measured using a touch-screen computer.

Based on previous evidence (e.g. Carter & Hartley, 2021; Hartley et al., 2019, 2020; Rothwell et al., accepted, see Study 2), we anticipated that both neurotypical children and autistic children would apply the ME principle to accurately perform referent selection regardless of whether distractors were more salient than targets. However, we predicted that children would be slowest to map labels during referent selection (e.g. Rothwell et al., accepted, see Study 2), particularly in the presence of target objects, which may be less interesting than the animal distractors (whereas target animals are equally interesting). Since the word learning mechanisms of autistic children are unimpaired, we also expected that children in both groups would accurately retain labels for both object and animal stimuli (Rothwell et al., accepted, see Study 2). Potentially, children with a particularly strong interest in animals could find it more difficult to retain names for target objects due to increased competition from distractor animal stimuli during training. In particular, autistic children may demonstrate poorer short- and long-term retention due to increased interest in animal non-target stimuli during encoding, especially during the object condition, where autistic children may have greater difficulty disengaging attention away from animal distractors to focus on target objects. By comparing neurotypical and autistic populations, this research will reveal the importance of attentional flexibility to word learning and determine the consequences of preferential biases to selective stimuli. Investigating whether word learning is affected by children's categorical interests will have significant implications for understanding their language acquisition.

2.4. Method

Participants

Participants were 18 autistic children (15 males, 3 females; *M* age = 90.61 months; *SD* = 19.97) recruited from specialist schools, and 19 neurotypical children (10 males, 9 females; *M* age = 51.68 months; *SD* = 17.38) recruited from mainstream schools, nurseries, and Lancaster University BabyLab (see Table 1). Please refer to Appendix E for additional information about the involvement of individual participants in multiple studies reported in this thesis. All participants had normal or corrected-to-normal colour vision, and were monolingual, native English speakers. Autistic children had received a pre-existing diagnosis from a qualified clinician, using standardised instruments (i.e. Autism Diagnostic Observation Scale and Autism Diagnostic Interview – Revised; Lord et al., 1994, 2002) and expert judgement. Diagnoses were confirmed via the Childhood Autism Rating Scale 2 (ASD *M* score = 33.78, *SD* = 10.23; NT *M* score = 16.34, *SD* = 1.91; Schopler et al., 2010). This measure was usually completed by class teachers, but it was completed by caregivers for ten neurotypical children who were tested at our BabyLab due to COVID-19 restrictions. Autistic children were significantly older, $t(35) = -6.33$, $p < .001$, $d = 2.08$, and had significantly higher CARS scores, $t(33) = -7.30$, $p < .001$, $d = 2.37$, than the neurotypical children.

Groups did not significantly differ on receptive vocabulary as measured by the British Picture Vocabulary Scale 2 (BPVS; ASD: *M* age equivalent = 56.78 months, *SD* = 23.54; NT: *M* age equivalent = 60.26 months, *SD* = 24.69; Dunn et al., 1997), $t(35) = 0.44$, $p = .66$. Receptive vocabulary was used to match the groups as it demonstrates children's ability to learn word-referent relationships (Bion et al., 2013). Children's expressive vocabulary abilities were measured using the Expressive Vocabulary Test 2 (EVT; ASD: *M* age equivalent = 51.22 months, *SD* = 29.45; NT: *M* age equivalent = 58.16 months, *SD* = 19.91; Williams, 2007), or the expressive language module of the Mullen's Scales of Early Learning (Mullen, 1995) for children who scored below the baseline on the EVT. The groups did not significantly differ on expressive vocabulary abilities, *t*(35) = 0.84, *p* = .40.

Children's non-verbal intellectual abilities were measured using the Leiter-3 (Roid et al., 2013). The average IQ score for the autistic group was *M* = 79.08 (*SD* = 13.29), and the average IQ of the neurotypical group was significantly higher at *M* = 103.13 (*SD* = 12.28), *t*(26) = 4.98, *p* <.001, *d* = 1.88. Scaled IQ scores could not be calculated for four neurotypical children as they were younger than 3-years-old. However, the groups' raw scores on the Leiter-3 did not significantly differ (ASD: *M* = 60.54, *SD* = 16.28; NT: *M* = 57.42 months, *SD* = 18.58), *t*(30) = -0.49, *p* = .63, suggesting that when age was not considered, their non-verbal cognitive abilities were similar at the time of testing. To assess attentional behaviours, the Conner's Teacher Rating Scale (CTRS-15; Pupura & Lonigan, 2009) was completed by children's class teachers, or the caregivers of the ten neurotypical children who were tested in our BabyLab. The mean raw scores for the autistic children (*M* = 18.62, *SD* = 11.72) were significantly higher than those of the neurotypical children (*M* = 9.63, *SD* = 5.78), *t*(33) = -2.95, *p* = .006, *d* = 0.97. The Repetitive Behaviour Questionnaire was completed by the participants' caregivers to assess the extent of their restrictive and repetitive behaviours (RBQ; Leekam et al., 2007). Autistic children (*M* score = 42.17*, SD* = 8.97) had significantly higher scores on the RBQ than neurotypical children (*M* score = 25.58, *SD* = 5.03), *t*(35) = -6.99, *p* <.001, *d* = 2.28.

Finally, to confirm that we recruited participants who were interested in animals and ensure animals would be a 'high interest' category for such children, we created a caregiver questionnaire assessing the extent of children's animal interests (min-max scores: 0-34). Autistic children (*M* score = 24.17, *SD* = 5.71) and neurotypical children (*M* score = 23.42, *SD* = 4.03; see Appendix A) did not differ significantly on this measure, $t(35) = -0.46$, $p = .65$. One autistic child was excluded from the study due to their lack of interest in animals.

An additional five participants were excluded from the study; one neurotypical participant who was unable to complete the touch-screen task, two neurotypical participants who scored above the 'low to minimal symptoms' threshold on the CARS-2, and two children who did not complete both experimental conditions due to school closures during the COVID-19 pandemic (one autistic child) or absence (one neurotypical child). Due to pandemic-related school closures, five autistic children did not complete the Leiter-3, and the teachers of two autistic children did not complete the CTRS-15 and CARS-2. These participants were retained in the study as they completed the experimental tasks and other participant variables.

All procedures in the present study were in accordance with the ethical standards of institutional and national research committees. Informed consent was obtained from caregivers prior to children's participation and a debrief was provided after participation.

Table 1

Characteristics of autistic and neurotypical Participants (SD and Ranges in Parentheses)

Pop.	N	Gender	Chron. Age (*M*, months)	BPVS. age equiv. (*M*, months)	Express. Lang. age equiv. (*M*, months)	CARS raw score (*M*)	Leiter-3 raw score (*M*)	CTRS raw score (*M*)	RBQ raw score (*M*)	Animal Interest score (*M*)
NT	19	9 females, 10 males	51.68 (17.38, 28-93)	60.26 (24.69, 34-101)	58.16 (19.91, 28-104)	16.34 (1.91, 15-21.5)	57.42 (18.58, 31-102)	9.63 (5.78, 2-20)	25.58 (5.03, 20-35)	23.42 (4.03, 16-34)
ASD	18	3 females, 15 males	90.61 (19.97, 67-132)	56.78 (23.54, 24-97)	51.22 (29.45, 5-92)	33.78 (10.23, 19-52)	60.54 (16.28, 38-83)	18.62 (11.72, 5-40)	42.17 (8.97, 29-59)	24.17 (5.71, 17-34)
Group comparison t-test (*p*)			<.001	.66	.40	<.001	.63	.006	<.001	.65

Note. NT: neurotypical; ASD: autism spectrum disorder; BPVS: British Picture Vocabulary Scale, CARS: Childhood Autism Rating Scale, CTRS: Conner's Teacher Rating Scale, RBQ: Repetitive Behaviour Questionnaire. Participants experienced both conditions.

Materials

The study was administered via a touch-screen computer running MATLAB. Audio stimuli for the word learning task included eight two-syllable unfamiliar words (kita, teebu, ipis, gazzer, colat, blicket, regli, zepper) selected from academic sources such as the NOUN database (Horst & Hout, 2016). Visual stimuli included high-resolution colour photographs of 4 unfamiliar objects, 4 unfamiliar (but real) animals (see Figure 1), 6 familiar objects, and 22 familiar animals, all presented on a grey background. All photographs were approximately 6cm^2 and 500 x 500 pixels when displayed on the screen. Unfamiliar stimuli were selected because children would not have pre-existing linguistic labels for them. Familiar objects and animals were selected on the basis that most children understand their linguistic labels by around 16 months (Fenson et al., 1994). Three pictures of six familiar objects and six familiar animals were employed in the warm-up trials (monkey, owl, tiger, squirrel, giraffe, mouse, clock, train, handbag, top, swing, key), mirroring the combinations of categories children would experience in the forthcoming trials. Thus, children saw a different combination of the three images dependent on whether they were viewing animals only, objects only, or a mixture of animals and objects in that test phase. Pictures of 16 familiar animals were presented during referent selection trials in both the animal and object conditions. These were divided into two sets and counterbalanced across conditions (1. turtle, bear, horse, dog, lion, frog, duck, butterfly; 2. pig, rabbit, cat, bird, bee, fish, elephant, cow). Familiar animals allocated to the two conditions were matched on mean comprehension age (15 months for both sets) and frequency of animals belonging to particular categories (e.g. insects, birds). Familiar animals within each set were divided into pairs and presented alongside an unfamiliar object or animal in referent selection trials of both conditions. In every trial type, three pictures were presented side by side. The names of stimuli presented together were selected on the basis that they were phonologically distinct, and their images clearly contrasted in shape and colour.

Stimuli names were recorded by a female speaker from the local area and presented through the computer's integrated speakers. The audio files were recorded using a Sony ECM-MS907 Digital Microphone, and the software Audacity 2.2.2. The auditory stimulus were edited for timing and clarity in Audacity, and the volume of all files normalised. The carrier phrases (e.g. "Can you see the [label]", "Touch the [label]") and the labels (e.g. "elephant", "kita") were edited separately, so they were all distinct files. However, when the MATLAB program was used to run the experiment, the audio files were presented sequentially. This was to ensure that there were no differences in the carrier phrases that may offer a hint to children regarding the forthcoming labels. Three web cameras attached to the left, right, and centre of the computer were used to record participants' visual attention and behaviour during the study, although these data are not reported in the present paper.

Figure 1

Sets of unfamiliar objects and animals used in the word learning task

Procedure

Since the present study aimed to investigate how animal interests, a common fascination of many autistic and neurotypical children (Martin & Farnum, 2002; Prothmann et al., 2009), affects word learning, caregivers completed a questionnaire about their child's interest in animals (see Appendix A). Examples of questions included: 'How often does your child enjoy interacting with real animals?' and 'How much does your child enjoy listening to stories about realistic animals?' (responses: 1 - they don't particularly enjoy it, 2 - they enjoy it a little, 3 - they enjoy it a lot, 4 - they really, really enjoy it).

Participants were tested individually in their own school or nursery, or Lancaster University BabyLab, and were accompanied by a familiar adult when required. Children were assessed using the BPVS, EVT or MSEL, and Leiter-3 by the researcher over multiple sessions on different days. Children completed two within-subjects conditions of the word learning task – animal stimuli and object stimuli – administered on different days (average of 7 days apart, order counterbalanced). The word learning task was delivered via a touch-screen computer. Children were seated approximately 50-70cm away from the screen on a height adjustable chair. The word learning task was very similar to that reported by Rothwell et al. (accepted, see Study 2) aside from stimuli used, and consisted of the following stages, presented in a fixed order: 1. Warm-up trials, 2. Referent selection trials, 3. Five-minute delay, 4. Retention trials, 5. 24-hour delay, 6. Retention trials (see Figure 2). The experimenter sat quietly while the participant was engaged in the task and offered verbal praise for attention and good behaviour.

Figure 2

Examples of trial types in the word learning task

Warm up trials

Prior to the study beginning, children were presented with a cartoon image of a hand that appeared in each of three touch-screen panels. The experimenter asked the child to "Put their hand on the picture" in order to encourage children to become comfortable touching the screen. Then, children completed three warm-up trials. Children were instructed to "Put your hand on the picture that the computer asks you to". During warm-up trials, children were presented with images of three familiar stimuli in the left, middle, and right sections of the touchscreen. After 2 seconds, participants heard "Look, '*2s gap*' [label]!", '*1s gap*', "Can you see the [label]?" '*1s gap*', "Touch the [label]!". Children then had 12 seconds to respond. The same instructions repeated up to six times if children did not respond. Responses were accepted only after the first label utterance, preventing children from skipping through trials without hearing the requested labels. Hence, children who took longer to respond heard more repetitions of the label (this factor is examined in our analyses). Children received feedback when they made their selection; either audio praise if they responded accurately (e.g. "Well done, you touched the [label]!") or corrective feedback if they responded inaccurately ("Actually, this is the [label]. Can you touch the [label]?"). Following inaccurate responses, the correct referent was highlighted by a green border and children could retry up to five times. Children in both groups responded significantly above chance levels at first attempt on the warm-up trials ($M_{NT} = 0.96$, $M_{ASD} = 0.91$) demonstrating that they understood the task requirements and were familiar with the requested referents. The location and order of requested stimuli were counterbalanced across participants. Children saw different sets of images for the warm-up trials in each condition and for different study phases, to mirror the categories of stimuli that were present in the upcoming study phase (order counterbalanced).

Referent selection trials

Following the warm-up trials, children completed eight referent selection trials. These followed the same format as the warm-up, except children did not receive feedback following their responses. In each condition, four novel words were taught via a fast-mapping paradigm based on Horst and Samuelson (2008). Children viewed four sets of pictures (each containing one unfamiliar picture and two familiar animals). Each set was presented twice; on one trial the novel picture (either an animal or object) was requested (novel name trial: “Look, regli! Can you see the regli? Touch the regli!”), and on another trial a familiar animal was requested (familiar name trial: “Look, cow! Can you see the cow? Touch the cow!”). Familiar name trials were included to prevent participants from demonstrating a novelty preference and only attending to novel items, and instead encourage them to examine every item. This is vital as fast mapping requires children to attend to known distractors in order to exclude them as referents for a novel word (Halberda, 2003). Novel name trials promoted active learning of new word-referent pairings; since participants have prior knowledge of the familiar stimuli labels, they can decipher the referent of the novel label by applying the mutual exclusivity principle.

Trial order was pseudo-randomised with the constraints that the same set of pictures, or the same trial type (familiar name or novel name), were never presented on more than two trials sequentially. Positioning of stimuli on the screen (left, middle, right) was pseudo-randomised across trials on the basis that the target did not appear in the same location more than twice consecutively. The eight novel words were divided into two sets (1. kita, teebu, ipis, gazzer; 2. colat, blicket, regli, zepper) and were counterbalanced across conditions. Novel words were pseudo-randomly allocated to the novel items, so different novel words represented different novel items between participants. Familiar animal stimuli were divided into two sets of eight to obtain a degree of control, but these were also counterbalanced across conditions.

5-minute delay

Immediately following referent selection, children engaged in an unrelated task for five minutes (e.g. building with blocks or colouring). None of the familiar or unfamiliar experimental stimuli were present during this stage.

Retention trials

After the five-minute delay, children completed one warm-up trial to re-engage their attention (exactly as described above). Eight retention trials immediately followed (see Figure 2 for an illustration of each trial type). Each novel word was tested on two retention trials, assessing whether children's retention of newly mapped word-referent associations differed between stimulus categories they were more interested in (animals), compared to those they were less interested in (objects). Trial order was pseudo-randomised, ensuring that the same set of stimuli was never presented on more than two trials consecutively. Positioning of stimuli on the screen (left, middle, right) was pseudo-randomised across trials with the constraint that the target did not appear in the same location more than twice sequentially. Each picture served as a target on two trials and as a foil on four trials.

24-hour retention trials

After a 24-hour delay, children completed a second block of eight retention trials. Due to practical constraints, not all children experienced exactly a 24-hour delay (*M* delay = 24.0 hours, range: 22.1 – 28.0 hours). The retention trials were preceded by three warm-up trials (as described above) to remind children of the task requirements and how to respond. These 24-hour retention trials were identical to the 5-minute retention trials with the exception that stimuli were presented in different orders and combinations.

2.5. Results

Accuracy and response time data were analysed via mixed-effects models using the glmer and lmer functions from the lme4 package in R (Bates et al., 2015). Population was contrast coded as -0.5 (neurotypical) and 0.5 (autistic). Condition was coded as -0.5 (object) and 0.5 (animal). Trial type was coded as -0.5 (familiar) and 0.5 (novel). By-word referent selection accuracy was coded as -0.5 (incorrect) and 0.5 (correct) when included as a fixed effect in retention accuracy analyses. Total referent selection accuracy was coded as 0-4, incorporating the number of novel referent selection trials that children were correct on. Total 5-minute retention accuracy was coded as 0-8. Trial-level accuracy as a dependent measure was coded as 1 (correct) or 0 (incorrect) for all analyses. Number of labels heard at referent selection per novel word was coded as 1-6 (autistic $M = 2.11$, $SD = 1.11$; neurotypical $M = 1.90$, $SD = 0.75$).

The likelihood of children responding correctly by chance on each trial was 33%. All models were built up sequentially, adding in one fixed effect at a time and comparing each model with the previous best-fitting model using log-likelihood tests. Each analysis started with a baseline model containing by-participant and by-word random intercepts, with a random slope of condition x trial type per participant for referent selection, or condition per participant for retention phases. If some models in a sequence failed to converge, the random effects were simplified until all models in the sequence successfully converged. Only final models are reported; please refer to Appendix B for full details of the model building sequences.

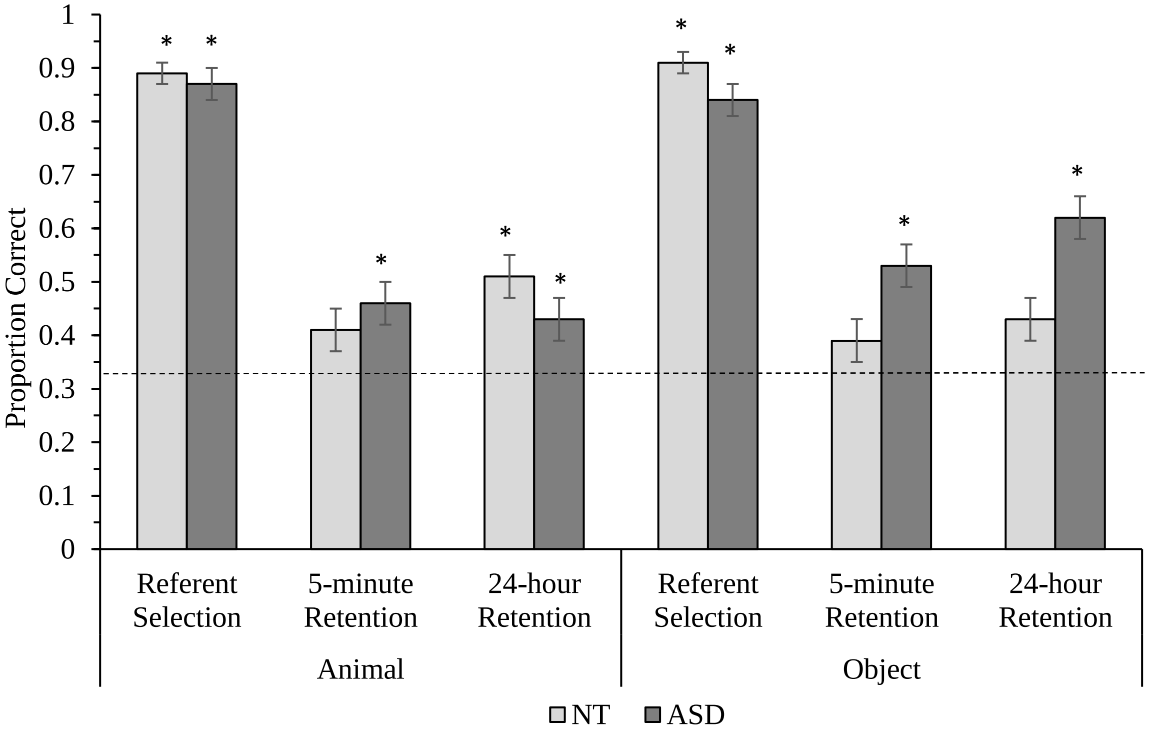
For analyses of individual differences, receptive vocabulary was coded as the participant's age equivalent on the British Picture Vocabulary Scale 2 (BPVS; Dunn et al., 1997). Expressive vocabulary was coded as the participant's age equivalent on the Expressive Vocabulary Test 2 (EVT; Williams, 2007) or Mullen's Scales of Early Learning for children who scored below the baseline on the EVT (MSEL; Mullen, 1995). Non-verbal intelligence was coded as the participant' raw score on the Leiter-3 (NVIQ; Roid et al., 2013; min-max range: 0-152). Autism severity was coded as the participant's raw score on the Childhood Autism Rating Scale 2 (CARS-2; Schopler et al., 2010; min-max range: 15-60). Attention score was coded as the participant's raw score on the Conner's Teacher Rating Scale (CTRS-15; Pupura & Lonigan, 2009; min-max range: 0-45). Repetitive behaviour was coded as the participant's raw score on the Repetitive Behaviour Questionnaire (RRB; Leekam et al., 2007; min-max range: 20-66). Animal interest was coded as the participant's raw score on our animal interest questionnaire (min-max range: 0-34). Chronological age was coded as the participant's age in months. Unfortunately, we were unable to test the effect of nonverbal intelligence, autism severity, or attention for the autistic group as interruptions in data collection caused by the COVID-19 pandemic prevented us obtaining these measures for some children.

Referent selection accuracy

Referent selection accuracy was analysed via generalised linear mixed-effects models testing the effects of population, condition, and trial type. Four trials from the neurotypical group and two trials from the autistic group were removed due to a technical error. An additional three trials were removed from the autistic group due to one child choosing multiple referents simultaneously with their head and hand. This analysis contained 283 data points from autistic children and 300 data points from neurotypical children. Descriptive statistics for referent selection accuracy are presented in Figure 3.

Figure 3

Mean referent selection, 5-minute retention, and 24-hour retention trial accuracy for neurotypical (NT) and autistic children (ASD), error bars show ± 1 SE. Stars *above columns indicate where performance was significantly different from chance, indicated by the dotted lines (*p <.05)*



The final model included a trial type x condition interaction ($z = -4.89$, $p <.001$; Table 2). This interaction was deconstructed by testing the effect of condition for familiar and novel trial types separately, revealing that children responded more accurately in the animal condition than the object condition on familiar trials ($z = 4.44$, $p <.001$; object condition $M = 0.91$; animal condition $M = 0.95$), but there was no significant effect of condition for novel trials ($z = 0.01$, $p = 1.00$; object condition $M = 0.84$; animal condition $M = 0.82$). It is noteworthy that both groups responded very accurately on novel trials with object and animal targets (neurotypical children, animal condition $M = 0.82$, neurotypical children, object condition $M = 0.92$; autistic children, animal condition $M = 0.81$, autistic children, object condition $M = 0.75$), demonstrating that they could use mutual exclusivity effectively.

We also examined whether individual differences between participants in each population predicted additional variability in their referent selection accuracy. Although the populations did not significantly differ on referent selection accuracy, it is possible that different factors contributed to their successful performance (autistic individuals can attain 'typical' performance on psychological tasks via 'atypical' routes and compensatory strategies; Happé, 1995; Norbury et al., 2010). As such, the following analyses were conducted on data from the autistic and neurotypical groups separately.

For autistic children, the final model included an additional fixed effect of expressive vocabulary ($z = 4.12$, $p <.001$; see Table 2). Higher expressive vocabulary scores were associated with more accurate referent selection in the autistic group. Note the inclusion of expressive vocabulary resulted in the trial type by condition interaction no longer being significant for the autistic children. This suggests that this effect in the combined model was driven primarily by the neurotypical children.

For neurotypical children, the final model included an additional fixed effect of nonverbal intelligence ($z = 2.72$, $p = .006$; see Table 2). Neurotypical children with higher non-verbal intelligence scores responded more accurately on referent selection trials.

Table 2

Summaries of the fixed effects in the final generalised linear mixed-effects models (log odds) of children's accuracy on referent selection trials, and how individual differences influence referent selection trial accuracy

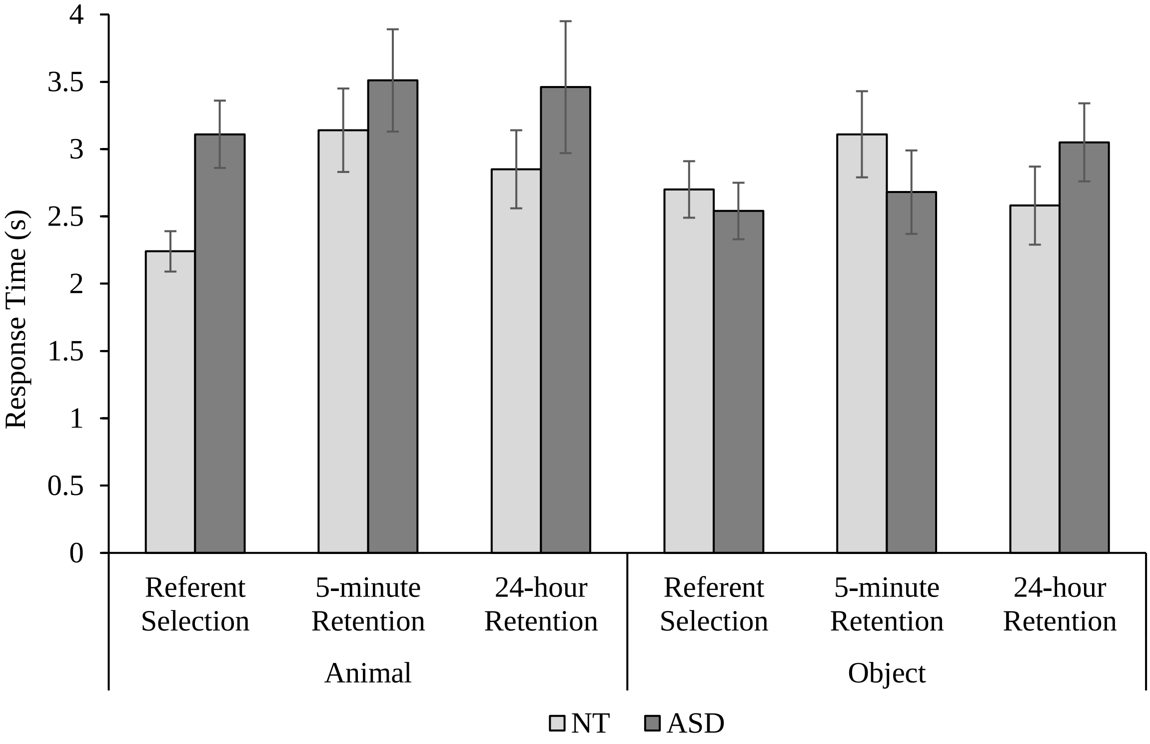
	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Referent Selection Accuracy	(Intercept)	5.59	0.81	6.88	<.001
	Trial Type	-6.97	1.41	-4.96	<.001
	Condition	4.46	0.91	4.91	<.001
	Trial Type x Condition	-8.75	1.79	-4.89	<.001
		AIC	**BIC**	**logLik**	**deviance**
		360.8	426.3	-165.4	330.8
	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Autistic Children	(Intercept)	2.21	2.00	1.10	.27
	Trial Type	-5.64	3.52	-1.60	.11
	Condition	3.91	3.26	1.20	.23
	Expressive Vocabulary	0.05	0.01	4.12	<.001
	Trial Type x Condition	-6.63	6.38	-1.04	.30
		AIC	**BIC**	**logLik**	**deviance**
		171.3	229.6	-69.6	139.3
	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Neurotypical Children	(Intercept)	0.98	1.39	0.71	.48
	Trial Type	-3.65	1.51	-2.41	.016
	Condition	1.95	1.20	1.63	.10
	Nonverbal Intelligence	0.07	0.02	2.72	.006
	Trial Type x Condition	-7.97	2.35	-3.39	<.001
		AIC	**BIC**	**logLik**	**deviance**
		183.3	242.6	-75.7	151.3

Referent selection response times

Children's response times for correctly answered referent selection trials were analysed using linear mixed-effects models, testing the effects of population, condition, and trial type. We calculated the average correct response time for each population in each trial type and condition, and removed outliers that were ≥ 3*SD* above the mean for the sub-group (e.g. autistic children in the animal condition responding to novel trials). We also removed one trial from an autistic participant who did not use their own hand to respond (they used their head). The models in these analyses included 235 of 241 (98%) correct responses from autistic children and 263 of 270 (97%) correct responses from neurotypical children. With outliers excluded, mean correct response times for each population are reported in Figure 4.

Figure 4

*Mean response times on correctly answered referent selection, 5-minute retention, and 24-hour retention trials for neurotypic*al (NT) *and autistic children (ASD), error bars show ± 1 SE*



The final model included a population x trial type interaction (t = -3.10, p = .004; see Table 3). We deconstructed this interaction by testing the effect of population for novel and familiar trials individually, and the effect of trial type on the neurotypical and autistic groups separately. Autistic children's response times did not significantly differ between familiar and novel trials (*t* = 1.60, *p* = .14), however neurotypical children were significantly quicker to respond to familiar trials than novel trials (*t* = 6.96, *p* <.001).

Table 3

Summary of the fixed effects in the final linear mixed-effects model of children's response times on correctly answered referent selection trials, predicted by trial type and population

Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
(Intercept)	2.75	0.24	11.62	<.001
Trial Type	0.95	0.18	5.17	<.001
Population	0.37	0.47	0.79	.43
Trial Type x Population	-1.13	0.36	-3.10	.004
	AIC	**BIC**	**logLik**	**deviance**
	2040.1	2107.5	-1004.0	2008.1

5-minute retention accuracy

Children’s retention accuracy after 5 minutes was analysed via generalised linear mixed-effects models testing the effects of population, condition, referent selection accuracy, by-word novel referent selection trial accuracy, and number of labels heard at referent selection for each target word. Six trials from neurotypical children and five trials from autistic children were excluded due to a technical issue. The models in these analyses contained 283 data points from autistic children and 298 data points from neurotypical children. The descriptive statistics are reported in Figure 3.

The final model included a population x referent selection accuracy interaction which approached significance (*z* = 1.81, *p* = .070; see Table 4). The interaction was deconstructed by testing the effect of referent selection accuracy on each population separately. Referent selection accuracy was significant for the autistic group (*z* = 2.01, *p* = .044), but not the neurotypical group (*z* = -0.03, *p* = .97). While autistic children responded significantly more accurately during 5-minute retention if they were previously more accurate during referent selection, referent selection accuracy did not predict 5-minute retention accuracy for neurotypical children. We must interpret this result with caution however due to the borderline significant interaction effect and model comparison (see Appendix B).

As for referent selection accuracy, we investigated how individual differences in participant characteristics influenced children's 5-minute retention accuracy. The structure of the baseline model was based on the preceding analysis, but as this analysis examined the autistic and neurotypical groups separately, the population effects were not included in the model. Since the effect of referent selection accuracy was not significant for the neurotypical group, this was also removed from the model for neurotypical children.

For the autistic group, the inclusion of animal interest as an additional fixed effect (z = -2.01, p = .045; see Table 4) predicted accuracy. Children with higher animal interest scores responded *less* accurately on 5-minute retention trials.

For the neurotypical group, the inclusion of expressive vocabulary (z = 2.99, p = .003) and receptive vocabulary (z = -2.07, p = .039; see Table 4) as additional fixed effects significantly improved model fit. Children with higher expressive vocabularies and lower receptive vocabularies were more accurate at retention after five minutes.

Table 4

Summaries of the fixed effects in the final generalised linear mixed-effects models (log odds) of children's accuracy on 5-minute retention trials, and how individual differences influence 5-minute retention trial accuracy

	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
5-Minute Retention Accuracy	(Intercept)	-0.29	0.17	-1.67	.10
	Population	0.15	0.29	0.51	.61
	Referent Selection Accuracy	0.24	0.25	0.96	.34
	Population x Referent Selection Accuracy	0.92	0.51	1.81	.07
		AIC	**BIC**	**logLik**	**deviance**
		793.5	828.4	-388.8	777.5
	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Autistic Children	(Intercept)	1.16	0.74	1.57	.12
	Referent Selection Accuracy	0.88	0.35	2.49	.013
	Animal Interest	-0.06	0.03	-2.01	.045
		AIC	**BIC**	**logLik**	**deviance**
		389.8	415.3	-187.9	375.8
Neurotypical Children	**Fixed effects**	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
	(Intercept)	-2.49	0.51	-4.86	<.001
	Expressive Vocabulary	0.08	0.03	2.99	.003
	Receptive Vocabulary	-0.04	0.02	-2.07	.039
		AIC	**BIC**	**logLik**	**deviance**
		390.4	416.3	-188.2	376.4

5-minute retention response times

Children's response times for correctly answered 5-minute retention trials were analysed using linear mixed-effects models. Outliers were identified and removed in the same way as described for referent selection trials. The models in these analyses included all 140

(100%) correct responses from autistic children and 118 of 119 (99%) correct responses from neurotypical children. With outliers excluded, mean correct response times for each population are reported in Figure 4.

The inclusion of fixed effects (population and condition) did not improve model fit.

24-hour retention accuracy

Children's retention accuracy after 24-hours was analysed via generalised linear mixed-effects models testing the effects of population, condition, by-word referent selection accuracy, accuracy on novel referent selection trials, number of labels heard at referent selection for each word, and total 5-minute retention accuracy (all coded as described previously). Two autistic children in the animal condition did not complete the 24-hour retention trials due to absence or non-compliance. Additionally, four trials from autistic children and four trials from neurotypical children were excluded due to a technical issue. The model in these analyses contained 268 data points from the autistic group and 300 data points from the neurotypical group. Descriptive statistics for 24-hour retention accuracy are presented in Figure 3.

The final model contained a population x condition interaction (z = -2.67, p = .008) and a fixed effect of total 5-minute retention accuracy (z = 3.59, p <.001; see Table 5). Children were more likely to respond correctly at 24-hour retention if they were more accurate at 5-minute retention. When deconstructing the interaction, the effect of population was significant in the object condition (z = 2.50, p = .012; autistic children responded significantly more accurately than neurotypical children), but not the animal condition (z = -1.31, p = .19; the groups did not significantly differ). Autistic children responded with significantly greater accuracy in the object condition compared to the animal condition (z = -

2.84, *p* = .005), but neurotypical children did not significantly differ in their response accuracy between conditions (*z* = 1.44, *p* = .15).

We also explored how individual differences in participant characteristics influenced children's 24-hour retention accuracy. The model for the autistic group contained 268 data points, while the model for the neurotypical group contained 300 data points. The structure of the baseline model was based on the preceding analysis, but as this analysis investigated the autistic and neurotypical groups separately, the population effects were not included in the model. Since the effect of condition was not significant for the neurotypical group, this was removed from the model for neurotypical children.

For the autistic group, the final model included an additional fixed effect of animal interest (*z* = 2.71, *p* = .007; see Table 5) alongside total 5-minute retention accuracy. Overall, children who were more interested in animals were more likely to respond correctly on 24-hour retention trials.

For the neurotypical group, the final model included additional fixed effects of non-verbal intelligence (*z* = 1.91, *p* = .056) and attention (*z* = -2.78, *p* = .006; see Table 5) alongside total 5-minute retention accuracy. Children who had higher non-verbal intelligence scores, and more typical attentional mechanisms, were more likely to respond correctly on 24-hour retention trials. The effect of total 5-minute retention accuracy reduced to marginal significance when separating the models by population, suggesting that the relationship between 5 minute and 24-hour retention was much stronger for autistic children than neurotypical children.

Table 5

Summaries of the fixed effects in the final generalised linear mixed-effects models (log odds) of children's accuracy on 24-hour retention trials, and how individual differences influence 24-hour retention trial accuracy

	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
24-hour Retention Accuracy	(Intercept)	-0.78	0.25	-3.19	.001
	Population	0.11	0.22	0.48	.63
	Condition	-0.22	0.21	-1.03	.30
	Total 5-minute Retention Accuracy	0.22	0.06	3.59	<.001
	Population x Condition	-1.14	0.43	-2.67	.008
		AIC	**BIC**	**logLik**	**deviance**
		763.9	803.0	-373.0	745.9
	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Autistic Children	(Intercept)	-2.86	0.86	-3.31	<.001
	Animal Interests	0.07	0.03	2.71	.007
	Total Accuracy at 5-minute Retention	0.32	0.09	3.35	<.001
	Condition	-0.80	0.35	-2.26	.024
		AIC	**BIC**	**logLik**	**deviance**
		356.2	384.9	-170.1	340.2
	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Neurotypical Children	(Intercept)	-0.84	0.53	-1.60	.11
	Nonverbal Intelligence	0.02	0.01	1.91	.056
	Attention	-0.06	0.02	-2.78	.006
	Total Accuracy at 5-minute Retention	0.16	0.09	1.77	.077
		AIC	**BIC**	**logLik**	**deviance**
		402.0	431.6	-193.0	386.0

24-hour retention response times

Children's response times for correctly answered 24-hour retention trials were analysed using linear mixed-effects models. Outliers were identified and removed in the same way as for previous analyses. These analyses included 141 of 143 (99%) correct responses from autistic children, and 140 of 141 (99%) correct responses from neurotypical children. With outliers excluded, mean correct response times for each population are reported in Figure 4.

The inclusion of fixed effects (population and condition) did not improve model fit.

2.6. Discussion

This study investigated whether autistic and neurotypical children can more easily identify and retain novel words associated with categories of stimuli they are interested in, compared to neutral stimuli, when presented alongside interesting distractor objects. We examined children's accuracy and correct response speed across three distinct stages of word learning: referent selection, 5-minute retention, and 24-hour retention after a period of sleep. Compared to neurotypical children matched on receptive vocabulary, autistic children did not significantly differ in their ability to disambiguate the meanings of novel words using ME or recall them after 5 minutes. Individual differences in children's animal interests affected autistic children's word learning accuracy; children who had stronger interests in animals were less accurate during 5-minute retention, but more accurate at 24-hour retention. Autistic children were more accurate to retain novel object labels than neurotypical peers during 24-hour retention, possibly due to the starker perceptual contrasts between familiar animal and novel object stimuli during encoding. During referent selection, autistic children responded at a similar pace on novel and familiar trials, whereas neurotypical children were quicker to respond on familiar compared to novel trials. Overall, autistic children's observed influence of animal interests over the longer timescale, and slower response times during encoding, may suggest that the presence of interesting non-target stimuli strengthen autistic children's encoding of novel word-referent associations by boosting their attention within learning contexts.

Our results demonstrate that children in both groups could accurately utilise ME-based referent selection to identify the meanings of novel words across conditions. This finding adds to a growing body of literature that autistic children can perform fast mapping with comparable accuracy to neurotypical children with similar receptive language abilities (e.g. Carter & Hartley, 2021; Hartley et al., 2019; Preissler & Carey, 2005; Rothwell et al., accepted, see Study 2). We found that neurotypical children responded more accurately during familiar trials than novel trials, as children could draw on existing representations of familiar words. However, autistic children did not significantly differ in their response accuracy between trial types. Both groups of children also responded more accurately in the animal condition than the object condition on familiar trials, but not novel trials. Potentially, since the perceptual contrasts between familiar and novel stimuli were more distinct in the object condition, familiar trial accuracy may be reduced due to children's attention being drawn to the more categorically salient novel object and away from requested familiar animal referents. By contrast, in the animal condition, the novel animal may have been relatively less distracting due to being categorically similar to the familiar animal stimuli.

During 5-minute retention, greater referent selection accuracy predicted more accurate retention for autistic children, but not neurotypical children. This was unexpected, as many studies often demonstrate that referent selection and short-term retention are underpinned by separable mechanisms (e.g. Carter & Hartley, 2021; Horst & Samuelson, 2008; McMurray et al., 2012; Rothwell et al., accepted, see Study 2). As such, it is possible that the mechanisms underlying referent selection and short-term retention may be less distinct for autistic children with concomitant language delays (for similar findings, see also Rothwell et al., accepted, Study 2, Study 3).

A striking finding at 5-minute retention was that autistic children with higher animal interests were *less* accurate to respond. Given the link between referent selection and 5-

minute retention accuracy, it is possible that autistic children with high animal interests paid more attention to familiar animal distractors during referent selection, processing and encoding information about them as well as the novel stimuli. Consequently, proportionally reduced focus on the targets may have had a detrimental impact on 5-minute retention accuracy for individuals with especially strong animal interests (Hilton et al., 2019). This aligns with previous research from Pomper and Saffran (2018) who discovered that neurotypical 3-year-olds were less accurate to identify and retain novel words that were taught in the presence of salient distractors, compared to non-salient distractors (see Axelsson et al., 2012 for similar results). Likewise, Horst et al. (2010) demonstrated that when neurotypical 30-month-olds learnt novel words in the presence of either two, three, or four competitor objects, children who encountered more competitors during training were significantly worse at retaining the novel word mappings (despite no influence on referent selection). These findings highlight that the increased competition from quantity or saliency of distractor stimuli experienced during training negatively affected children's short-term novel word retention. This is likely due to greater division of attention between target and distractor stimuli reducing the opportunity for novel word-referent encoding. However, in the present study, the poorer retention performance of autistic children with greater animal interests were only demonstrated at 5-minute retention (not 24-hour retention). Given the relatively small body of literature investigating retention in autistic children, further research should aim to address this theory further.

Remarkably, the animal interests of autistic children had the converse effect on their retention accuracy after a 24-hour delay compared to a 5-minute delay. After 24 hours, autistic children with greater animal interests were more likely to accurately retain novel word meanings. As animal distractors were present on all referent selection trials, autistic children with greater animal interests may have focused more intensively due to the presence of highly interesting distractors, leading to greater attention towards the visual scene and superior information encoding. Studies suggest that children require just the right amount of contextual support and variability within word learning environments to encode a rich representation of the target, a so-called ‘Goldilocks effect’ (Horst et al., 2013; Kidd et al., 2012, 2014). Too few attentional demands can be as detrimental to word learning as too many attentional demands, both meaning that encoding and retention suffer (Zosh et al., 2013). This process may be operating in the present study – when autistic children are presented with interesting stimuli, the learning environment is optimal for robust encoding which benefits later retrieval only after a period of consolidation has occurred.

After a 24-hour delay, autistic children retained novel animal words with comparable accuracy to neurotypical peers, and *more* novel object words than neurotypical children. This suggests that word learning mechanisms in autistic children remain unimpaired in both short and long-term learning phases, despite their delayed vocabulary development. Consistent with our similar studies (Rothwell et al., accepted, Study 2, Study 3), we discovered that children who responded more accurately during 5-minute retention also responded more accurately at 24-hour retention. This suggests that the mechanisms underpinning ‘short-term’ and ‘long-term’ word learning are related for both populations (McMurray et al., 2012). It is plausible that, since word acquisition gradually accumulates following multiple exposures to the novel word, correct 5-minute retention could reinforce word-referent mappings, improving subsequent 24-hour retention accuracy (McMurray et al., 2012).

The presence of a significant population x condition interaction term at 24-hour retention suggested that autistic children responded significantly more accurately than neurotypical children in the object condition (M_{ASD} = 0.62, M_{NT} = 0.43), but the groups did not significantly differ in their response accuracy in the animal condition (M_{ASD} = 0.43, M_{NT} = 0.51). This provides further insight into how autistic children’s attentional mechanisms influence their longer-term learning. When presenting novel object stimuli alongside familiar animal stimuli, the perceptual contrasts between stimuli are more distinct than when presenting novel and familiar categorically matched animal stimuli. Consequently, autistic children could have encoded more information about the novel objects due to this striking distinction, thus affording greater long-term retention of novel word-object representations. This is an important insight as it has implications for learning contexts. Potentially, the whole environment is influential for learning success, including the distractor stimuli and contrast between distractors and targets, as all interact to impact autistic children's attention and likelihood of learning.

The difference in autistic children's novel object retention accuracy between 5 minute and 24-hour retention indicates that their word learning was influenced by the facilitative effect of sleep. Numerous studies have shown that children's novel word retention improves after a nap (e.g. Axelsson et al., 2018; Horváth et al., 2015; Williams & Horst, 2014) and after a night's sleep in both autistic and neurotypical populations (Fletcher et al., 2020; Henderson et al., 2012, 2014; Rothwell et al., accepted, see Study 2). These findings can be explained by active system consolidation theory which states sleep protects new declarative memories from decay and reactivates novel words during sleep to integrate them into long term memory (Diekelmann & Born, 2010; Gais & Born, 2004). As such, after sleeping overnight, autistic children's word-object representations were consolidated for greater retrieval at 24-hour retention (more so than word-animal representations). This finding aligns with Rothwell et al. (accepted, see Study 2), where autistic children also demonstrated larger improvements in novel word retrieval between 5-minutes and 24-hours in the object condition than the animal condition. Both studies from Rothwell et al. (Study 1, Study 2, accepted) highlight that the benefits of interesting animal stimuli are relatively short lasting, and sleep consolidates the less interesting stimuli into memory for retrieval after 24-hours. Future research should examine whether the benefit of category interests and long-term retention would apply for more complex learning environments - for example, when children are required to remember more detailed, semantic information about animals or other categories of interest.

Whereas response accuracy indicates whether children successfully identified and retained word-referent pairings, how quickly children generated correct responses provides insight into the *efficiency* of children's information processing. Interestingly, at referent selection, autistic children responded at similar speed for both novel and familiar trials, whereas neurotypical children were significantly slower to respond correctly during novel trials compared to familiar trials. Neurotypical children's response times reflect the use of ME, as to decipher the correct referent, children were required to evaluate and eliminate familiar distractors, and shift their attention to the novel stimuli (Halberda, 2006). Despite this, autistic children took similar times to respond to requests for both familiar and novel stimuli, suggesting that they spent longer studying animal distractor stimuli across trial types than neurotypical children. Given that at least two thirds of the referent selection stimuli were animals pertaining to children's interests, the longer reaction times present here could be due to difficulties disengaging from stimuli that appeal to their interests (e.g. Akechi et al., 2011; Chawarska et al., 2012; Elsabbagh et al., 2009). In a similar study, Rothwell et al. (accepted, see Study 2) demonstrated that autistic children took longer to correctly identify animal stimuli that pertained to their interests due to heightened attentional allocation. It is plausible that in the present study, our findings reflect children's decision making whilst engaging in mutual exclusivity. When stimuli that appeal to children's interests are present, it may be more difficult for autistic children to ignore such stimuli, and thus may exhibit greater scrutiny of stimuli when excluding incorrect referents and deciphering the target stimuli.

Autistic children were not significantly slower than neurotypical children to generate correct responses on 5-minute or 24-hour retention trials. Previous research investigating autistic children’s response times during word learning tasks is limited and therefore the picture is mixed. Some studies find that autistic children are slower to correctly recall word meanings at 5-minute retention (e.g. Hartley et al, 2020), although other studies suggest that differences between populations are not significant for all stimulus types (e.g. Rothwell et al., accepted, Study 2, Study 3). In the present study, the presence of high interest stimuli during referent selection potentially strengthened encoding of word-referent representations, with stronger representations resulting in quicker retrieval. Previous studies focus on distractor objects that are not categories of interest (Hartley et al, 2020; Rothwell et al., accepted, Study 2, Study 3), so it is not surprising that categories of interest may enhance encoding of novel word-referents during referent selection. As such, the presence of familiar animals may provide the optimal amount of attentional division.

The knowledge gained in this study could potentially inform the development of interventions designed to scaffold autistic children’s word learning. Our findings reveal that capitalising upon the interests of children during referent selection may increase their retention of novel information after a period of sleep, as well as the speed with which these novel representations are retrieved. Whilst heightened visual attention affords more robust novel word encoding, it is possible that reduced attention to stimuli in word learning environments, due to factors such as lack of interest, may hinder vocabulary development. This finding highlights the importance of considering children’s individual interests within learning contexts. Practitioners can utilise children’s interests to craft interesting, stimulating environments to maximise autistic children’s likelihood of learning novel words, and help them pay attention to the right things at the right time. The present study adds to a growing body of literature that suggests touch-screen computers are an effective platform to teach children new words associated with different types of stimuli when distractions are minimised (e.g. Hartley et al., 2020; Rothwell et al, accepted, see Study 2). Moreover, the active nature of ME-based referent selection, where children utilise their existing knowledge to decipher the correct referent, effectively facilitated the word learning of autistic children with language delay. Consequently, these principles could be applied in clinical and educational interventions, as well as natural learning environments where possible, to scaffold the language development of autistic children with concomitant language delay.

Of course, we must address the limitations of this study. Firstly, our study set up was not reminiscent of a naturalistic word learning context – stimuli presentation was constrained, and distractions were minimised. The increased cognitive demands present in natural communicative situations could lead to reduced or incorrect visual-auditory input, and consequent language delay (McMurray et al., 2012; Venker et al., 2018; Yu & Smith, 2012). Thus, it is possible that the accurate word learning of our autistic children could diminish when tested in more naturalistic environments. As such, future research should investigate autistic children's word learning under more challenging conditions by displaying more complex and diverse arrays of stimuli and presenting stimuli more rapidly. Secondly, we chose to match our autistic and neurotypical children on receptive vocabulary, rather than chronological age (our autistic sample were significantly older) because this study aimed to compare word learning abilities across populations when delays in language development were controlled for. When matched on chronological age, studies often show that autistic children are delayed on multiple facets of language development in comparison with similarly aged neurotypical children (e.g. Charman et al., 2003). We acknowledge that our autistic sample may have attained poorer accuracy on our word learning measures than age-matched neurotypical children. Finally, although we only included participants who had pre-existing animal interests, we did not specifically recruit autistic children with a 'special interest' in animals. Special interests are preferred topics that autistic individuals attend to with a particularly intensive focus (American Psychiatric Association, 2013; Attwood, 2003;

Klin et al., 2007). It is possible that comparing neutral interest stimuli against special-interest stimuli would have afforded greater condition and population differences.

Overall, this study has advanced understanding of how autistic and neurotypical children identify and retain novel word meanings, and how these processes are impacted by interest in stimulus categories (concerning both targets and, for the first time, distractors). Despite our autistic participants' delayed language development, they responded at least as accurately as vocabulary-matched neurotypical children on measures of referent selection and 5-minute retention, and more accurately when identifying object stimuli at 24-hour retention. This suggests that the fundamental mechanisms underpinning word learning are not qualitatively atypical in autism. Generally, both autistic and neurotypical children accurately identified and retained words at a comparable pace. Since increased animal interests in autistic children negatively influenced 5-minute retention, but lead to more accurate longer-term retention, it is possible that animal distractors strengthen encoding of novel word-representations in autistic children. However, the contrast between distractor and target saliency did not appear to affect neurotypical children's novel word learning. Together, these findings highlight the potential benefits of incorporating stimuli that pertain to autistic children's categories of interests during learning events. To investigate attentional mechanisms further, we plan to examine the looking behaviours of autistic and neurotypical children during this word learning task. We will explore between-group differences in gaze during learning in each condition, to elucidate how within-trial looking behaviour predicts word learning accuracy when stimuli differ in categorical salience.

Chapter 3: How do autistic and neurotypical children's interests influence their visual attention and accuracy during novel word learning?

3.1. Chapter Introduction

Novel word acquisition is a dynamic process that is influenced by numerous factors, including lexical heuristics (Mathée-Scott et al., 2021), attention allocation (Ackermann et al., 2020; Hartley et al., 2019), and sleep consolidation (Henderson et al. 2014; Williams & Horst, 2014). Research continues to establish exactly *how* these factors affect word learning. Heightened visual attention through increased interest to stimuli has been demonstrated to result in the formation of more robust word-referent relationships that are less susceptible to decay in neurotypical children (e.g. Ackermann et al., 2020; Bion et al., 2013). However, limited research has examined this theory in autism.

Study 1 highlighted that categorically interesting distractor stimuli aided novel word retention of autistic children by increasing their attention to the whole scene at encoding. The present chapter (Study 2) employed a novel combination of touchscreen and in-trial visual attention data to investigate how category interests impact word learning when distractor stimuli do not pertain to children's interests. Specifically, neurotypical and autistic children learnt words for novel animal (high interest) and novel object (neutral interest) stimuli in the presence of familiar objects at referent selection. Examining these stimuli and methodological combinations will elucidate how children's preferential biases influence learning outcomes.

Author contribution: *Charlotte Rothwell*: study design, data collection, analysis, writing, review. *Gert Westermann*: study design, review. *Calum Hartley:* study design, review.

A version of this study reporting the accuracy and response time data has been accepted for publication in the Journal of Autism Developmental Disorders (Rothwell et al., accepted, 'How do autistic and neurotypical children's interests influence their accuracy during novel word learning?'). The in-trial looking behaviour data will be submitted for publication in a separate follow-up paper.

3.2. Abstract

Children's word learning is influenced by their visual attention to stimuli. Thus, difficulties acquiring language experienced by some autistic children may be due to differences in attention and interests, rather than atypical word learning mechanisms. Employing a novel combination of accuracy and looking measures, we investigated whether autistic children with delayed language development and neurotypical children matched on receptive vocabulary differ in accuracy and/or visual attention when learning words associated with animals (high interest stimuli) and objects (neutral interest stimuli). In a fast-mapping task, both groups identified meanings of novel words associated with unfamiliar animals and objects with comparable accuracy. After five minutes, autistic children retained animal names with greater accuracy than neurotypical children. Autistic children also showed a greater increase in their accuracy between 5 minute and 24-hour retention tests and outperformed neurotypical children after a night's sleep. However, autistic children demonstrated slower response times than neurotypical children at each word learning stage. Analyses of in-trial visual attention revealed that autistic children looked at targets significantly more often than neurotypical children at each word learning stage. Across groups, retention accuracy was predicted by visual attention and auditory input at referent selection, indicating a relationship between fast mapping and retention mechanisms. Overall, these findings indicate that differences in visual attention do not have a detrimental impact on autistic children's word learning under experimental conditions when expectations are based on receptive vocabulary. However, the observed differences in visual attention and response times signal reduced efficiency of processing audio-visual input that could impair naturalistic language acquisition.

Keywords: Word learning; Autism; Interests; Attention; Referent Selection; Retention

3.3. Introduction

Word learning is one of the most important milestones in children's cognitive development (Carpenter et al., 1998). While neurotypical children can map word-referent associations from 6 months (Friedrich & Friederici, 2011), and know the meanings of approximately 200 words before 2-years of age (Dale & Fenson, 1996), autism is often characterised by significant delays in receptive vocabulary development (Artis & Arunachalam, 2023; Kover et al., 2013). Recent studies investigating the causes of autistic children's difficulties acquiring vocabulary have demonstrated that fundamental word learning mechanisms function and inter-relate in a manner that resembles neurotypical development (Carter & Hartley, 2021; Hartley et al., 2019, 2020). Thus, it may be that autistic children's word learning difficulties can be attributed to attentional differences that affect their intake of visual and auditory input (Arunachalam & Luyster, 2018; Venker et al., 2018). Here, we directly test this theory by systematically investigating how autistic children's interests in stimuli influence multiple word learning mechanisms. Here, we directly test this theory by systematically investigating how autistic children's interest in stimuli affects multiple word learning mechanisms, exploring whether their visual attention differs during word learning, and examining how their looking behaviour impacts on learning accuracy.

When a child detects a novel word in speech, successful word learning is contingent on accurately identifying its intended meaning (referent selection; Spiegel & Halberda, 2011; Vlach & Sandhofer, 2012). The child must then store the correct word-referent association in memory for later retrieval (retention; Gleitman, 1990). According to the 'dynamic associative account', referent selection and retention utilise separate 'fast mapping' and 'slow learning' mechanisms that operate on different timescales (McMurray et al., 2012; Kucker et al., 2015; Samuelson & McMurray, 2017).

Fast mapping occurs when children overcome the challenge of referential ambiguity (there are often multiple potential targets for a newly heard word; Markman, 1989) by correctly inferring meaning from linguistic and environmental cues (Carey & Bartlett, 1978). For example, by 2 years, neurotypical children map new word-referent associations on the basis that each word has only a single referent (they employ the principle of 'mutual exclusivity' (ME); Markman, 1989; Markman & Wachtel, 1988). Children's use of ME is commonly tested by presenting an unfamiliar object amongst familiar objects and asking them to identify the referent of a novel word (Dautriche et al., 2015). As the familiar objects already have known labels, neurotypical children deduce that the unfamiliar object must be the referent for the novel word (Frank & Poulin-Dubois, 2002).

Although referent selection is an important first step towards vocabulary acquisition, children are considered to have 'learnt' a new word only when they can retrieve its meaning after a delay (Gleitman, 1990; Vlach & DeBrock, 2019). Crucially, accurate referent selection does *not* guarantee retention; Horst and Samuelson (2008) demonstrated that neurotypical toddlers who perform at ceiling on a fast-mapping task often fail to retain novel words after five minutes (also see Bion et al., 2013; Gurteen et al., 2011). While referent selection represents a process of attentional narrowing, retention is underpinned by basic associative learning mechanisms that gradually strengthen as statistical input increases (Hartley et al., 2020; McMurray et al., 2012). Newly formed word-referent associations are also strengthened by sleep. School-aged neurotypical children's novel word retention significantly improves after a night's sleep (Brown et al., 2012; Henderson et al., 2012; Ma et al., 2022), and preschool children who nap shortly after exposure to novel words are more likely to retain their meanings (Horváth et al., 2015; Williams & Horst, 2014). These effects are explained by 'active system consolidation theory', which proposes that sleep enhances retention by reactivating recently encoded word-referent representations, facilitating their integration into memory networks by strengthening synaptic connections (Diekelmann & Born, 2010; Gais & Born, 2004).

Importantly, children’s word learning is intrinsically related to visual attention. During fast mapping, children must focus their attention on a novel word’s intended referent while excluding non-target competitors (Twomey et al., 2016; Zosh et al., 2013). This requires children to navigate their attention across multiple components of the learning environment *and* coordinate their attention to corresponding audio-visual stimuli during naming events (Samuelson et al., 2017). Children’s attention during referent selection then influences their subsequent retention of novel words (Hilton et al., 2019; Hilton & Westermann, 2017). Ackermann et al. (2020) recently reported that neurotypical 30-month-olds find it easier to learn names for novel referents belonging to categories they are particularly interested in, such as animals. This finding suggests that heightened attention to interesting objects increases children’s focus, which in turn benefits their encoding of word-referent representations. Direct manipulation of children’s attention via external cues can also impact neurotypical children’s novel word retention. For example, Axelsson et al. (2012) found that illuminating the target of a novel word after neurotypical children had selected a referent elicited superior retention of label-object pairings. Conversely, inattentiveness has been associated with reduced retention of words. Smith and Yu (2013) showed that neurotypical children who do not coordinate their looking behaviour with a social partner, and thus do not focus on intended referents for a sustained period of time, often fail to retain their corresponding labels. Together, these studies demonstrate a robust positive relationship between visual attention during exposure to novel word-referent associations and likelihood of successful learning (Bion et al., 2013; Spiegel & Halberda, 2011).

Early studies investigating autistic children’s referent selection identified reduced sensitivity to social-pragmatic cues as a potential cause of their language learning difficulties (e.g. Baron-Cohen et al.,1997; Preissler & Carey, 2005). However, an array of recent studies have demonstrated that autistic children with varying language abilities can successfully utilise social cues to inform accurate referent selection (Bean Ellawadi & McGregor, 2016; Hani et al., 2013; Luyster & Lord, 2009; McGregor et al., 2013). Furthermore, autistic children – including those with delayed language development – can accurately identify novel word meanings via lexical heuristics such as ME (de Marchena et al., 2011; Parish-Morris et al., 2007; Preissler & Carey, 2005).

In contrast to referent selection, few studies have investigated retention of newly learned words in autistic children with delayed language development. In two recent exceptions, Hartley et al. (2019, 2020) studied the relationship between identification and retention of novel word meanings and explored how these processes are influenced by attentional cues. In their 2019 paper, language delayed autistic children and neurotypical children (ASD *M* age = ~8 years; NT *M* age = ~5 years) matched on receptive vocabulary (NT *M* = ~5 years; ASD *M* = ~5 years) identified the names of novel objects in a ME-based fast-mapping task. After a 5-minute delay, autistic children responded at least as accurately as neurotypical children on a retention test. In Hartley et al. (2020), similar samples disambiguated word meanings by tracking statistical word-object co-occurrences with equivalent accuracy and the groups did not differ on retention tests. However, autistic children were significantly slower to identify correct referents under both cued and non-cued learning conditions. These findings suggest that fundamental mechanisms supporting word learning, and the relationships between them, may not be qualitatively atypical in language delayed autistic children. Rather, differences in response time may indicate that autism impacts the speed at which children process stimuli during word learning (Arunachalam & Luyster, 2018; Hartley et al., 2019, 2020; Tenenbaum et al., 2017).

Whereas neurotypical children can flexibly navigate attention across their environment, many autistic children have difficulties allocating sustained/selective attention and shifting focus between stimuli (e.g. Courchesne et al., 1994; Noterdaeme et al., 2002; Ozonoff et al., 1994). These differences in visual attention have been linked to domain-general deficits in executive functioning (Burgess et al., 2006; Ozonoff et al., 2004), which in turn have been implicated as a potential cause of diagnosis-defining restricted and repetitive behaviours and interests (RRBIs; Honey et al., 2012; Kanner, 1943; Richler et al., 2007). RRBIs result in children focusing intensely and repeatedly on very specific interests and activities in their daily lives. Such is the intensity of their RRBIs, many autistic children experience difficulty disengaging from preferred stimulus categories and may be reluctant to attend to stimuli that they find less interesting (Bryson et al., 1990; Leekam et al., 2011). Since environmental input is carefully selected and restricted by the child's interests, attentional focal points are narrowed (Elsabbagh et al., 2009, 2013; Landry & Bryson, 2004; Sacrey et al., 2014) and sensitivity to valuable information and informants in the environment may be suppressed (McGregor et al., 2013).

During word learning, differences in visual attention may prevent autistic children from attending to all stimuli in an array, thus affecting the accuracy and/or strength of newly encoded word meanings (e.g. Hartley et al., 2019; Hilton & Westermann, 2017). Many autistic children experience 'sticky' attentional fixations, and their focus is often captured by salient perceptual features to an atypical degree (Hartley & Allen, 2014; Pierce et al., 2011, 2016). For instance, Venker et al. (2022) recently reported that autistic children's novel referent selection was disrupted more than their neurotypical peers when stimuli were of high perceptual salience. This aligns with previous looking time studies showing that autistic children are slower, or unable, to disengage their attention from irrelevant stimuli (Elsabbagh et al., 2009, 2013; Sacrey et al., 2014). These differences in visual attention could have profound implications for language acquisition – if an autistic child hears a novel word whilst attending to a non-target object that has captured their attention, the misalignment in auditory-visual statistics may inhibit accurate referent selection and potentially generate a spurious association between the word and non-target object (Baron-Cohen et al., 1997; Tenenbaum et al., 2017; Venker et al., 2018).

Regarding retention, differences in autistic children's visual attention could either hinder or facilitate encoding and consolidation of word-referent associations under different circumstances. On one hand, if to-be-learned stimuli do not align with autistic children's interests, reduced attention may result in weak or incorrect representations of word-referent relationships (e.g. Tenenbaum et al., 2014, 2017). Alternatively, if stimuli appeal to their interests, heightened attentional focus could lead to the formation of more robust word-referent relationships that are less susceptible to decay (e.g. Ackermann et al., 2020). Such differences in encoding strength may be evident in contrasting patterns of autistic children's visual attention to stimuli they find more or less interesting during referent selection, and in predictive relationships between looking behaviour and retention accuracy. However, to our knowledge, no research to date has investigated whether autistic children's interests in stimuli influence inter-relationships between visual attention and direct measures of referent selection and retention.

For the first time, the present study researched how interests associated with specific categories of stimuli and differences in visual attention affects multiple word learning mechanisms in autistic children with delayed language development. Autistic children and neurotypical children matched on receptive vocabulary identified the meanings of novel words in a computer-based ME referent selection task with two within-subjects conditions. In one condition, children learned the names for relatively interesting stimuli – unfamiliar animals (participants' interest in animals was confirmed via a questionnaire). It is well-documented that children generally prefer animal stimuli over non-animal stimuli (Ackermann et al., 2020; Celani, 2002; Prothmann et al., 2009) and many autistic individuals are particularly fond of animals (Martin & Farnum, 2002). In another condition, children learnt the names for unfamiliar objects – generic experimental stimuli that are less likely to align with children's pre-existing interests. Retention of novel words was tested after 5 minutes and 24 hours. The retention tests following a 24-hour delay allowed us to investigate a) the robustness of novel word representations relating to different categories, and b) how sleep influences lexical consolidation in autistic children with concomitant language delay. Autism is often characterised by problematic sleep disorders, including bedtime resistance, sleep anxiety, difficulties falling asleep, and parasomnia (Díaz-Román et al., 2018; Souders et al., 2009). Given that sleep plays a critical role in protecting newly acquired declarative memories against decay in neurotypical development (Axelsson et al., 2018), such difficulties could impact autistic children's consolidation of recently mapped word-referent associations. Although previous studies have identified benefits of sleep for autistic children's lexical retention, these have exclusively recruited intellectually able participants with high IQs who do not have language-learning difficulties (e.g. Fletcher et al., 2020; Henderson et al., 2014; Norbury et al., 2010). Therefore, this study is the first to test whether overnight memory consolidation of novel words differs for autistic children with delayed language development.

Crucially, we recorded children's looking behaviour during referent selection and both retention phases, enabling us to directly compare the populations' visual engagement at each stage of the task and assess whether variability predicted learning accuracy. While previous studies focus on looking time alone, often inferring longer looking to be indicative of more accurate learning (e.g. Potrzeba et al., 2015; Tenenbaum et al., 2014, 2017), this study is the first to examine how multiple measures of visual attention and direct behavioural accuracy measures inter-relate. This novel methodological approach will yield important insights into how autism affects children's intake of stimuli during word learning *and* allow us to draw comparisons between accuracy and looking measures.

As numerous studies have shown that autistic and neurotypical children spontaneously apply ME when fast mapping (e.g. Carter & Hartley, 2021; Hartley et al., 2019), we did not necessarily expect any between-population or between-condition differences in accuracy during referent selection. Based on evidence for positive relationships between attentional focus and word learning (Ackermann et al., 2020; Axelsson et al., 2012; Bion et al., 2013; Yu & Smith, 2012), we predicted that children in both populations would retain names for unfamiliar animals with greater accuracy than names for unfamiliar objects. However, we anticipated that effects of stimuli may be greater for autistic children due to more substantial differences in visual attention between conditions. After 24 hours, we tentatively predicted that sleep-induced benefits for retention would be weaker for autistic children with delayed language development than neurotypical controls. We also anticipated that autistic children would be slower to generate correct responses than their neurotypical peers across all word learning stages, indicating less efficient processing of audio-visual input (e.g. Hartley et al., 2020).

During referent selection, we expected autistic children to spend longer looking towards novel items regardless of whether they were intended targets – particularly in the animal condition – due to difficulties disengaging attention from interesting stimuli (Elsabbagh et al., 2009, 2013; Landry & Bryson, 2004; Sacrey et al., 2014). We also anticipated that autistic children might make more frequent looks to target stimuli due to greater uncertainty and/or longer processing times required to generate correct responses. We predicted that increased visual attention to targets would predict response accuracy across conditions, groups, and task stages, and that increased attention at referent selection would be associated with superior retention. However, we were mindful of the possibility that between-population differences in visual attention may not necessarily translate to significant differences in learning outcomes. Indeed, it is well-documented that autistic children and adults can generate correct responses on a range of cognitive tasks despite differences in processing or the use of atypical strategies (Happé, 1995; Norbury et al., 2010). Importantly, this research will advance theoretical understanding of word learning by disentangling the influence of preferential biases to selective stimuli and visual attention in both autism and neurotypical development.

3.4. Method

Participants

Participants were 15 autistic children (13 males, 2 females; *M* age = 91.87 months; *SD* = 21.30) recruited from specialist schools, and 16 neurotypical children (6 males, 10 females; *M* age = 52.31 months; *SD* = 18.88), recruited from mainstream schools, nurseries, and Lancaster University BabyLab (see Table 1). All participants were monolingual, English was their native language, and had normal or corrected-to-normal colour vision. Autistic children were previously diagnosed by a qualified educational or clinical psychologist, using standardised instruments and expert judgement (i.e. Autism Diagnostic Observation Scale and Autism Diagnostic Interview – Revised; Lord et al., 1994, 2002). Diagnoses were confirmed via the Childhood Autism Rating Scale 2 (CARS; autistic *M* = 34.70, *SD* = 10.23; neurotypical *M* = 16.78, *SD* = 2.56; Schopler et al., 2010). This measure was usually completed by class teachers, but for eight neurotypical children who were tested at our BabyLab due to COVID-19 restrictions, it was completed by caregivers. Autistic children were significantly older, $t(29) = -5.48$, $p < .001$, $d = 1.97$, and had significantly higher CARS scores, $t(29) = -6.79$, $p < .001$, $d = 2.40$, than the neurotypical children.

Groups did not significantly differ on receptive vocabulary as measured by the British Picture Vocabulary Scale 2 (BPVS; autistic *M* age equivalent = 53.27 months, *SD* = 22.48; neurotypical *M* age equivalent = 60.31, *SD* = 27.44; Dunn et al., 1997), *t*(29) = 0.78, *p* = .44. Receptive vocabulary was selected as our group matching criterion as it reflects children's ability to learn word-referent relationships (Bion et al., 2013). Expressive vocabulary was measured using the Expressive Vocabulary Test 2 (EVT; Williams, 2007), or the expressive language module of the Mullen's Scales of Early Learning (MSEL; Mullen, 1995) for children who scored below the baseline on the EVT. Autistic (*M* age equivalent = 48.47 months, *SD* = 27.70) and neurotypical children (*M* age equivalent = 60.31 months, *SD* = 22.76) did not significantly differ on expressive vocabulary, *t*(29) = 1.30, *p* = .20.

Children's non-verbal intellectual abilities were measured using the Leiter-3 (Roid et al., 2013). The neurotypical group's average non-verbal IQ score (*M* = 101.38, *SD* = 7.84) was significantly higher than the autistic group's (*M* = 77.67, *SD* = 11.73), *t*(23) = 5.99, *p* <.001, *d* = 2.38. Scaled IQ scores could not be calculated for three neurotypical children as they were below the age of three years. However, the groups' raw scores on the Leiter-3 did not significantly differ (autistic *M* = 60.33, *SD* = 15.57; neurotypical *M* = 57.25, *SD* = 17.93), *t*(26) = -0.48, *p* = .64, suggesting that their non-verbal cognitive abilities were similar at the time of testing (when age was not considered). Three autistic children did not complete the Leiter-3 due to school closure during the COVID-19 pandemic, but they were retained in the study as they completed all other measures. To assess attentional behaviours, the Conner's Teacher Rating Scale (CTRS-15; Pupura & Lonigan, 2009) was completed by children's class teachers, or the caregivers of the eight neurotypical children who were tested in our BabyLab. The mean raw scores for the autistic children (*M* = 17.27, *SD* = 11.04) and neurotypical children (*M* = 12.25, *SD* = 6.03) did not significantly differ, *t*(29) = -1.58, *p* = .12. The Repetitive Behaviour Questionnaire was completed by participants' caregivers to assess the extent of their restrictive and repetitive behaviours (RBQ; Leekam et al., 2007). Autistic children (*M* = 43.87*, SD* = 8.37) had significantly higher scores than neurotypical children (*M* = 27.00, *SD* = 5.80), *t*(29) = -6.56, *p* <.001, *d* = 2.34.

Finally, caregivers completed a bespoke questionnaire assessing the extent to which children were interested in animals (min-max scores: 0-34; autistic *M* score = 23.93, *SD* = 5.55, neurotypical *M* score = 23.31, *SD* = 2.80; see Appendix A). The purpose of this measure was to ensure that we recruited participants who were interested in animals, validating our categorisation of stimuli in the animal condition as 'high interest'. One autistic child was excluded from the study due to their lack of interest in animals. The groups did not differ significantly on this measure, *t*(29) = -0.40, *p* = .69.

An additional four participants were excluded from the study; one neurotypical participant who was unable to complete the touch-screen task, one neurotypical participant who scored above the 'low to minimal symptoms' threshold on the CARS-2, and two autistic participants who did not complete both experimental conditions due to school closures during the COVID-19 pandemic.

All procedures performed in this study involving human participants were in accordance with the ethical standards of institutional and national research committees. Informed consent was obtained from caregivers prior to children's participation and a debrief was provided after participation.

Table 1

Characteristics of autistic and neurotypical Participants (SD and Ranges in Parentheses)

Group	N	Gender	Chron. Age (*M* months)	BPVS age equiv. (*M* months)	Express. Lang. age equiv. (*M* months)	CARS raw score (*M*)	Leiter-3 raw score (*M*)	CTRS raw score (*M*)	RBQ raw score (*M*)	Animal Interest score (*M*)
NT	16	6 males, 10 females	52.31 (18.88; 27-94)	60.31 (27.44; 36-118)	60.31 (22.76; 35-104)	16.78 (2.56; 15-24)	57.25 (17.93; 40-95)	12.25 (6.03; 2-26)	27.00 (5.80; 20-35)	23.31 (2.80; 19-29)
ASD	15	13 males, 2 females	91.87 (21.30; 67-136)	53.27 (22.48; 24-97)	48.47 (27.70; 5-82)	34.70 (10.23; 20-52)	60.33 (15.57; 38-83)	17.27 (11.04; 5-36)	43.87 (8.37; 30-59)	23.93 (5.55; 17-34)
Group comparison *t*-test (*p*)			<.001	.44	.20	<.001	.64	.12	<.001	.69

Note. NT: neurotypical; ASD: autism spectrum disorder; BPVS: British Picture Vocabulary Scale, CARS: Childhood Autism Rating Scale, CTRS: Conner's Teacher Rating Scale, RBQ: Repetitive Behaviour Questionnaire.

Materials

The study was administered via a touch-screen computer running MATLAB. Audio stimuli for the word learning task included eight two-syllable unfamiliar words (manu, tanzer, boskot, virdex, toma, fiffin, chatten, modi) selected from the NOUN database (Horst & Hout, 2016) and other academic sources. Visual stimuli included high-resolution colour photographs of 4 unfamiliar objects, 4 unfamiliar animals (see Figure 1), and 22 familiar objects, all presented on a grey background. All photographs were approximately 6cm^2 and 500 x 500 pixels when displayed on the screen. Unfamiliar objects were selected on the basis that children would not know their linguistic labels. Familiar objects were selected on the basis that most children understand their linguistic labels by 16 months (Fenson et al., 1994).

Three pictures of six familiar objects were employed in the warm-up trials (tree, door, light, slide, pram, top). Pictures of 16 familiar objects were presented during referent selection trials in the object condition and animal condition. These were divided into two sets and counterbalanced across conditions (1. bottle, hat, pillow, toothbrush, rock, balloon, truck, bath; 2. telephone, ball, chair, spoon, bed, window, fridge, towel). Familiar objects allocated to the two conditions were matched on mean comprehension age (13.5 months for both sets) and frequency of objects belonging to particular categories (e.g. toys, furniture). Familiar objects within each set were divided into pairs and presented alongside an unfamiliar object or unfamiliar animal in referent selection trials (depending on condition). In every trial type, three pictures were presented side by side. We ensured that names of stimuli presented together were phonologically distinct and their images clearly contrasted in shape and colour.

Stimuli names were recorded by a female speaker from the local area and presented through the computer's integrated speakers. Audio files were recorded and edited using a Sony ECM-MS907 Digital Microphone and Audacity 2.2.2 software. Auditory stimuli were edited for timing and clarity, and the volume of all files was normalised. The carrier phrases (e.g. "Can you see the [label]?", "Touch the [label]!") and the labels (e.g. "tree", "fiffin") were edited separately, so they were all distinct files. However, when the MATLAB programme ran the experiment, the audio files were presented sequentially. This was to ensure that there were no differences in the carrier phrases that may offer a hint to children regarding the labels that were about to be presented.

Figure 1

Sets of unfamiliar objects and animals used in the word learning task

Three web cameras attached to the left, right, and centre of the frame were used to record participants' visual attention and behaviour during the study. Recording was done using the 'Open Broadcaster Software' version 23.2.1, which allowed recording from all three cameras simultaneously. The cameras positioned to the left and right of the frame were 15-megapixel Logitech C920 HD Pro Webcams and recorded at a rate of 30 frames-per-second. The centre camera was built into the iMac (1.2 megapixels) and also recorded at 30 frames-per-second. The red recording lights were obscured from participants using black tape to avoid distraction.

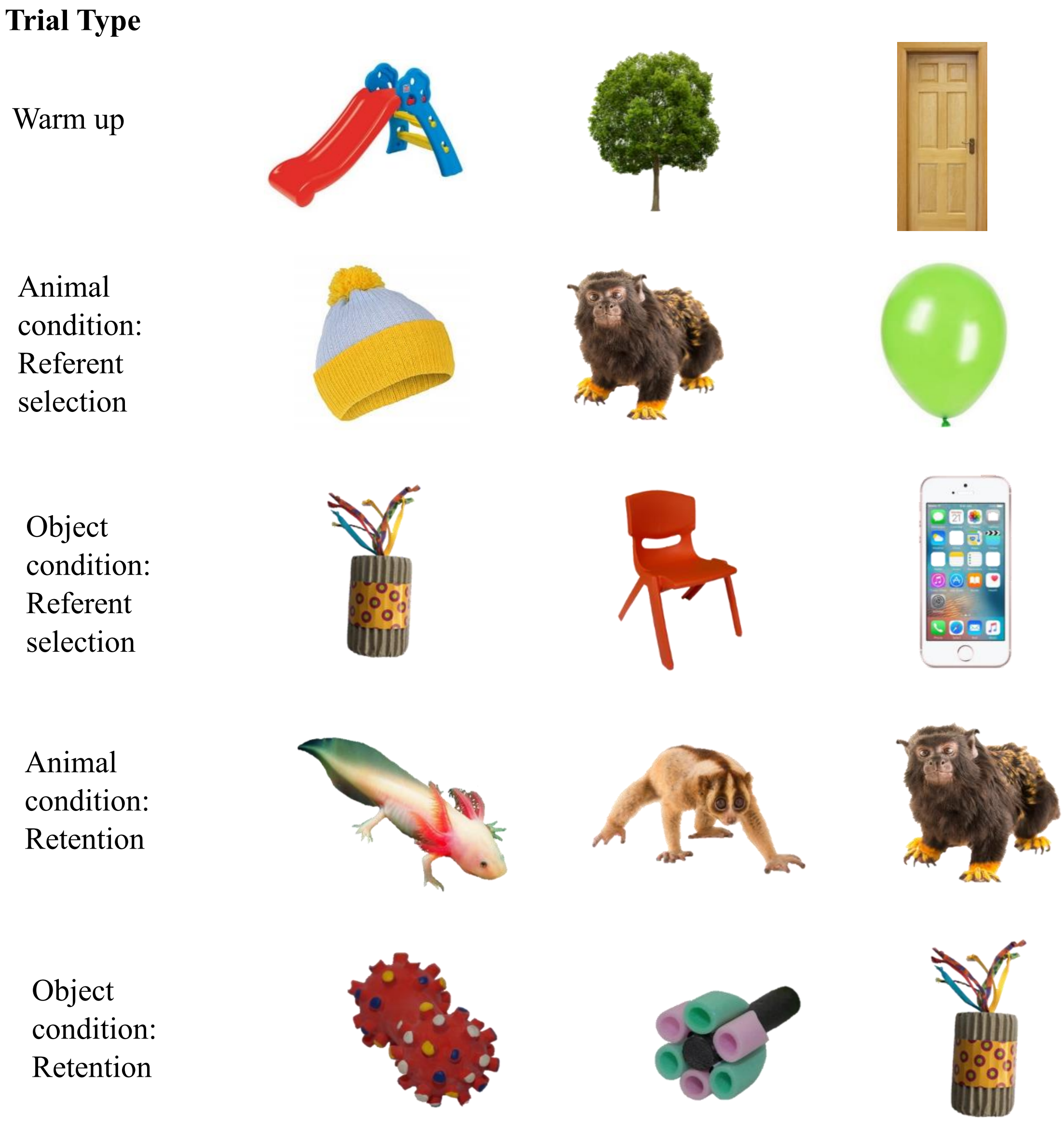
Procedure

The task in the present study was very similar to that in Study 1. During recruitment, caregivers completed a questionnaire about their child's interest in animals (see Appendix A). Animals are a common interest of many autistic and neurotypical children (Martin & Farnum, 2002; Prothmann et al., 2009), and our objective was to explore how this interest would influence their relative performance in the two word learning conditions. Examples of questions included: 'How much does your child like animals?' (responses: 1 - they don't mind animals, 2 - they like animals a little, 3 - they like animals a lot, 4 - they really, really like animals) and ‘How much does your child enjoy watching television programmes, videos, and films involving realistic animals?’ (responses: 1 - they don’t particularly enjoy it, 2 - they enjoy it a little, 3 - they enjoy it a lot, 4 - they really, really enjoy it).

Participants were tested individually in their own school or nursery, or our BabyLab, and were accompanied by a familiar adult when required. Children were assessed using the Leiter-3, BPVS, and EVT or MSEL by the researcher over multiple sessions on different days. Children completed two within-subjects conditions of the word learning task – novel animals and novel objects – administered on different days (average of six days apart, order counterbalanced). The word learning task was delivered via a customised touch-screen computer. Children were seated approximately 50-70 cm away from the screen on a height-adjustable chair. The word learning task consisted of the following stages, presented in a fixed order: 1. Warm-up trials, 2. Referent selection trials, 3. Five-minute delay, 4. Retention trials, 5. 24-hour delay, 6. Retention trials (see Figure 2). The experimenter sat quietly while the participant was engaged in tasks and offered verbal praise for attention and good behaviour.

Figure 2

Examples of trial types in the word learning task



Warm up trials

Before the study started, children were presented with a cartoon image of a hand that appeared in each of three touch-screen panels, one by one. To encourage children to feel comfortable touching the screen, the experimenter asked them to "Put their hand on the picture." Then, children completed three warm-up trials. Children were instructed to "Put your hand on the picture that the computer asks you to." During warm-up trials, children were presented with images of three familiar objects in the left, middle, and right sections of the touchscreen. After 2 seconds, participants heard "Look, '*2 s gap*' [label]!", '*1 s gap*', "Can you see the [label]?" '*1 s gap*', "Touch the [label]!". Children then had 12 seconds to respond. The same instructions played up to six times if children did not respond. Responses were accepted only after the first label utterance, preventing children from skipping through trials without hearing the requested labels. Children received feedback when they made their selection; either audio praise if they responded accurately (e.g. "Well done, you touched the [label]!") or corrective feedback if they responded inaccurately ("Actually, this is the [label]. Can you touch the [label]?"). Following inaccurate responses, the correct referent was highlighted by a green border and children could retry up to five times. Children in both groups responded significantly above chance levels at first attempt on the warm-up trials ($M_{NT} = 0.95$, $M_{ASD} = 0.90$) demonstrating that they were familiar with the requested referents and understood the task requirements. The location and order of requested objects were counterbalanced across participants.

After the warm-up trials, children were video recorded to measure their visual attention. To assist with coding, LEDs on the three video cameras flashed to signify the start of the experiment, transitions between trials, and when participants touched the screen. However, the LEDs were invisible to participants as they were covered with black tape.

Referent selection trials

After the warm-up trials, children completed eight referent selection trials. These followed exactly the same format, except children did not receive feedback following their responses. Four novel words were taught via a fast-mapping paradigm based on Horst and Samuelson (2008). Children viewed four sets of pictures (each containing one unfamiliar picture and two familiar pictures). Each set was presented twice; on one trial the novel picture was requested (novel name trial: “Look, modi! Can you see the modi? Touch the modi!”), and on another trial a familiar picture was requested (familiar name trial: “Look, ball! Can you see the ball? Touch the ball!”). Familiar trials were included to detect whether participants’ responses were biased by a preference for novelty and to encourage them to examine every item in each array (accurate fast mapping requires children to attend to familiar competitors in order to exclude them as referents for a novel word; Halberda, 2003). Novel name trials promoted active learning of new word-referent pairings; since participants already knew labels for the familiar pictures, they could identify the referent of the novel label by applying the ME principle. During this stage, familiar stimuli were always objects, and novel stimuli were either animals (high interest) or objects (neutral interest), condition dependent.

Trial order was pseudo-randomised with the constraints that the same set of pictures, or the same trial type (familiar name or novel name), was not presented on more than two trials sequentially. Positioning of stimuli on the screen (left, middle, right) was pseudo-randomised across trials with the constraint that the target did not appear in the same location more than twice consecutively. The eight novel words were divided into two sets (1. manu, tanzer, boskot, virdex; 2. toma, fiffin, chatten, modi) and were counterbalanced across conditions. Novel words were pseudo-randomly allocated to novel referents, so different novel words represented different novel referents across participants. Familiar stimuli were divided into two sets of eight to obtain a degree of control, but these were also counterbalanced across conditions.

5-minute delay

Immediately after referent selection, children engaged in an unrelated task for five minutes (e.g. colouring or building with blocks). None of the familiar or unfamiliar experimental stimuli were visible during this stage.

Retention trials

Following the five-minute delay, children completed one warm-up trial to re-engage their attention (exactly as described above). Eight retention trials immediately followed (see Figure 2 for an illustration of each trial type). Each novel word was tested on two retention trials. These trials enabled us to assess whether children's retention of newly mapped word-referent associations differed between high interest (animals) and neutral interest (objects) stimuli. Trial order was pseudo-randomised, ensuring that the same set of stimuli was never presented on more than two trials sequentially. Positioning of stimuli on the screen (left, middle, right) was pseudo-randomised across trials with the constraint that the target did not appear in the same location more than twice consecutively. Each picture served as a target on two trials and as a foil on four trials.

24-hour retention trials

After a 24-hour delay, children completed a second block of eight retention trials. Due to practical constraints, not all children experienced exactly a 24-hour delay (*M* delay = 23.8 hours, range: 20.5 – 25.6 hours). These retention trials were preceded by three warm-up trials (as described above) to remind children of the task requirements and how to respond. The 24-hour retention trials were identical to the 5-minute retention trials with the exception that stimuli were presented in different orders and combinations.

Coding and data cleaning

Videos were coded using the software Blender 2.78, with a customised version of the python script ultra-coder added on (see https://github.com/dmbasso/misc-blender-tools/blob/master/ultra_coder.py for original). Coders were blind to the location of the target stimuli on each trial. Children's visual fixations were coded frame-by-frame with a precision of 16.7ms, and looks were coded as left, right, centre, away, or not visible. The LEDs that flashed to signify the beginning of the experiment and transitions between trials, as well as participant touches, were vital for coding. Of the 184 videos recorded across the three experimental stages, 25% were reviewed by two independent coders. Coders agreed on looking direction on 98% of frames. A custom Matlab programme then calculated the primary dependent variables (see Table 2). These variables were calculated 233ms after the label onset to allow for saccade initiation latencies (Swingley et al., 1999; Swingley, 2009).

We analysed three distinct looking measures from our coded videos. 'Proportion of looking towards the target' was selected as this measure is commonly used in both the neurotypical and autism word learning literatures (e.g. Ackermann et al., 2020; Akechi et al., 2011, 2013; Bion et al., 2013; Potrzeba et al., 2015). It captures looking behaviour to the target relative to the distractor and is often interpreted as a measure of learning accuracy in the absence of an explicit behavioural response. 'Number of looks towards the target' was selected as a measure of children's exploration of stimuli, which is known to be atypical in autism (e.g. Sasson et al., 2008, 2011). A measure of longest look was selected to provide a categorical indication of children's attentional preference based on one look alone, which is more comparable to behavioural accuracy (Ambridge & Rowland, 2013), and accounts for children's attention likely decreasing over the trial (Bailey & Plunkett, 2002; Ma et al., 2011; Schafer & Plunkett, 1998). 'Longest look to novel or familiar stimuli' was coded at referent selection to investigate how word learning was influenced by attentional allocation to novel stimuli during encoding, and 'Longest look to target or foil' was coded at retention to disentangle whether children allocated their longest look to correctly encoded stimuli.

Table 2

Looking measures recorded during the experimental tasks

Looking Measure	Coding	Experimental Phase	Outcome
Proportion of time spent looking at the target stimuli	0 to 1	Referent selection 5-minute retention 24-hour retention	Proportion of time spent looking at target stimuli, compared to the proportion of time spent looking at distractor stimuli.
Number of looks to the target stimuli	0-14	Referent selection 5-minute retention 24-hour retention	Every new look towards target stimuli was counted as a new look.
Longest look to novel or familiar stimuli	0 = familiar 1 = novel	Referent selection	If the longest look was to novel stimuli or familiar stimuli.
Longest look to target or foil stimuli	0 = foil 1 = target	5-minute retention 24-hour retention	If the longest look was to target stimuli or foil stimuli.

3.5. Results and Discussions

Due to the variety of measures and the comprehensive nature of our analyses, we present Results and Discussions in two sections. The first section analyses and interprets data concerning children's word learning accuracy and response times. The second section compares the groups' visual attention at each stage of the word learning task and assesses whether variability predicted learning accuracy.

3.5.1. Do autistic and neurotypical children differ in accuracy and response times when learning names for high interest and neutral interest stimuli?

Accuracy and response time data were analysed via mixed-effects models using the glmer and lmer functions from the lme4 package in R (Bates et al., 2015). Population was contrast coded as -0.5 (neurotypical) and 0.5 (autistic). Condition was coded as -0.5 (novel object) and 0.5 (novel animal). Trial type was coded as -0.5 (familiar) and 0.5 (novel). By-word referent selection accuracy was coded as -0.5 (incorrect) and 0.5 (correct) when included as a fixed effect in retention accuracy analyses. Total accuracy at referent selection for novel trials was coded as 0-4. Number of labels heard at referent selection for each novel word was coded as 1-7 (autistic $M = 2.27$, $SD = 1.00$; neurotypical $M = 1.84$, $SD = 0.89$). Total accuracy at 5-minute retention was coded as 0-8. Trial-level accuracy as a dependent measure was coded as 1 (correct) or 0 (incorrect) for all analyses.

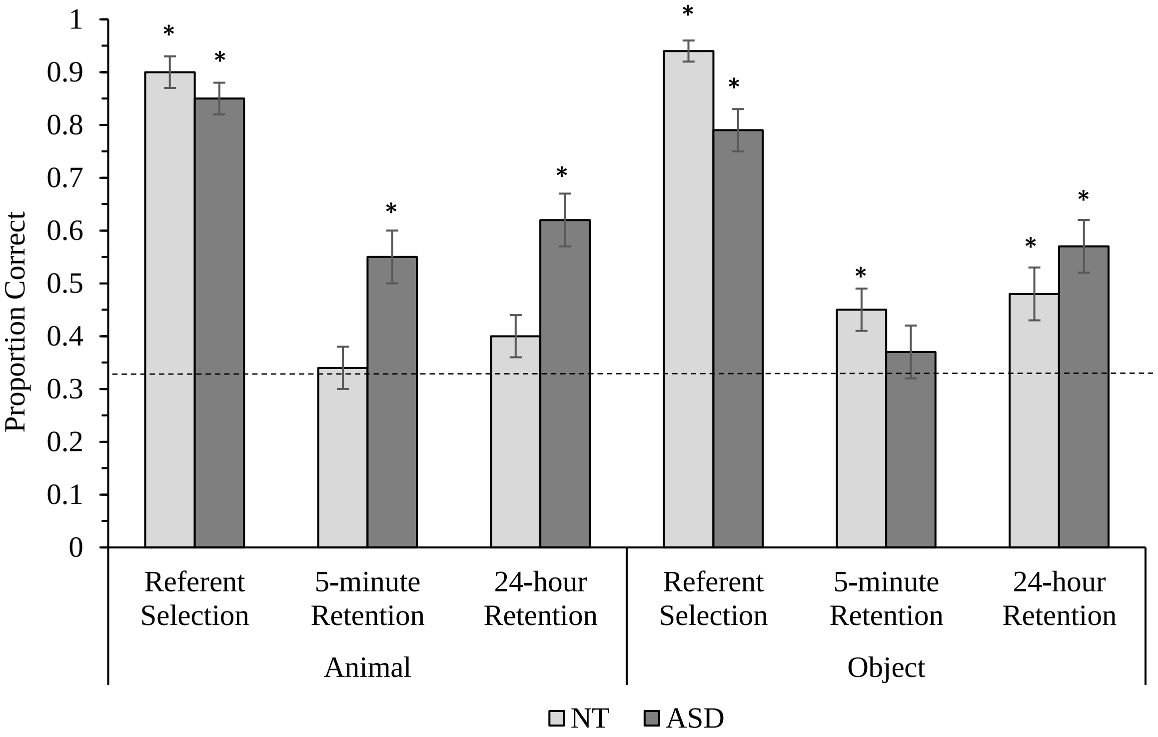
The likelihood of children responding correctly by chance on each trial was 33%. All models were built up sequentially, adding in one fixed effect at a time and comparing each model with the previous best-fitting model using log-likelihood tests. Each analysis started with a baseline model containing by-participant and by-word random intercepts, with a random slope of condition x trial type per participant for referent selection, or condition per participant for retention phases. If some models in a sequence failed to converge, the random effects were simplified until all models in the sequence successfully converged. Only final models are reported; please refer to Appendix C for full details of the model building sequences and analyses of individual differences.

Referent selection accuracy

Referent selection accuracy was analysed via generalised linear mixed-effects models testing the effects of population, condition, and trial type. Five trials were excluded from autistic participants who simultaneously responded to different locations with their head and hands. This analysis contained 491 data points. Descriptive statistics for referent selection accuracy are presented in Figure 3.

Figure 3

*Mean referent selection, 5-minute retention, and 24-hour retention trial accuracy for neurotypical (NT) and autistic children (ASD), error bars show ± 1 SE. Stars above columns indicate where performance was significantly different from chance, indicated by the dotted lines (*p <.05)*



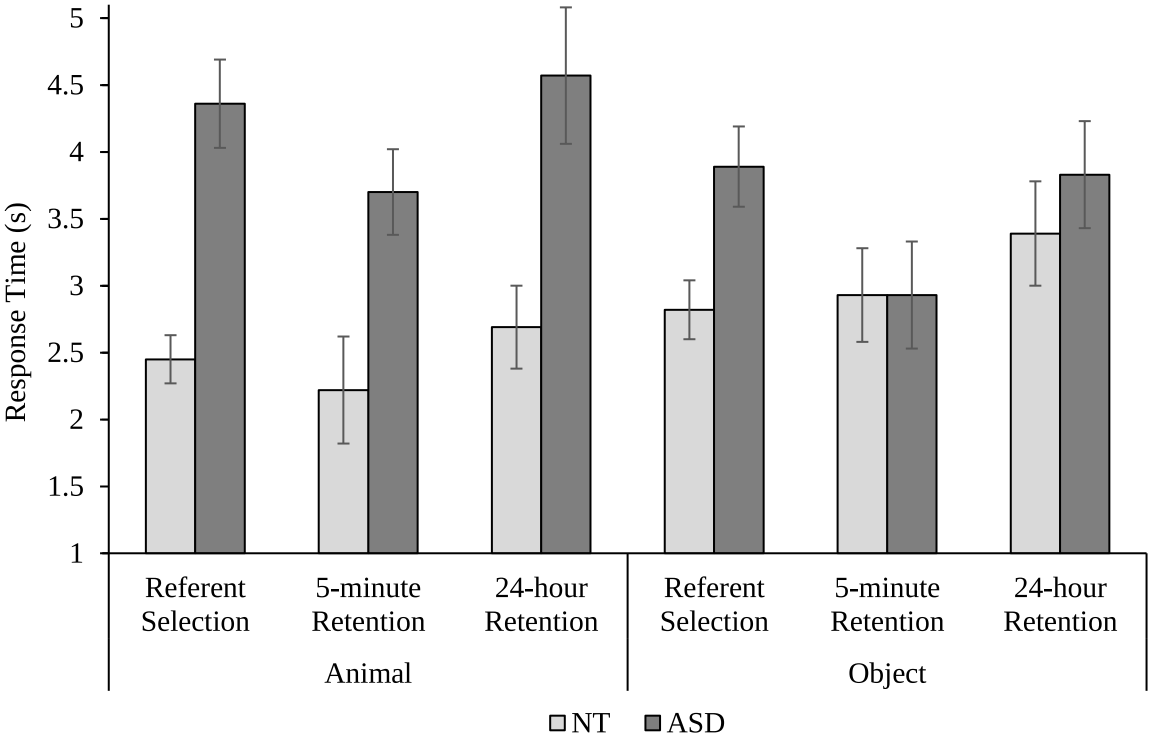
The best fitting model included a fixed effect of trial type ($z = -5.19$, $p < .001$; see Table 3) indicating that autistic children and neurotypical children responded with significantly greater accuracy on familiar trials than novel trials. However, it is noteworthy that both groups responded well above chance on novel trials with both object and animal targets (neurotypical children, animal condition $M = 0.86$, neurotypical children, object condition $M = 0.88$; autistic children, animal condition $M = 0.81$, autistic children, object condition $M = 0.73$), demonstrating their effective use of mutual exclusivity.

Referent selection response times

Children's response times for correctly answered referent selection trials were analysed using linear mixed-effects models, testing the effects of population, condition, and trial type. We calculated the average correct response time for each population in each trial type and condition, and removed outliers that were ≥ 3*SD* above the mean for the sub-group (e.g. autistic children in the animal condition responding to novel trials). We also removed three trials from autistic children who did not use their hand to respond (e.g. they responded hand-over-hand or using their head). The models in these analyses included 185 of 193 (96%) correct responses from autistic children, and 233 of 235 (99%) correct responses from neurotypical children. With outliers excluded, mean correct response times for each population are reported in Figure 4.

Figure 4

Mean response times on correctly answered referent selection, 5-minute retention, and 24-hour retention trials for neurotypical (NT) and autistic children (ASD), error bars show ± 1 SE



The best fitting model included significant fixed effects of trial type (*t* = 4.40, *p* <.001), population (*t* = 2.19, *p* = .037) and a population x condition interaction (*t* = 2.65, *p* = .008; see Table 3). Children in both populations were slower to generate correct responses for novel trials than familiar trials. Autistic children took significantly longer than neurotypical children to respond correctly across both conditions, but the difference between groups was greater in the animal condition than in the object condition.

Table 3

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) of children's accuracy on referent selection trials, and response times on correctly answered referent selection trials

	Fixed effects	Estimated coefficient	Std. error	*z*	Pr(> |z|)
Accuracy	(Intercept)	3.42	0.48	7.13	<.001
	Trial Type	-2.59	0.50	-5.19	<.001
		AIC	**BIC**	**logLik**	**deviance**
		326.1	380.7	-150.1	300.1
	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
Response Times	(Intercept)	3.41	0.33	10.35	<.001
	Population	1.44	0.66	2.19	.037
	Condition	0.07	0.19	0.38	.70
	Trial Type	0.85	0.19	4.40	<.001
	Population x Condition	1.03	0.39	2.65	.008
		AIC	**BIC**	**logLik**	**deviance**
		1836.1	1864.4	-911.1	1822.1

5-minute retention accuracy

Children's retention accuracy after five minutes was analysed via generalised linear mixed-effects models testing the effects of population, condition, referent selection accuracy, novel referent selection trial accuracy, and number of labels heard at referent selection for the target word. We excluded nine trials for autistic participants and three trials for neurotypical participants due to a technical error (8) and ambiguous responses (4). The models in these analyses contained 484 data points. The descriptive statistics are reported in Figure 3.

The final model was predicted by an interaction between population and condition ($z = 2.94$, $p = .003$; see Table 4). The effect of population was significant in the animal condition ($z = 2.50$, $p = .013$; autistic children responded significantly more accurately than neurotypical children), but not the object condition ($z = -1.24$, $p = .22$; the groups did not significantly differ). Autistic children responded with significantly greater accuracy in the animal condition compared to the object condition ($z = 2.08$, $p = .038$), but neurotypical children did not significantly differ in their response accuracy between conditions ($z = -1.69$, $p = .09$).

Table 4

Summary of the fixed effects in the final generalised linear mixed-effects models (log odds) of children's accuracy on 5-minute retention trials

Fixed effects	Estimated coefficient	Std. error	*z*	Pr(> |z|)
(Intercept)	-0.32	0.15	-2.17	.03
Population	0.28	0.25	1.13	.26
Condition	0.12	0.21	0.59	.56
Population x Condition	1.24	0.42	2.94	.003
	AIC	**BIC**	**logLik**	**deviance**
	656.7	690.2	-320.4	640.7

5-minute retention response times

Children's response times for correctly answered 5-minute retention trials were analysed using linear mixed-effects models. Outliers were identified and removed in the same way as described for referent selection trials. The models in these analyses included 102 of 106 (96%) correct responses from autistic children and 99 of 100 (99%) correct responses from neurotypical children. With outliers excluded, mean correct response times for each population are reported in Figure 4.

The inclusion of fixed effects (population and condition) did not improve model fit.

24-hour retention accuracy

Children's retention accuracy after 24-hours was analysed via generalised linear mixed-effects models testing the effects of population, condition, referent selection accuracy, accuracy on novel referent selection trials, number of labels heard at referent selection for the target word, and total 5-minute retention accuracy (all coded as described previously). Two autistic children in the animal condition and one neurotypical child in the object condition did not complete the 24-hour retention trials due to absence. We excluded 8 trials for autistic participants due to non-completion (1), simultaneous selection of multiple referents with their head and hand (4), and technical issues (3). The models in these analyses contained 464 data points. Descriptive statistics for 24-hour retention accuracy are presented in Figure 3.

The best fitting model contained fixed effects of population ($z = 1.92$, $p = .055$), total accuracy at 5-minute retention ($z = 4.43$, $p <.001$), referent selection accuracy ($z = 2.83$, $p = .005$), and number of labels heard at referent selection per word ($z = 3.18$, $p = .001$; see Table 5). These results show that (1) autistic children responded more accurately than neurotypical children (marginally significant), (2) children with higher 5-minute retention accuracy were significantly more likely to respond correctly on 24-hour retention trials, (3) children who heard more labels during referent selection were significantly more likely to respond correctly at 24-hour retention, and (4) children with higher referent selection accuracy for individual novel words were significantly more likely to respond correctly after 24 hours. Note the condition effect present at 5-minute retention for autistic children disappeared after 24-hours.

24-hour retention reaction times

Children's response times for correctly answered 24-hour retention trials were analysed using linear mixed-effects models. Outliers were identified and removed in the same way as for previous analyses. These analyses included 121 of 128 (95%) correct responses from autistic children, and 107 of 109 (98%) correct responses from neurotypical children. With outliers excluded, mean correct response times for each population are reported in Figure 4.

The best fitting model included a significant population x condition interaction (*t* = 2.82, *p* = .005; see Table 5). Autistic children took longer than neurotypical children to respond correctly in the animal condition, but not in the object condition.

Table 5

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) of children's accuracy on 24-hour retention trials, and response times on correctly answered 24-hour retention trials

	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Accuracy	(Intercept)	-2.14	0.44	-4.89	<.001
	Population	0.49	0.26	1.92	.055
	Total Accuracy at 5-minute Retention	0.35	0.08	4.43	<.001
	Referent Selection Accuracy	0.79	0.28	2.83	.005
	Number of Labels at Referent Selection	0.41	0.13	3.18	.001
		AIC	**BIC**	**logLik**	**deviance**
		604.8	642.0	-293.4	586.8
	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
Response Times	(Intercept)	4.01	0.50	7.98	<.001
	Population	1.89	1.00	1.88	.072
	Condition	0.26	0.32	0.82	.41
	Population x Condition	1.81	0.64	2.82	.005
		AIC	**BIC**	**logLik**	**deviance**
		1106.1	1126.7	-547.1	1094.1

3.5.2. Discussion: Accuracy and Response Times

These analyses assessed whether autistic and neurotypical children differ in their ability to disambiguate and retain novel words associated with high interest and neutral interest stimulus categories. Importantly, we examined children's accuracy and response speed across three distinct stages of word learning: referent selection, 5-minute retention, and 24-hour retention after a period of sleep. In comparison to neurotypical controls matched on receptive vocabulary, autistic children did not significantly differ in accuracy when spontaneously disambiguating the meanings of novel words using ME across conditions. After five minutes, autistic children retained significantly more novel word-referent mappings for animal stimuli compared to object stimuli, whereas neurotypical children retained novel words for both animals and objects with comparable accuracy. Autistic children also retained more novel animal names after a 5-minute delay than neurotypical children. Surprisingly, after a 24-hour delay, autistic children retained novel word-referent mappings with greater accuracy than neurotypical children (marginally significant difference). However, autistic children demonstrated slower response times than neurotypical children at each word learning stage (with significant differences detected at referent selection and 24-hour retention).

As predicted, our participants' referent selection across conditions demonstrates that both autistic and neurotypical children can employ ME to accurately identify the meanings of novel words, regardless of whether intended referents correspond with categories of interest. These findings, alongside existing evidence, show that autistic children can perform ME-based referent selection with similar accuracy to neurotypical children when expectations are based on receptive language ability (e.g. Carter & Hartley, 2021; Preissler & Carey, 2005). Using ME to actively disambiguate word meanings may represent a critical strategy through which both neurotypical and autistic children establish correct word-referent associations, increasing the quality of their audio-visual input and potentially contributing to long-term vocabulary development (Hartley et al., 2019). Unsurprisingly, children in both populations responded more accurately on familiar trials than novel trials, as children had pre-existing representations of referents for requested words.

Also in support of our predictions, effects of stimulus condition were observed at 5-minute retention. Here, autistic children achieved significantly greater accuracy in the animal condition – where they surprisingly exceeded neurotypical children – than in the object condition. As autistic children tend to process high interest stimuli with greater focus and intensity (Sasson et al., 2011), it may be that their interest in animals facilitated encoding of more robust word-referent representations that were less vulnerable to decay after five minutes. Indeed, previous studies have demonstrated positive relationships between children's attentional focus during word-referent mapping and subsequent retention accuracy (Bion et al., 2013; Hilton et al., 2017, 2019). It is also well-documented that many autistic individuals are adept at memorising information associated with topics and events of heightened interest (Bölte & Poustka, 2008; Happé, 1999). By contrast, neurotypical children's 5-minute retention accuracy did not significantly differ between conditions. These findings show that autistic children experience short-term retention benefits for words associated with high interest stimuli. Our subsequent analyses examining how differences in looking behaviour during referent selection predict 5-minute retention accuracy across the two groups will address this theory. However, autistic children's superior retention accuracy of animal stimuli compared to neurotypical children was not maintained after 24 hours due to a substantial increase in accuracy in the object condition (this point is addressed in the General Discussion).

In contrast to 5-minute retention, after 24 hours we observed that autistic children retained novel words for *both* objects and animals with greater accuracy than neurotypical children, and the condition effect on autistic children's retention accuracy disappeared. For autistic children, overnight improvement in the object condition compared to the animal condition may be attributed to sleep having more beneficial consolidation effects on weaker memory representations (Drosopoulos et al., 2007; Williams & Horst, 2014). Sleep plays a critical role in protecting newly acquired declarative memories against decay, and many studies have demonstrated that neurotypical children retain words more accurately after sleep (e.g. Axelsson et al., 2018; Williams & Horst, 2014). Active system consolidation theory (Diekelmann & Born, 2010) posits that sleep enhances novel word retention through the reactivation of recently encoded word-referent representations. New word-referent representations are initially fragile, but reactivation during sleep facilitates their integration into memory networks enabling longer-term retention (Gais & Born, 2004). While limited evidence suggests that novel word retention in intellectually able autistic children with age-expected language abilities may benefit from overnight sleep (e.g. Fletcher et al., 2020; Henderson et al., 2014), this study is the first to show a similar effect in autistic children with delayed language development.

One explanation for the observed between-population difference in 24-hour retention accuracy concerns chronological age. Children experience shorter sleep cycles than adults until 6 years (Hill et al., 2007; Montgomery-Downs et al., 2006), but longer sleep cycles are more beneficial for novel word consolidation (Diekelmann & Born, 2010). Therefore, it is possible that our autistic participants benefited more from overnight sleep because their average age exceeded 6 years, while the average age of the neurotypical children was significantly younger at just over 4 years. However, it is important to note that autism is commonly characterised by sleep disorders (e.g. bedtime resistance, sleep anxiety, difficulties falling asleep, parasomnia) that have the potential to negatively impact on overnight lexical consolidation and long-term vocabulary development (Díaz-Román et al., 2018; Souders et al., 2009). As no previous studies have tested 24-hour retention in autistic children with delayed language development, further research is required to replicate this effect and draw comparisons against neurotypical children matched on chronological age (in addition to children matched on receptive vocabulary) to control for developmental differences in sleep cycles. We also recommend that future studies investigate whether individual differences in sleep quality, duration, and disturbances predict variability in overnight consolidation of novel words for autistic children with language impairments.

At 24-hour retention, we found that both autistic and neurotypical children responded more accurately when they had heard more label repetitions during referent selection. This result highlights an important relationship between fast mapping and longer-term retention – *quantity* of auditory input received during referent selection influences the likelihood of successful memory consolidation. As proposed by the dynamic associative model (McMurray et al., 2012), successful identification of meaning may not necessarily support retention unless sufficient statistical input has been experienced. Cross-situational word learning studies show how increasing exposures to correct word-referent pairings can increase children's uptake from input and support encoding of word-referent representations that can be retrieved after delays (Hartley et al., 2020). Thus, for both autistic and neurotypical children, repeated exposures to novel word-referent associations may be critical to successful vocabulary acquisition, emphasising the importance of repetition as a component of communication interventions.

While response accuracy indicates whether children successfully identified and retained word-referent pairings, the time taken to generate correct responses provides insight into the speed of children’s information processing. At referent selection, children in both populations were quicker to respond correctly on familiar trials than novel trials. As children already knew the meanings of familiar words, correct responding simply required visual recognition of familiar referents. However, on hearing a novel word, children had to disambiguate the meaning of the word via mutual exclusivity. This required children to eliminate familiar competitors, ruling them out as targets, and shifting their attention to the novel object (Halberda, 2003). Since this task is more cognitively demanding, it is unsurprising that children were slower to make their selections on the novel trials (Bion et al., 2013).

Critically, autistic children took significantly longer than younger neurotypical children to generate correct responses, particularly in the animal condition. This finding aligns with previous evidence (e.g. Hartley et al., 2020) and suggests that, although word learning mechanisms appear to be intact, autistic children may require longer to process audio-visual stimuli in the service of word learning. Delays in processing stimuli could be attributed to general learning difficulties or differences in visual attention disrupting children's intake of information (Arunachalam & Luyster, 2018; Venker et al., 2018). On the other hand, autistic children's particularly slow responses in the animal condition across test stages could be due to their heightened interest in the novel stimuli (i.e. they chose to spend longer studying items in the array before identifying referents). Longer response times at referent selection may have ultimately benefitted their subsequent retention accuracy by affording more time to encode each target's perceptual features and providing the opportunity to hear more repetitions of the corresponding label. By extension, it is possible that neurotypical children's retention accuracy would have increased if they had also taken longer to respond on referent selection trials. Thus, we recommend that future research investigates potential speed-accuracy trade-offs across word learning mechanisms in autism and neurotypical development.

3.5.3. Does autistic and neurotypical children's visual attention differ while learning names for high and neutral interest stimuli, and does variability in visual attention predict learning accuracy?

In the following analyses we examine whether autistic and neurotypical children differed in their looking behaviour during each stage of the word learning task. To elucidate the relationship between visual attention and word learning performance, we also investigated how variability in children’s in-trial looking behaviour, and looking behaviour during referent selection, predicted their response accuracy.

All models were conducted using the glmer and lmer functions from the lme4 package in R (Bates et al., 2015). Population, condition, and trial type were coded as described previously. Proportion of time spent looking at the target object on each trial was scored between 0 and 1. Number of looks to the target object on each trial ranged from 0 to 14, with every new look towards the target stimuli being counted. Longest look to novel or familiar stimuli was coded as 0 (familiar) or 1 (novel) for analyses examining between-group differences in this measure, and -0.5 (longest look to familiar stimuli) or 0.5 (longest look to novel stimuli) when included as a fixed effect in analyses predicting accuracy. This variable was only included in referent selection analyses, as the retention phases only involved novel stimuli. Longest look to target or foil stimuli was coded as 0 (foil) or 1 (target) for analyses examining between-group differences in this measure at retention, and -0.5 (longest look to foil stimuli) and 0.5 (longest look to target stimuli) when included as a fixed effect in analyses predicting accuracy.

All analyses were undertaken following the same modelling procedures described for the accuracy and response time data. Each analysis started with a baseline model containing by-participant and by-word random intercepts, with a random slope of condition x trial type per participant for referent selection, or condition per participant for retention phases. As only final models are reported, please refer to Appendix C for full details of the model building sequences.

Referent selection

Linear mixed-effects models (unless otherwise specified) testing whether effects of population, condition, and trial type predicted variability in each visual attention measure during referent selection contained 496 data points (see Table 6 for descriptive statistics).

Generalised linear mixed-effects models testing whether children's in-trial visual attention predicted their behavioural response accuracy at referent selection contained 491 data points (exclusions were the same as described for the previous accuracy analyses).

Table 6

Mean values for the individual visual attention measures at referent selection, split by condition, trial type, and population

Looking Measure	Condition	Trial Type	ASD	NT
Proportion of time spent looking at target stimuli (0-1)	Animal	Familiar	0.56	0.65
		Novel	0.53	0.61
	Object	Familiar	0.59	0.74
		Novel	0.49	0.53
Number of looks towards target stimuli (0-14)	Animal	Familiar	2.13	1.27
		Novel	2.25	1.48
	Object	Familiar	2.27	1.34
		Novel	2.38	1.89
Longest look to novel or familiar stimuli (0/1)	Animal	Familiar	0.28	0.09
		Novel	0.62	0.73
	Object	Familiar	0.15	0.08
		Novel	0.60	0.64

Note. NT: neurotypical; ASD: autism spectrum disorder.

Proportion of time spent looking at the target

Proportion of time spent looking at the target referent was predicted by a fixed effect of population (*t* = -2.55, *p* = .016) and a trial type x condition interaction (*t* = 2.73, *p* = .006; see Table 7). Neurotypical children looked significantly longer at target stimuli than autistic children. The trial type x condition interaction was deconstructed by testing the effect of trial type on the animal and object condition separately. Children looked significantly more towards targets on familiar trials than novel trials in the object condition (*t* = -5.45, *p* <.001), but proportional looking did not differ across trial types for the animal condition (*t* = -1.12, *p* = .26).

Referent selection accuracy was predicted by a fixed effect of the looking measure (*z* = 7.16, *p* <.001; see Table 7). Across populations and conditions, as children's proportion of looking towards the target increased, so too did their referent selection accuracy.

Table 7

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) for proportion of time spent looking at the target stimuli during referent selection

	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
Between-group Differences	(Intercept)	0.59	0.02	32.66	<.001
	Trial Type	-0.10	0.02	-4.37	<.001
	Condition	-0.001	0.02	-0.01	1.00
	Population	-0.09	0.04	-2.55	.016
	Trial Type x Condition	0.12	0.04	2.73	.006
		AIC	**BIC**	**logLik**	**deviance**
		47.5	77.0	-16.8	33.5
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	-1.98	0.73	-2.73	.006
	Proportion of time spent looking at the target	13.23	1.85	7.16	<.001
		AIC	**BIC**	**logLik**	**deviance**
		185.7	240.2	-79.8	159.7

Number of looks to the target

Number of looks to target stimuli was predicted by fixed effects of population (t = 3.20, p = .003) and trial type (t = 2.34, p = .026; see Table 8). Across conditions, autistic children made more looks to target stimuli than neurotypical children, and children in both groups made more looks towards the target during novel trials than familiar trials.

Referent selection accuracy was predicted by a looking measure x population interaction (z = -2.01, p = .044; see Table 8). This interaction was deconstructed by testing the looking measure effect on autistic and neurotypical children separately. Across conditions, children who made more frequent looks towards the target during referent selection responded more accurately, but this effect was stronger for the autistic group (z = 4.37, p <.001) than the neurotypical group (z = 3.69, p <.001).

Table 8

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) for number of looks at the target stimuli during referent selection

	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
Between-group Differences	(Intercept)	1.87	0.12	15.38	<.001
	Population	0.77	0.24	3.20	.003
	Trial Type	0.25	0.11	2.34	.026
		AIC	**BIC**	**logLik**	**deviance**
		1624.0	1657.7	-804.0	1608.0
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	1.44	0.61	2.36	.018
	Number of looks to target	1.68	0.34	4.96	<.001
	Population	-0.43	1.19	-0.36	.72
	Number of looks x Population	-1.35	0.67	-2.01	.044
		AIC	**BIC**	**logLik**	**deviance**
		298.1	361.0	-134.0	268.1

Longest look to novel or familiar stimuli

This variable was analysed via generalised linear mixed effects models. Whether children's longest look was towards novel or familiar stimuli was predicted by a population x trial type interaction (*z* = -2.36, *p* = .018; see Table 9). This interaction was deconstructed by exploring the effect of population on familiar and novel trials separately, and trial type for neurotypical and autistic groups separately. Autistic children looked significantly longer at novel stimuli than neurotypical children on familiar trials (*z* = 2.34, *p* = .019), but not novel trials (*z* = -1.32, *p* = .19). Both autistic children (*z* = 4.34, *p* <.001) and neurotypical children (*z* = 8.30, *p* <.001) looked longer at the novel object on novel trials, however, this effect was larger for neurotypical children.

Referent selection accuracy was predicted by a looking measure x trial type interaction (*z* = 6.20, *p* <.001; see Table 9). This interaction was deconstructed by testing the looking measure effect on familiar and novel trials separately. Accuracy on novel trials significantly increased when children's longest look was towards the novel stimuli (*z* = 5.78, *p* <.001). However, on familiar trials, accuracy significantly decreased when children's longest look was towards the novel stimuli (*z* = -3.58, *p* <.001).

Table 9

Summaries of the fixed effects in the final generalised linear mixed-effects models (log odds) for the longest look to novel or familiar stimuli during referent selection

	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |t|)**
Between-group Differences	(Intercept)	-0.60	0.12	-4.84	<.001
	Population	0.33	0.25	1.31	.19
	Trial Type	2.80	0.32	8.90	<.001
	Population x Trial Type	-1.49	0.63	-2.36	.018
		AIC	**BIC**	**logLik**	**deviance**
		535.0	598.1	-252.5	505.0
	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	3.32	0.55	6.07	<.001
	Longest look novel or familiar	-0.21	0.48	-0.43	.66
	Trial Type	-2.34	0.70	-3.33	<.001
	Longest look x Trial Type	6.02	0.97	6.20	<.001
		AIC	**BIC**	**logLik**	**deviance**
		275.3	338.3	-122.7	245.3

5-minute retention

Linear mixed-effects models (unless otherwise specified) testing whether effects of population and condition predicted variability in each looking measure at 5-minute retention contained 493 data points (see Table 10 for descriptive statistics). Due to a technical error, three trials were removed for one neurotypical participant in the object condition.

Generalised linear mixed-effects models testing whether children's in-trial looking, and looking at referent selection, predicted their behavioural response accuracy at 5-minute retention contained 489 data points. We excluded three trials due to a technical error, and four trials due to ambiguous responses.

Table 10

Mean values for the individual visual attention measures at 5-minute retention, split by condition and population

Looking Measure	Condition	ASD	NT
Proportion of time spent looking at the target stimuli (0-1)	Animal	0.43	0.37
	Object	0.36	0.39
Number of looks towards the target stimuli (0-9)	Animal	1.81	1.00
	Object	1.62	1.30
Longest look to target or foil stimuli (0/1)	Animal	0.49	0.38
	Object	0.33	0.45

Proportion of time spent looking at the target stimuli

This looking measure did not significantly differ between populations or conditions; the inclusion of fixed effects did not improve fit in comparison with the baseline model.

Children's 5-minute retention accuracy was predicted by a fixed effect of this looking measure ($z = 11.21$, $p < .001$; see Table 11). Across groups and conditions, children who looked proportionately longer at the target stimuli responded more accurately during 5-minute retention.

Variability in proportion of time spent looking at the target stimuli during referent selection did not predict 5-minute retention accuracy.

Table 11

Summary of the fixed effects in the final generalised linear mixed-effects model (log odds) of 5-minute retention accuracy, predicted by proportion of time spent looking at the target stimuli at 5-minute retention

Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
(Intercept)	-3.65	0.37	-9.78	<.001
Proportion of time looking at target	8.12	0.72	11.21	<.001
	AIC	**BIC**	**logLik**	**deviance**
	366.0	391.1	-177.0	354.0

Number of looks to the target stimuli

This looking measure was predicted by a population x condition interaction ($t = 2.24$, $p = .033$; see Table 12). This interaction was deconstructed by exploring the effect of population in the animal and object conditions separately, and condition for neurotypical and autistic groups separately. While neurotypical children looked towards the target significantly more often in the object condition compared to the animal condition ($t = -2.50$, $p = .013$), number of looks towards the target by autistic children did not significantly differ between conditions ($t = 1.19$, $p = .23$). Autistic children made significantly more looks towards the target than neurotypical children in the animal condition ($t = 2.70$, $p = .011$), but the groups did not significantly differ in the object condition ($t = 1.30$, $p = .20$).

Children's 5-minute retention accuracy was predicted by a looking measure x condition interaction ($z = 2.60$, $p = .009$; see Table 12). Across populations, children who looked more frequently towards target stimuli achieved significantly higher 5-minute retention accuracy, but this effect was larger in the animal condition ($z = 5.51$, p <.001) than the object condition ($z = 3.80$, p <.001).

Across populations and conditions, children who looked more frequently towards target stimuli during referent selection responded with significantly greater accuracy at 5-minute retention (*z* = 2.20, *p* = .028; see Table 12).

Table 12

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) for 5-minute retention, predicted by number of looks to the target stimuli

	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
Between-group Differences	(Intercept)	1.43	0.13	11.15	<.001
	Population	0.56	0.24	2.31	.028
	Condition	-0.05	0.11	-0.41	.68
	Population x Condition	0.50	0.23	2.24	.033
		AIC	**BIC**	**logLik**	**deviance**
		1570.7	1608.5	-776.4	1552.7
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	-1.41	0.25	-5.58	<.001
	Number of looks to target	0.74	0.11	6.91	<.001
	Condition	-0.63	0.37	-1.70	.09
	Number of looks x Condition	0.52	0.20	2.60	.009
		AIC	**BIC**	**logLik**	**deviance**
		609.4	642.9	-296.7	593.4
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Referent Selection Looking Predicting Accuracy	(Intercept)	-0.70	0.21	-3.32	<.001
	Number of looks to target at referent selection	0.17	0.08	2.20	.028
		AIC	**BIC**	**logLik**	**deviance**
		658.3	683.4	-323.1	646.3

Longest look to target or foil

This looking measure did not significantly differ between populations or conditions; the inclusion of fixed effects did not improve fit in comparison with the baseline model.

Children's 5-minute retention accuracy was predicted by this looking measure (z = 12.19, p <.001; see Table 13). Looking longest at the target object (rather than a foil) was associated with significantly greater 5-minute retention accuracy across populations and conditions.

Whether children looked longest at the target or a foil during referent selection did not predict 5-minute retention accuracy.

Table 13

Summary of the fixed effects in the final generalised linear mixed-effects models (log odds) for longest look to the target/foil stimuli during 5-minute retention

Fixed effects	**Estimated coefficient**	**Std. error**	z	**Pr(> |z|)**
(Intercept)	-0.19	0.18	-1.04	.30
Longest look target	3.03	0.25	12.19	<.001
	AIC	**BIC**	**logLik**	**deviance**
	483.8	508.9	-235.9	471.8

24-hour retention

Linear mixed-effects models (unless otherwise specified) testing whether effects of population and condition predicted variability in each looking measure at 24-hour retention contained 471 data points (see Table 14 for descriptive statistics). Data from two autistic children in the animal condition, and one neurotypical child in the object condition, were excluded from this phase due to absence. One additional trial from an autistic participant in the object condition was removed due to non-completion.

Generalised linear mixed-effects models testing whether children's in-trial looking, and looking at referent selection, predicted their behavioural response accuracy at 24-hour retention contained 467 data points. As for the accuracy analyses, we excluded five trials for autistic participants in the object condition due to non-completion (1), and simultaneous selection of multiple referents with their head and hand (4).

Table 14

Mean values for the individual visual attention measures at 24-hour retention, split by condition and population

Looking Measure	Condition	ASD	NT
Proportion of time spent looking at the target stimuli (0-1)	Animal	0.45	0.37
	Object	0.42	0.42
Number of looks towards the target stimuli (0-11)	Animal	1.89	1.17
	Object	1.68	1.27
Longest look to target or foil stimuli (0/1)	Animal	0.50	0.41
	Object	0.49	0.48

Proportion of time spent looking at the target stimuli

This looking measure did not significantly differ between populations or conditions; the inclusion of fixed effects did not improve fit in comparison with the baseline model.

Children's 24-hour retention accuracy was predicted by a fixed effect of this looking measure ($z = 11.56$, $p < .001$; see Table 15). Across groups and conditions, as children's proportion of looking at the target object increased, so did their 24-hour retention accuracy.

The final best-fitting model testing whether children's 24-hour retention accuracy was predicted by proportion of time spent looking at the target during referent selection detected a looking measure x population x condition interaction ($z = -2.03$, $p = .042$; see Table 15). This interaction was deconstructed by separately testing the looking measure effect on each population in each condition. For autistic children, greater proportion of time spent looking at the target during referent selection predicted more accurate 24-hour retention in the object condition ($z = 2.73$, $p = .006$), but not the animal condition ($z = -0.83$, $p = .41$). However, this looking measure did not predict neurotypical children's 24-hour retention accuracy in either the object ($z = 0.35$, $p = .73$) or animal ($z = 1.08$, $p = .28$) condition.

Table 15

Summaries of the fixed effects in the final generalised linear mixed-effects models (log odds) for 24-hour retention accuracy, predicted by proportion of time spent looking at the target stimuli

	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	-2.85	0.29	-9.84	<.001
	Prop. of time looking to target	7.20	0.62	11.56	<.001
		AIC	**BIC**	**logLik**	**deviance**
		375.3	400.1	-181.6	363.3
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Referent Selection Looking Predicting Accuracy	(Intercept)	-0.43	0.28	-1.53	.13
	Prop. of time looking at target at referent selection	0.96	0.41	2.31	.021
	Condition	0.22	0.50	0.44	.66
	Population	0.56	0.54	1.04	.30
	Proportion x Condition	-0.71	0.84	-0.85	.39
	Proportion x Population	0.47	0.83	0.57	.57
	Condition x Population	2.45	1.00	2.46	.014
	Proportion x Condition x Population	-3.38	1.66	-2.03	.042
		AIC	**BIC**	**logLik**	**deviance**
		635.2	685.0	-305.6	611.2

Number of looks to the target stimuli

Number of looks to the target was predicted by a fixed effect of population. Across conditions, autistic children made significantly more looks towards the target than neurotypical children ($t = 2.35$, $p = .025$; see Table 16).

Children's 24-hour retention accuracy was predicted by a looking measure x population interaction ($z = -3.02$, $p = .002$; see Table 16). More frequent looks to the target object were associated with significantly higher 24-hour retention accuracy in both populations, however, this effect was larger for neurotypical children ($z = 6.01$, $p <.001$) than autistic children ($z = 4.02$, $p <.001$).

Across populations and conditions, children who looked more frequently towards target stimuli during referent selection responded with significantly greater accuracy at 24-hour retention ($z = 3.11$, $p = .002$; see Table 16).

Table 16

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) for 24-hour retention, predicted by number of looks towards the target stimuli

	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
Between-group Differences	(Intercept)	1.49	0.13	11.87	<.001
	Population	0.51	0.21	2.35	.025
		AIC	**BIC**	**logLik**	**deviance**
		1582.2	1611.3	-784.1	1568.2
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	-1.14	0.23	-5.04	<.001
	Number of looks	0.90	0.12	7.39	<.001
	Population	1.35	0.45	3.02	.003
	Number of looks x Population	-0.74	0.24	-3.02	.002
		AIC	**BIC**	**logLik**	**deviance**
		566.8	600.0	-275.4	550.8
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Referent Selection Looking Predicting Accuracy	(Intercept)	-0.51	0.26	-1.99	.047
	Number of looks to target at referent selection	0.29	0.09	3.11	.002
		AIC	**BIC**	**logLik**	**deviance**
		628.2	653.1	-308.1	616.2

Longest look to target or foil stimuli

Generalised linear mixed-effects models revealed that this looking measure did not significantly differ between populations or conditions; the inclusion of fixed effects did not improve fit in comparison with the baseline model.

Children's 24-hour retention accuracy was predicted by a looking measure x population interaction ($z = -2.78$, $p = .005$; see Table 17). Looking longest at the target was associated with significantly higher 24-hour retention accuracy in both populations, however, this effect was larger for neurotypical children ($z = 9.48$, $p <.001$) than autistic children ($z = 5.84$, $p <.001$).

When children's longest look during referent selection was towards the target rather than a competitor, their 24-hour retention accuracy increased ($z = 2.57$, $p = .010$; see Table 17).

Table 17

Summaries of the fixed effects in the final generalised linear mixed-effects models (log odds) for 24-hour retention accuracy, predicted by longest look to the target/foil stimuli

	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	0.16	0.15	1.07	.28
	Longest look target	2.64	0.24	10.97	<.001
	Population	0.68	0.28	2.40	.016
	Longest look target x Population	-1.34	0.48	-2.78	.005
		AIC	**BIC**	**logLik**	**deviance**
		485.0	518.2	-234.5	469.0
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Referent Selection Looking Predicting Accuracy	(Intercept)	0.0002	0.18	0.001	1.00
	Longest look to target at referent selection	0.56	0.22	2.57	.010
		AIC	**BIC**	**logLik**	**deviance**
		633.0	657.9	-310.5	621.0

3.5.4. Discussion: Looking during word learning and relationships with accuracy

We compared the visual attention of autistic and neurotypical children during novel word learning and investigated whether a range of measures predicted their retention accuracy. Neurotypical children spent longer looking at target stimuli during referent selection than autistic children, and both populations looked longer at target stimuli during familiar trials in comparison to novel trials. However, autistic children made more individual looks towards targets than neurotypical children during all three experimental phases. At 5-minute retention, autistic children looked more frequently at target stimuli in the animal condition than neurotypical children. Across groups and conditions, children's in-trial visual attention predicted accuracy at all three stages. Importantly, we discovered that children's looking behaviour at referent selection predicted their 5 minute and 24-hour retention accuracy, demonstrating that visual attention at encoding directly influences the likelihood of memory consolidation (see also Hilton et al., 2019 for similar results).

Children's proportional looking patterns during referent selection reflect their use of ME to inform fast mapping (Carter & Hartley, 2021; Preissler & Carey, 2005). Potentially, children spent more time looking towards targets during familiar trials compared to novel trials because they did not need to eliminate competitors to identify the requested referent. On the other hand, successful engagement in ME during novel trials required children to distribute their attention more evenly across stimuli in order to rule out non-target competitors and settle on the unfamiliar object as the intended referent (de Marchena et al., 2011; Horst & Samuelson, 2008; Zosh et al., 2013).

Neurotypical children looked proportionately longer towards targets than autistic children during referent selection, and autistic children were significantly more likely than neurotypical children to look longest at non-target novel stimuli on familiar trials. However, despite these differences in looking – which could be interpreted as evidence for less accurate identification of referents in autistic children – our behavioural data demonstrate that the groups' referent selection accuracy did not significantly differ. Importantly, these findings show that population differences in fast mapping accuracy should not necessarily be inferred from contrasting profiles of looking behaviour. Indeed, autistic children's heightened attention to novel stimuli across both referent selection trial types could have afforded more robust encoding of novel word-object associations (e.g. Axelsson et al., 2012), potentially supporting their superior accuracy at 24-hour retention.

Our analyses examining frequency of looks towards target stimuli during referent selection revealed that autistic children made more looks than neurotypical children, and both populations made more looks in novel trials (vs. familiar trials). It is possible that increased frequency of looks reflected participants' uncertainty and need to check multiple stimuli before making a confident response. For example, children likely made more looks during novel trials because the requested referent was unfamiliar and they needed to employ ME to rule out competitor objects (Halberda, 2003; Lewis et al., 2020). Given that our autistic sample was characterised by delayed language development, these children may have required more looks towards target stimuli due to uncertainties associated with comprehension and, potentially, the need for greater input to support processing.

At 5-minute retention in the animal condition, recall that autistic children responded with significantly greater accuracy and outperformed vocabulary-matched neurotypical children. Correspondingly, significant differences were detected in children's looking behaviour – autistic children looked at targets significantly more frequently than neurotypical children in the animal condition. These findings align with previous evidence that autistic children tend to explore stimuli that relate to their interests in a more detailed fashion than less interesting stimuli (Sasson et al. 2008, 2011). Given that predictive relationships were identified between all in-trial looking measures and response accuracy at 5-minute retention, it may be that group differences in visual attention reflected variability in short-term memory consolidation.

However, at 5 minute and 24-hour retention, proportion of time spent looking at target stimuli did not significantly differ across populations or conditions. This aligns with a growing body of evidence demonstrating that neurotypical and autistic children perform comparably well on word learning tasks when utilising proportion of looking time as an indicator of accuracy (e.g. Akechi et al., 2013; Swensen et al., 2007; Venker, 2019). Our proportional looking data could be interpreted as evidence for a lack of group differences in word learning, yet group differences in accuracy were identified at both 5 minute and 24-hour retention. Consequently, this finding calls into question the validity of drawing conclusions about retention accuracy in autistic and neurotypical children based exclusively on data representing proportion of time spent looking at requested referents.

3.6. General Discussion

This study investigated whether autistic children with delayed language development and neurotypical children differ in accuracy and/or visual attention when learning words associated with high interest (animals) and neutral interest (objects) stimuli. Importantly, this study is the first to examine how direct measures of word learning accuracy and visual attention inter-relate in autism. Our data show that autistic children can identify and retain the meanings of novel animals *and* novel objects as accurately as neurotypical children matched on receptive vocabulary. Indeed, autistic children retained names for novel animals with greater accuracy than neurotypical children after five minutes. Unexpectedly, autistic children showed a greater increase in their accuracy between 5 minute and 24-hour retention tests and outperformed neurotypical children after a night's sleep. However, autistic children demonstrated slower response times than neurotypical children at each word learning stage, particularly towards animal stimuli. Analyses of in-trial visual attention revealed that autistic children looked at targets significantly more often than neurotypical children at each word learning stage, and this measure predicted response accuracy across populations. Despite observing population differences in retention accuracy, but not referent selection accuracy, autistic and neurotypical children significantly differed in their proportion of time spent looking at targets during referent selection, but not retention. We also discovered that children's 5 minute and 24-hour retention accuracy was predicted by visual attention at referent selection. Children's 24-hour retention accuracy was also predicted by auditory input. These findings indicate a relationship between fast mapping and slow learning mechanisms.

In partial support of our predictions, effects of stimulus condition were observed at 5-minute retention. Here, autistic children achieved significantly greater accuracy in the animal condition – where they exceeded neurotypical children – than in the object condition. We also identified complementary between-group differences in visual attention; autistic children looked more frequently at animal targets than neurotypical children. As autistic children tend to process high interest stimuli with greater focus and intensity (Elison et al., 2012; Sasson et al. 2008, 2011), it may be that their interest in animals facilitated encoding of more robust word-referent representations that were less vulnerable to decay after five minutes. By contrast, neurotypical children's 5-minute retention accuracy did not significantly differ between conditions, suggesting that differences in interests associated with target stimuli did not influence the strength of their encoded word-referent representations after a short delay. However, condition effects on autistic children's 5-minute retention were not maintained after 24 hours due to their accuracy in the object condition markedly increasing ($M_{5\text{-}minute} = 0.37$; $M_{24\text{-}hour} = 0.57$). Autistic children's overnight improvement in the object condition may be attributed to sleep having more beneficial consolidation effects on weaker memory representations (Drosopoulos et al., 2007). Previous studies have shown that neurotypical children's retention of fragile word-referent associations encoded under relatively difficult conditions improves more substantially after a period of sleep than their retention of strongly encoded word-referent associations (Axelsson et al., 2021; Williams & Horst, 2014). Overall, these findings demonstrate that autistic children can effectively identify and retain novel word meanings across varied stimuli, and that benefits conferred by high interest stimuli are relatively short-term.

This study is one of the few to investigate both referent selection *and* retention in autism – a necessary requirement to comprehensively study children's novel word learning (Hartley et al., 2019; Horst & Samuelson, 2008). Despite our autistic participants' delays in language development, they did not respond with significantly lower accuracy than neurotypical children at referent selection, 5-minute retention, or 24-hour retention (for similar results, see: Carter & Hartley, 2021; Haebig et al., 2017; Hartley et al., 2019, 2020; Luyster & Lord, 2009). Indeed, their unexpectedly superior 24-hour retention accuracy potentially signposts the importance of sleep to their long-term lexical consolidation (as is the case for neurotypical development; Axelsson et al., 2016; Brown et al., 2012; Henderson et al., 2012; Horváth et al., 2015; Williams & Horst, 2014). As no previous studies have tested 24-hour retention in autistic children with delayed language development, further research is required to replicate this effect and draw comparisons against neurotypical children matched on chronological age (in addition to children matched on receptive vocabulary) to control for developmental differences in sleep cycles. Crucially, the response profile of autistic children across word learning accuracy measures demonstrates that fundamental mechanisms subserving their identification and retention of meaning are intact and yield learning outcomes commensurate with expectations based on their receptive vocabulary (Carter & Hartley, 2021; Hartley et al., 2019, 2020). Thus, if autistic children do not qualitatively differ in terms of word learning mechanisms, delays in their naturalistic language learning and vocabulary development must have an alternative cause.

Our data indicated group differences in visual attention and speed of processing during word learning. In comparison with neurotypical controls, autistic children made more looks towards target stimuli during all three word learning phases, and their attention was increasingly drawn towards novel stimuli during referent selection. While these differences in looking behaviour may not have had a detrimental impact on learning accuracy, they may explain why autistic children took longer to generate correct responses. Frequency of looks towards target stimuli at referent selection had a stronger predictive effect on autistic children’s response accuracy, suggesting that repeated examination of novel target stimuli may have been necessary to support their correct identification of meaning. These differences in looking behaviour and response times potentially signal that autistic children are less *efficient* at processing audio-visual stimuli when disambiguating novel word meanings than neurotypical children with similar language abilities (Ferman, 2021; Hartley et al., 2020; Ricketts et al., 2015).

This study's combination of behavioural accuracy and looking measures represents an important methodological advancement that has generated new insight into autistic children's word learning. Specifically, we discovered that these measures both complement and contradict one another. On one hand, across populations and conditions, increased looking predicted accuracy at each word learning stage. On the other hand, population differences in certain aspects of looking behaviour did not necessitate population differences in learning accuracy (or vice versa). It could be inferred from the analyses of time spent looking towards target stimuli that autistic children responded with significantly lower accuracy on referent selection trials than neurotypical children but did not differ on either 5-minute or 24-hour retention. On the contrary, the groups' referent selection accuracy did not significantly differ, and autistic children responded more accurately at retention. These data suggest that proportion of time spent looking at targets may be less representative of children's genuine learning outcomes than alternative measures of visual attention, such as frequency of looking towards targets. Indeed, our findings spotlight the number of looks towards targets as an important measure – more frequent looks at referent selection predicted greater retention accuracy at 5 minutes and 24 hours, and population differences in this variable complemented between-group differences at both retention stages. Thus, we recommend that future research investigating autistic children's word learning using looking paradigms should carefully consider their selection of visual attention measures, as studies may draw (potentially erroneous) conclusions exclusively from time spent looking at targets.

Our findings also provide new insight into the relationship between fast mapping and retention. Previous research with neurotypical children suggests that these word learning mechanisms are distinct (e.g. Horst & Samuelson, 2008), but our findings indicate a link between referent selection and overnight retention that is mediated by exposure to visual and auditory input. We discovered that superior 24-hour retention accuracy was predicted by hearing more label repetitions at referent selection, as well as looking at targets longest and more frequently. Together, these findings show that *how* children experience stimuli during encoding directly predicts the likelihood of longer-term learning. As posited by McMurray et al.'s (2012) dynamic associative model of learning, increased exposures and statistical input in both visual and auditory domains are critical for supporting consolidation of newly acquired word-referent associations into children's vocabularies. Moreover, our data are congruent with recent studies showing that autism does not impair statistical learning mechanisms when deciphering word meanings (e.g. Hartley et al., 2020; McGregor et al., 2013) or processing visual stimuli (Foti et al., 2015; Roser et al., 2015).

Thinking practically, our findings have the potential to inform the development of interventions designed to scaffold autistic children's word learning. While autistic children are often highly motivated to interact with touch-screen technology, evidence of effective learning via this platform has been mixed (Allen et al., 2016; Wainwright et al., 2020). Our study demonstrates that it is possible to teach children novel words associated with different types of stimuli using a touch-screen computer when distractions are minimised. Additionally, we have shown that employing ME-based referent selection is an effective way to facilitate autistic children's word learning. Presenting limited options helps children to utilise their existing vocabulary to engage in active learning, deciphering which novel referent is associated with a novel word. Furthermore, progression through trials was dependent on the speed of children's responses, enabling them to engage with stimuli at their own pace. In natural environments, speech occurs at a rate of approximately 150 words-per-minute (Studdert-Kennedy, 1986), significantly faster than in most experimental contexts. The increased rate of stimuli presentation and greater attentional demands in natural communicative situations could create a processing bottleneck for autistic children, reducing the quality of their visual-auditory input and strength of associations between words and referents (Hartley et al., 2020; McMurray et al., 2012). As such, applying unrestricted processing times in clinical and educational interventions, as well as natural learning environments where possible, may facilitate autistic children's vocabulary acquisition.

This study is not without limitations. Firstly, we must reflect on the implications of matching autistic and neurotypical children on receptive vocabulary, but not chronological age (the autistic sample was significantly older than the neurotypical sample). We selected these matching criteria because the study's purpose was to compare word learning abilities across populations when delays in language development were controlled for. Previous studies comparing various aspects of language development in autism against chronological age norms for neurotypical children have consistently found deficits (e.g. Charman et al., 2003; Luyster et al., 2007). However, these differences could be due to various factors, including neurotypical children's generally superior vocabulary learning abilities and differences in nonverbal intelligence. Thereby, matching on receptive vocabulary allows us to identify whether autistic children fundamentally differ in *how* they learn words relative to neurotypical children with similar vocabularies. Secondly, our recruitment of participants was hindered by school closures and lockdown restrictions associated with the COVID-19 pandemic which occurred whilst the study was underway. Thus, we would recommend that future studies combining behavioural and looking measures seek to replicate our findings with larger samples.

In summary, this study has advanced understanding of how autistic and neurotypical children identify and retain novel word meanings, and how these processes are influenced by interest in stimulus categories. Despite our autistic participants' delayed language development, they responded at least as accurately as vocabulary-matched neurotypical children on measures of referent selection, 5-minute retention, and 24-hour retention. Differences between neutral- and high interest stimuli were only observed at 5-minute retention, where autistic children recalled animal names significantly more accurately than object names. This condition advantage disappeared after 24 hours, suggesting that superior learning of words associated with high interest stimuli was relatively short-term. Thus, under favourable experimental conditions, differences in attention to stimuli that are perceived to be more or less interesting may not be detrimental to autistic children's word learning. Although these results demonstrate that fundamental word learning mechanisms are not atypical in autism, autistic children were slower than neurotypical children to generate correct responses, particularly in the animal condition. Autistic children also differed in their visual engagement with stimuli, potentially indicating less efficient processing of visuo-auditory input. As children responded at their own pace and processing times were unrestricted, spending longer studying stimuli may have benefited autistic children's accuracy (i.e. a speed-accuracy trade-off). However, restricted processing times and the rapid pace of input during naturalistic communicative interactions could place strain on autistic children's word learning mechanisms and impact on their accuracy. Our results highlight the risk of drawing inaccurate conclusions about autistic children's learning from looking time alone and show the importance of including multiple measures of learning outcomes and visual attention. Overall, these findings indicate environmental conditions to scaffold word learning in clinical and educational contexts.

Chapter 4: Do autistic and neurotypical children's stimulus preferences influence their accuracy and visual attention when learning novel words?

4.1. Chapter Introduction

It is established that children's interests in specific categories influence novel word learning due to enhanced attentional allocation to such stimuli (Ackermann et al., 2020, Rothwell et al., Study 1, Study 2). However, such interests are not the only element that may drive word learning. For example, we know that - in addition to specific categories of stimuli - children often prefer unique properties of stimuli. Research suggests that allowing neurotypical children to learn about objects of their choice affords better learning outcomes (Begus & Southgate, 2012; Partridge et al., 2015). However, these theories have not yet been tested in autistic populations.

Whilst Studies 1 and 2 of this thesis explore how categorical interests impact word learning, Study 3 (Chapter 4) investigated the effects of autistic and neurotypical children's idiosyncratic preferences for object stimuli. Children identified novel stimuli as being 'liked' or 'disliked' in a session prior to the experimental tasks, and these stimuli were allocated to liked and disliked conditions respectively. Children were then tested on their ability to learn novel names for objects that differed on this preference dimension. Like Study 2, here we examined touch-screen responses alongside in-trial visual attention. Addressing this gap in knowledge will advance theoretical understanding of word learning in autism and neurotypical development, highlighting how individual preferential biases to selective stimuli influence learning and visual attention, and how these measures inter-relate.

Author contribution: *Charlotte Rothwell*: study design, data collection, analysis, writing, review. *Gert Westermann*: review. *Calum Hartley*: study design, review.

4.2. Abstract

To decipher novel word-referent mappings, it is imperative that children pay attention to the right things at the right times during word learning. Disruptions to attention may therefore contribute to word learning difficulties experienced by some autistic children. Here, we combine accuracy and looking measures to investigate whether autistic children with delayed language development and neurotypical children matched on receptive vocabulary differ in accuracy and/or visual attention when learning words associated with self-identified liked and disliked object stimuli. In a computer-based mutual exclusivity referent selection task, neurotypical children mapped novel word-object associations with greater accuracy than autistic children. However, after 5 minutes and 24 hours, autistic and neurotypical children retained novel object names with comparable accuracy irrespective of stimulus preferences. Success on previous word learning stages predicted subsequent retention accuracy, indicating a relationship between short- and long-term word learning mechanisms. Across word learning stages, response speed of autistic and neurotypical children did not significantly differ. However, analyses of visual attention revealed that autistic children looked at targets significantly more often than neurotypical children. Moreover, conditional differences in visual attention at referent selection did not result in conditional differences in behavioural response accuracy across word learning stages. Together, these findings suggest that idiosyncratic preferences to stimuli do not influence autistic or neurotypical children's attention allocation enough to impact their novel word retention. Interestingly, even after a 24-hour delay, sleep did not consolidate novel object labels in either population. Differences in visual attention in autism did not lead to diminished novel word retention, suggesting that attentional allocation does not necessitate altered word learning accuracy.

Keywords: Word learning; Autism; Preferences; Attention; Referent Selection; Retention

4.3. Introduction

Word learning is a vital component of children's language acquisition (Patael & Diesendruck, 2008). Neurotypical children can understand approximately 200 words before 2-years (Dale & Fenson, 1996), increasing to approximately 2600 words by school age (Kidd & Donnelly, 2020). By contrast, many autistic children are delayed in their receptive and expressive vocabulary development in comparison to neurotypical peers (Artis & Arunachalam, 2023; Kover et al., 2013). Perhaps surprisingly, recent evidence shows that word learning mechanisms – and the relationships between them – are not atypical in many autistic children, including those with delayed language development (Foti et al., 2015; Hartley et al., 2019, 2020; Roser et al., 2015). An alternative explanation for delays in language acquisition in autism could be atypical attentional behaviours that reduce the quality of visual and auditory input and inhibit children's capacity to learn from their environment (Arunachalam & Luyster, 2015, 2018; Hartley et al., 2019, 2020). However, few studies to date have empirically studied how children's attentional mechanisms influence word learning in autism or neurotypical development. Here, we investigate how preferences for specific objects and endogenous attentional biases affect identification and retention of novel words in neurotypical and autistic children.

Word learning is a multi-stage process, whereby children must take in cues from their linguistic and non-linguistic environment and identify the correct mapping between a novel word and its meaning ('referent selection;' Carey & Bartlett, 1978; McMurray et al., 2012; Monaghan, 2017). Following this, the child must store the correct word-referent association for later retrieval ('retention;' Gleitman, 1990). The 'dynamic associative account' explains the relationship between mechanisms, proposing that referent selection and retention utilise separate 'fast mapping' and 'slow associative learning' processes that occur on different timescales (McMurray et al., 2012).

‘Fast mapping’ occurs when children form rapid, in-the-moment, associations between novel words and meanings (Kucker et al., 2015; Samuelson & McMurray, 2017). During fast mapping, children must overcome the challenge of referential ambiguity – a novel word could have multiple potential referents (Cartmill et al., 2013; Markman, 1989). One way that children resolve referential ambiguity is by applying the mutual exclusivity principle (ME). ME refers to the assumption that each referent only has a single label. Thus, when a single unfamiliar object is presented alongside one or more familiar objects, children use ME to deduce that a novel word must refer to the unfamiliar object (Markman & Wachtel, 1988; Merriman & Bowman, 1989).

However, referent selection is only one component of word learning – true word learning requires children to *retain* associations between novel words and their referents (McMurray et al., 2012). Learning, characterised by encoding and memory consolidation (Vlach & DeBrock, 2019), is only evidenced when children can accurately identify the referent of a novel word after a delay. Indeed, many studies demonstrate that neurotypical toddlers who fast map with excellent accuracy often forget novel words after a five-minute delay (Bion et al., 2013; Gurteen et al., 2011; Horst & Samuelson, 2008). These findings reveal that referent selection and retention are separate word learning processes. It is proposed that retention is underpinned by associative learning mechanisms that gradually strengthen word-object relationships over multiple exposures across situations and contexts (Hartley et al., 2020; McMurray et al., 2012; Monaghan, 2017). Moreover, sleep enhances retention by protecting new memories against decay (Diekelmann & Born, 2010; Gais & Born, 2004), and studies have shown that both autistic and neurotypical children recall novel word-referent associations more accurately after sleeping (e.g. Henderson et al., 2012, 2014; Horváth et al., 2015; Rothwell et al., accepted, see Study 2; Williams & Horst, 2014).

Crucially, children's word learning and visual attention are inter-related. To correctly learn a word, children must examine an array of stimuli and identify the intended referent whilst concurrently ruling out competitors and attending to cues such as gaze and gesture (Hollich et al., 2000; Preissler & Carey, 2005; Samuelson et al., 2017). Hilton et al. (2019) suggest that individual characteristics influence visual attention and word learning success. They demonstrated that shy children, who spent less time looking towards target stimuli, subsequently had poorer novel word retention (see also Hilton & Westermann, 2017 for similar results). Axelsson et al. (2012) exogenously manipulated neurotypical toddlers' attention and examined how this affected their novel word retention. When target objects were made more salient through visual illumination, children achieved superior retention accuracy, demonstrating that increased attention is beneficial for word learning via referent selection. Together, these studies show that children's visual attention during encoding impacts their subsequent retention of word-referent associations (Ackermann et al., 2020; Bion et al., 2013; Smith & Yu, 2013).

Children's attentional biases towards specific stimuli can also influence their word learning. Children often prefer objects that are larger, display bright and multi-colours, or are varied in texture and shape (Pereira et al., 2014), and can retain more information about these kinds of stimuli (Smith et al., 1996). Heightened interest in particular objects increase children's attentional focus, which can in turn benefit their encoding of more robust word-referent representations (Ackermann et al., 2020; Begus & Southgate, 2012). In Begus et al. (2014), 16-month-old children selected one of two novel objects, and then learnt about the function of either their chosen or non-chosen object. The children retained significantly more information about their chosen objects than their non-chosen objects, demonstrating that children's active engagement and selection of information leads to superior learning. Further, Partridge et al. (2015) investigated self-directed learning in neurotypical pre-schoolers and found that children who chose which objects they learned labels for achieved superior accuracy at training and test than their counterparts who did not choose which objects they learned about. However, with increased task difficulty, this conditional effect diminished. Thus, when a task is of low to moderate difficulty, children may find it easier to learn words for objects they are interested in due to a curiosity-driven desire to seek out more information about these items. Similarly, Ackermann et al. (2020) utilised questionnaires and pupillometry to identify neurotypical 2-year-olds' preferred specific stimuli and categories of stimuli. Children then learnt four novel word-referent pairings via ME-based referent selection. Overall, children demonstrated better novel word retention of stimuli belonging to categories they were more interested in, and for specific novel stimuli of interest. This indicates the positive role of interests and attention on neurotypical children's word learning.

Early theories attributed autistic children's delayed language development to difficulties interpreting social cues that attune children's attention to the correct target, thus informing referent selection (e.g. Baron-Cohen et al. 1997; Preissler & Carey 2005). However, more recent research demonstrates that autistic children with wide-ranging language abilities can use social informants to identify novel referents under controlled experimental conditions (Hani et al., 2013; Hartley et al., 2020; McGregor et al., 2013). Studies have also shown that, in comparison to neurotypical peers, autistic children with concomitant language delay are unimpaired in their use of lexical heuristics such as ME (Carter & Hartley, 2021; de Marchena et al., 2011; Preissler & Carey, 2005; Rothwell et al., accepted, see Study 2). These findings suggest that autistic children are capable of employing social and lexical strategies in the service of referent selection, as neurotypical children do.

Whilst referent selection has been widely researched, very few studies have investigated retention in autism. Norbury et al. (2010) report that verbally able autistic children can retain word-object mappings as accurately as neurotypical controls, but their memory of semantic features may degrade more rapidly. Hartley and colleagues found that autistic children with delayed language development retained novel names at least as accurately as vocabulary-matched neurotypical children after utilising social and non-social feedback in an ME task (Hartley et al., 2019) and tracking statistical co-occurrences between words and objects (Hartley et al., 2020). These data reveal that fundamental word learning mechanisms in language delayed autistic children are not impaired. However, the delayed response times in autistic children may represent reduced pace of word learning mechanisms in autism.

Studies report that whilst neurotypical children can allocate attention freely across learning contexts, autistic children often demonstrate attentional processes that are 'sticky' and challenging to disengage (Elsabbagh et al., 2009, 2013; Pierce et al., 2011, 2016; Sacrey et al., 2013, 2014). These behaviours are often attributed to restricted and repetitive behaviours and interests that are prevalent in many autistic children, causing them to perseverate their attention towards a specific range of topics or interests (RRBIs; Kanner, 1943; Honey et al., 2012; Richler et al., 2007). Compared to neurotypical children, autistic children often intensively allocate their visual attention towards salient perceptual features or preferred categories, at the detriment of other stimuli (Hartley & Allen, 2014; Pierce et al., 2011, 2016; Venker et al., 2022). Previous research also indicates that autistic children examine high interest and preferred stimuli in greater detail than neurotypical peers (Sasson et al., 2011). For instance, Pierce et al. (2016) suggested that autistic toddlers had a heightened preference for geometric shapes compared to neurotypical peers, leading to atypical attentional distributions particularly towards preferred shapes. These differences in visual attention could have profound implications for language acquisition. If a non-target object captures the attention of an autistic child, they may not disengage their attention and redirect towards the to-be-learned stimuli when hearing the novel word. Consequently, the misalignment in auditory-visual statistics may result in either incorrect word learning, or cause word learning to not take place at all (e.g. Baron-Cohen et al., 1997; Tenenbaum et al., 2017; Venker et al., 2018).

Differences in visual attention during identification of meaning may affect encoding of novel words (e.g. Axelsson et al., 2012; Bion et al., 2013; Hartley et al., 2019, Hilton et al., 2019). If target stimuli appeal to autistic children's interests or preferences, perseverative allocation of attention may generate robust encodings of word-referent associations that are less susceptible to decay (e.g. Ackermann et al., 2020, Rothwell et al., Study 2). However, if target stimuli are not interesting to autistic children, their attentional allocation may be reduced, resulting in fragile encodings of word-referent associations in memory (e.g. Tenenbaum et al., 2014, 2017). To our knowledge, our Study 2 is the only study to directly investigate how autistic children's interests influence their visual attention and retention accuracy in a word learning task. We demonstrated that neurotypical and language delayed autistic children matched on receptive vocabulary could successfully identify the names of both novel objects (neutral interest stimuli) and novel animals (high interest stimuli) via mutual exclusivity in a fast-mapping task. After a 5-minute delay, autistic children remembered names for more novel animals than neurotypical children, demonstrating that autistic children could more easily remember novel names associated with categories they were interested in. After a 24-hour delay, autistic children retained more novel object and animal names than neurotypical children. Despite children's word learning success, autistic children took longer to generate correct responses than vocabulary-matched neurotypical children, particularly for stimuli of interest. Autistic children also looked significantly more frequently at target stimuli regardless of whether they belonged to a category of interest (animals) or a neutral category (objects). For both groups, visual attention during encoding at referent selection significantly predicted retention accuracy. It remains unclear how visual attention may affect autistic children’s word learning when all target stimuli belong to the same category (objects) or differ in terms of children’s self-identified preferences, rather than pre-existing categorical interests. Investigating this would elucidate whether stimuli that pertain to children’s interests could be used to scaffold learning in different contexts.

The present study investigates how preferences for specific stimuli influence word learning in autistic and neurotypical children. Autistic and neurotypical children matched on receptive vocabulary first completed an object-sorting task where they selected their four favourite and four least favourite stimuli from a pool of twelve novel objects. Children then identified the meanings of novel words in a computer-based ME referent selection task with two-within subject conditions. In one condition, children learnt the names for their four favoured novel objects, and in the other condition they learnt the names for their four least favoured novel objects. Retention of novel words was tested after 5 minutes and 24 hours. The retention tests following a 24-hour delay allowed us to examine the robustness of novel word representations associated with favoured and less favoured stimuli, and how sleep determines lexical consolidation in autistic children with concomitant language delay. Since sleep disorders are prevalent in autism (Díaz-Román et al., 2018; Souders et al., 2009), it is possible that these difficulties could impact autistic children’s consolidation of recently mapped word-referent associations. However, Rothwell et al. (accepted, see Study 2) recently reported that language delayed autistic children benefitted from a period of sleep *more* than neurotypical peers matched on receptive vocabulary when learning names for animals and objects. Here, we explore whether the effect of sleep on lexical consolidation differs for stimuli within the same category that are liked or disliked. We recorded children’s looking behaviour during referent selection and both retention measures, plus their behavioural response times and accuracy. This allowed us to compare the populations’ visual engagement at each stage of the task and investigate whether variability in visual engagement predicted learning accuracy. While many studies utilise looking time as the sole dependent variable, with longer looking usually interpreted as more accurate learning (e.g. Axelsson et al., 2020; Tenenbaum et al., 2014, 2017), our novel study advances methodology by examining how multiple measures of visual attention and behavioural accuracy inter-relate.

As previous evidence demonstrates that both neurotypical children and autistic children spontaneously apply the ME principle when mapping novel labels to novel objects (e,g, Carter & Hartley, 2021; Hartley et al., 2019; Rothwell et al., accepted, see Study 2), we did not expect between-population or between-condition differences in accuracy during referent selection. Given that heightened attentional focus can enhance word learning (Axelsson et al., 2012; Hilton & Westermann, 2017; Rothwell et al., Study 2), we predicted that children in both populations would consequently retain novel names associated with their liked novel stimuli with greater accuracy than for their disliked stimuli. After 24 hours, we anticipated that both autistic and neurotypical children would retain novel words at least as accurately as during 5-minute retention, due to sleep-induced consolidation. Based on the findings of Rothwell et al. (accepted, see Study 2), we tentatively predicted that autistic children would experience a greater benefit of sleep consolidation than their neurotypical peers due to longer sleep cycles associated with older chronological age. We also anticipated that autistic children would be slower to generate correct responses than neurotypical children, especially during referent selection stages, suggesting that autistic children may process audio-visual stimuli at a slower pace (Hartley et al., 2020; Rothwell et al., accepted, see Study 2; Ricketts et al., 2015).

During referent selection, we expected that children would look longer and more frequently at their self-identified liked stimuli than disliked stimuli, and that this effect would be stronger for autistic children due difficulties disengaging attention from interesting stimuli (Elsabbagh et al., 2009, 2013; Landry & Bryson, 2004; Sacrey et al., 2014). We expected increased visual attention to targets would predict response accuracy across conditions, groups, and task stages. In particular, we predicted that increased attention at referent selection would afford superior retention for both populations, especially during 24-hour retention (e.g. Rothwell et al., Study 2). However, based on our previous findings (e.g. Rothwell et al., Study 2), we predicted that between-population differences in visual attention may not directly translate to poorer word learning accuracy. This would reflect autistic individuals' ability to utilise alternative strategies to generate correct responses on a range of cognitive tasks (Happé, 1995; Norbury et al., 2010). Importantly, this research will advance theoretical understanding of word learning in autism and neurotypical development, highlighting how individual preferential biases to selective stimuli influence learning and visual attention, and how these measures inter-relate.

4.4. Method

Participants

Participants were 17 autistic children (14 males, 3 females; *M* age = 94.82 months; *SD* = 20.90) recruited from specialist schools, and 17 neurotypical children (10 males, 7 females; *M* age = 52.12 months; *SD* = 17.81) recruited from mainstream schools, nurseries, and Lancaster University BabyLab (see Table 1). Participants were monolingual native English speakers and had normal or corrected-to-normal colour vision. Autistic children had a pre-existing diagnosis from a qualified clinician, using standardised instruments (i.e. Autism Diagnostic Observation Scale and Autism Diagnostic Interview – Revised; Lord et al., 1994, 2002) and expert judgement. Diagnoses were confirmed via the Childhood Autism Rating Scale 2 (ASD *M* score = 29.74, *SD* = 7.07; NT *M* score = 16.88, *SD* = 2.72; Schopler et al., 2010). This measure was usually completed by class teachers, but for eight neurotypical children who were tested at the Lancaster University BabyLab due to COVID-19 restrictions, it was completed by caregivers. Autistic children were significantly older, $t(32) = -6.41$, *p*

<.001, *d* = 2.20, and had significantly higher CARS scores, *t*(32) = -7.00, *p* <.001, *d* = 2.40, than the neurotypical children.

Autistic (*M* age equivalent = 60.24 months, *SD* = 21.85) and neurotypical children (*M* age equivalent 61.29 months, *SD* = 25.14) did not significantly differ on receptive vocabulary abilities as measured by the British Picture Vocabulary Scale 2 (BPVS; Dunn et al., 1997; *t*(32) = 0.13, *p* = .90). The groups were matched on receptive vocabulary abilities because they demonstrate children's aptitude for learning word-referent relationships (Bion et al., 2013, Rothwell et al., accepted, see Study 2). Children's expressive vocabulary abilities were measured using the Expressive Vocabulary Test 2 (EVT; ASD: *M* age equivalent = 53.38 months, *SD* = 24.46; TD: *M* age equivalent 61.47 months, *SD* = 21.97; Williams, 2007), or the expressive language module of the Mullen's Scales of Early Learning (Mullen, 1995) for children who scored below the baseline on the EVT, and did not significantly differ between groups, *t*(31) = 1.00, *p* = .32.

Children's non-verbal intellectual abilities were measured using the Leiter-3 (Roid et al., 2013). The neurotypical group's average non-verbal IQ score (*M* = 104.77, *SD* = 11.45) was significantly higher than the autistic group's (*M* = 80.79, *SD* = 15.91), *t*(25) = 4.46, *p* <.001, *d* = 1.73. Scaled IQ scores could not be calculated for four neurotypical children as they were below three years of age. However, autistic children (*M* = 64.71, *SD* = 18.78) and neurotypical children (NT: *M* = 59.18 months, *SD* = 18.16) did not significantly differ on their Leiter-3 raw scores *t*(29) = -0.83, *p* = .41, suggesting that their non-verbal cognitive abilities were comparable at time of testing (when age was not considered). To assess attentional behaviours, the Conner's Teacher Rating Scale (CTRS-15; Pupura & Lonigan, 2009) was completed by children's class teachers, or the caregivers of the eight neurotypical children who were tested in our BabyLab. The mean raw scores for the autistic children (*M* =

13.82, *SD* = 8.52) and neurotypical children (*M* = 11.88, *SD* = 5.68) did not significantly differ, (*t*(32) = -0.78, *p* = .44). The Repetitive Behaviour Questionnaire was completed by the participants' caregivers to assess the extent of their restrictive and repetitive behaviours (RBQ; Leekam et al., 2007). Autistic children (*M* score = 42.12*, SD* = 8.65) had significantly higher scores on the RBQ than neurotypical children (*M* score = 26.12, *SD* = 5.52), *t*(32) = -6.43, *p* <.001, *d* = 2.21.

Due to school closure during the COVID-19 pandemic, three autistic participants did not complete the Leiter-3 and one autistic participant did not complete the EVT-2. These participants were retained in the study as they completed all other assessments and experimental tasks. An additional four participants were excluded from the study; one neurotypical participant who was unable to complete the touch-screen task, one neurotypical participant who scored above the 'low to minimal symptoms' threshold on the CARS-2, and two autistic children who did not complete both experimental conditions due to school closures during the pandemic.

All procedures performed in this study involving human participants were in accordance with the ethical standards of institutional and national research committees. Informed consent was obtained from caregivers prior to children's participation and a debrief was provided after participation.

Table 1

Characteristics of autistic and neurotypical Participants (SD and Ranges in Parentheses)

Pop.	N	Gender	Chron. Age (*M*, months)	BPVS. Age equiv. (*M*, months)	Express. Lang. age equiv. (*M*, months)	CARS raw score (*M*)	Leiter-3 raw score (*M*)	CTRS raw score (*M*)	RBQ raw score (*M*)
NT	17	10 females, 7 males	52.12 (17.81, 28-94)	61.29 (25.14, 36-101)	61.47 (21.97, 35-104)	16.88 (2.72, 15-24)	59.18 (18.16, 40-102)	11.88 (5.68, 3-26)	26.12 (5.52, 20-35)
ASD	17	3 females, 14 males	94.82 (20.90, 66-132)	60.24 (21.85, 29-97)	53.38 (24.46, 5-82)	29.74 (7.07, 19-43)	64.71 (18.78, 40-106)	13.82 (8.52, 5-31)	42.12 (8.65, 30-59)
Group comparison t-test (*p*)			<.001	.90	.32	<.001	.41	.44	<.001

Note. NT: neurotypical; ASD: autism spectrum disorder; BPVS: British Picture Vocabulary Scale, CARS: Childhood Autism Rating Scale, CTRS: Conner's Teacher Rating Scale, RBQ: Repetitive Behaviour Questionnaire. Participants experienced both conditions.

Materials

The study was administered via a touch-screen computer running MATLAB. Audio stimuli for the word learning task included eight two-syllable unfamiliar words (hizzard, jefa, coodle, nelby; koba, adet, plumbus, erag) selected from the NOUN database (Horst & Hout, 2016) and other academic sources. Visual stimuli included high-resolution colour photographs of twelve unfamiliar objects (see Figure 1) and twenty-two familiar objects, all presented on a grey background. All photographs were approximately 6cm^2, and 500 x 500 pixels when displayed on the screen. Unfamiliar objects were selected on the basis that children would not be familiar with their pre-existing linguistic labels. Familiar objects were selected on account that most children understand their linguistic labels by 16 months (Fenson et al., 1994). Pictures of six familiar objects were employed in the warm-up trials (tree, clock, light, train, key, swing). Images of 16 familiar objects were presented during referent selection trials in the liked and disliked conditions. These were divided into two sets and counterbalanced across conditions (1. book, highchair, table, bowl, doll, flower, brush, car; 2. aeroplane, cot, blanket, tv, shoe, block, hoover, cup). Familiar objects allocated to the two conditions were matched on mean comprehension age (13.5 months for both) and frequency of objects belonging to particular categories (e.g. toys, furniture). Familiar objects within each set were divided into pairs and presented alongside an unfamiliar object in referent selection trials. In every trial type, three pictures were presented side by side. Combinations of objects presented together were phonologically distinct in their labels, and clearly contrasted in shape and colour.

Stimuli names were recorded by a female speaker from the local area and presented through the computer's integrated speakers. The audio files were recorded using a Sony ECM-MS907 Digital Microphone and the software Audacity 2.2.2. Auditory stimuli were edited for timing and clarity, and the volume of all files was normalised. The carrier phrases (e.g. "Can you see the [label]", "Touch the [label]") and the labels (e.g. "car", "nelby") were edited separately, so they were all distinct files. However, when the MATLAB program was used to run the experiment, the audio files were presented sequentially. This was to ensure that there were no differences in the carrier phrases that may offer a hint to children regarding the labels that were about to be presented.

Figure 1

Unfamiliar objects used in the word learning task

Three web cameras attached to the left, right, and centre of the screen were used to record participants' visual attention and behaviour during the study. Recording was done using the 'Open Broadcaster Software' version 23.2.1, which allowed recording from all three cameras simultaneously. The cameras positioned to the left and right of the screen were 15-megapixel Logitech C920 HD Pro Webcams and recorded at a rate of 30 frames-per-second. The centre camera was built into the iMac (1.2 megapixels) and also recorded at 30 frames-per-second. The red recording lights were obscured from participants using black tape to avoid distraction.

Procedure

The task in the present study was very similar to that in Studies 1 and 2. Participants were tested individually in their own school or nursery, or in our BabyLab, and were accompanied by a familiar adult when required. Children were assessed using the BPVS, EVT or MSEL, and Leiter-3 by the researcher over multiple sessions on different days. Children also completed a preliminary 'object preference task' to ascertain which objects would be presented to them in the liked and disliked conditions of the word learning task. Children completed two within-subjects conditions of the word learning task – liked and disliked objects – administered on different days (average of seven days apart, order counterbalanced). The word learning task was delivered via a touch-screen computer. Children were seated approximately 50-70cm away from the screen on a height adjustable chair. The word learning task was the same as Rothwell et al. (accepted, see Study 2) except for the visual and audio stimuli displayed. The task consisted of the following stages, presented in a fixed order: 1. Warm-up trials, 2. Referent selection trials, 3. Five-minute delay, 4. Retention trials, 5. 24-hour delay, 6. Retention trials (see Figure 2).

Object preference task

The experimenter presented children with twelve novel objects, one by one. Children were asked to hold and look at each object for approximately 15 seconds, and then place it down. When handing over the objects, the experimenter told the child to "Look at this toy and see what you think of it. Then I will tell you to put it down and pass you the next toy to look at". While children were handling the objects, the experimenter remained quiet, so the child was not distracted. If children were not looking at or holding the objects, the experimenter prompted them by saying "Please pick up the toy and look at it", handing the child the object again if necessary. Children were praised during the task for good behaviour and following the instructions.

Once children had examined all 12 objects, they were next shown two boxes – a 'good' box and a 'bad' box. The good box was green, with green smiley faces on each side, and contained toy cars and dinosaurs. The bad box was red, with red sad faces on each side, and contained 'rubbish' (e.g. empty crisp packets, toilet roll tubes, and scrunched-up paper). The experimenter communicated the difference between the boxes by showing children what was inside (e.g. "This is the bad box. This is for the worst things, or for things we don't like and don't want to play with" and "This is the good box. This is for the best things, or for things we like and want to play with"). The experimenter confirmed whether children understood the distinction between the boxes by asking them to choose which box a toy car and a scrunched-up piece of paper should be placed in. If a child placed these items in the correct boxes (i.e. the car in the 'good' green box and the paper in the red 'bad' box), we inferred that they associated the two boxes with appropriate concepts and progressed with the task. If children placed the items in the incorrect box, the experimenter provided corrective feedback (e.g. "I think this is rubbish, so it would go in the red box for things we don't like") and demonstrated the correct response. The experimenter then gave the child another try at sorting the example items into the correct boxes.

Then, children were asked to "put the four best objects that you would most like to play with in the green box" and "put the four worst objects that you would least like to play with in the red box." These two sets of four novel objects were assigned to the participant's "liked" and "disliked" conditions respectively. Thus, the novel stimuli that participants experienced in each condition of the word learning task was dependent on their individual preferences.

Figure 2

Examples of trial types in the word learning task

Warm up trials

Prior to the beginning of the study, children saw a cartoon image of a hand that appeared in each of three touch-screen panels, one by one. The experimenter asked the child to “Put their hand on the picture” to enable them to become comfortable with touching the screen. Then, children completed three warm-up trials. Children were instructed to “Put your hand on the picture that the computer asks you to”. During warm-up trials, children were presented with images of three familiar objects in the left, middle, and right panels of the touchscreen. After 2 seconds, participants heard “Look, ‘*2 s gap*’ [label]!”, ‘*1 s gap*’, “Can you see the [label]?” ‘*1 s gap*’, “Touch the [label]!”. Children then had 12 seconds to respond. The same instructions repeated up to six times if children did not respond. Responses were accepted only after the first label utterance, preventing children from advancing through trials without hearing the requested labels. Consequently, children who took longer to respond heard more repetitions of the label (this factor is examined in our analyses). Children received feedback when they made their selection; either audio praise if they responded correctly (e.g. “Well done, you touched the [label]!”) or corrective feedback if they responded incorrectly (“Actually, this is the [label]. Can you touch the [label]?”). Following incorrect responses, the correct referent was highlighted by a green border and children could retry up to five times. Children in both groups responded significantly above chance levels at first attempt on the warm-up trials ($M_{NT} = 1.00$, $M_{ASD} = 0.96$) demonstrating that they understood the task requirements and knew the requested labels. The location and order of requested objects were counterbalanced across participants.

After the warm-up trials, children were video recorded to measure their visual attention. To assist with coding, LEDs on the three video cameras flashed to signify the start of the experiment, transitions between trials, and when participants touched the screen. However, the LEDs were invisible to participants as they were covered with black tape.

Referent selection trials

Following the warm-up, children completed eight referent selection trials. These followed the same format, except children did not receive feedback following their responses. Four novel words were taught via a fast-mapping paradigm similar to that used in Rothwell et al. (accepted, see Study 2). Children viewed four sets of pictures (each containing one unfamiliar object and two familiar objects). The unfamiliar objects depicted the participants four most liked or disliked novel objects, depending on the condition. Each set was presented twice; on one trial the novel picture was requested (novel name trial: “Look, erag! Can you see the erag? Touch the erag!”), and on another trial a familiar picture was requested (familiar name trial: “Look, doll! Can you see the doll? Touch the doll!”). Familiar name trials were included to ensure participants did not demonstrate a novelty preference and encourage them to examine every item in the array, which is imperative as fast mapping requires children to attend to known competitors and exclude them as referents for a novel word (Halberda, 2003). Novel name trials promoted active learning of new word-object pairings; participants used their pre-existing labels for familiar stimuli to decipher the referent of the novel label by applying the mutual exclusivity principle.

Trial order was pseudo-randomised with the constraints that the same set of pictures, or the same trial type (familiar name or novel name), were never presented on more than two trials sequentially. Positioning of objects on the screen (left, middle, right) was pseudo-randomised across trials with the constraint that the target object did not appear in the same location more than twice consecutively. The eight novel words were divided into two sets (1. hizzard, jefa, coodle, nelby; 2. koba, adet, plumbus, erag), and were counterbalanced across conditions. Novel words were pseudo-randomly allocated to the novel objects, so different novel words represented different novel objects. Familiar objects were divided into two sets of eight to obtain a degree of control, but these were also counterbalanced across conditions.

5-minute delay

Immediately after referent selection, children engaged in an unrelated task, such as colouring, for 5-minutes. None of the familiar or unfamiliar experimental stimuli were visible during this stage.

Retention trials

After five minutes, children completed one warm-up trial to re-engage their attention (exactly as described above). Eight retention trials immediately followed (see Figure 2 for an illustration of each trial type). Each novel word was tested on two retention trials and served as a foil on four trials. These trials assessed whether children's retention of newly mapped word-referent associations differed between stimuli they liked, compared to those they disliked. Trial order was pseudo-randomised, ensuring that the same set of objects was never presented on more than two trials sequentially. Positioning of objects on the screen (left, middle, right) was pseudo-randomised across trials with the constraint that the target object did not appear in the same location more than twice consecutively.

24-hour retention trials

After a 24-hour delay, children completed a second block of eight retention trials. Due to practical constraints, not all children experienced exactly a 24-hour delay, with the average delay being 24.0 hours (range 23.2 – 26.0 hours). These retention trials were preceded by three warm-up trials (as described above) to remind children of the task requirements and how to respond. The 24-hour retention trials were identical to the 5-minute retention trials with the exception that stimuli were presented in different orders and combinations.

Coding and data cleaning

As described in Rothwell et al. (Study 2), videos were coded using the software Blender 2.78, with a customised version of the python script ultra-coder added on (see https://github.com/dmbasso/misc-blender-tools/blob/master/ultra_coder.py for original). Coders were blind to the location of the target stimuli on each trial. Children's visual fixations were coded frame-by-frame with a precision of 16.7ms, and looks were coded as left, right, centre, away, or not visible. The LEDs that flashed to signify the beginning of the experiment and transitions between trials, as well as participant touches, were imperative for coding. Of the 200 videos recorded across the three experimental stages, 25% were reviewed by two independent coders. Coders agreed on looking direction on more than 97% of the frames. A custom Matlab programme then calculated the primary dependent variables (see Table 2). These variables were calculated 233ms after the label onset to allow for saccade initiation latencies (Swingley et al., 1999; Swingley, 2009).

As in Rothwell et al. (Study 2), three distinct looking measures were examined from our coded videos. ‘Proportion of looking towards the target’ is a frequently used looking measure in both the neurotypical and autism word learning literature, particularly in the absence of an explicit behavioural response (e.g. Akechi et al., 2011, 2013; Ackermann et al., 2020; Bion et al., 2013; Potrzeba et al., 2015). This measure captures looking behaviour to the target relative to competitors. ‘Number of looks towards the target’ was utilised as a measure of children’s visual exploration of stimuli across the scene, a feature that commonly differs in autism in comparison to the neurotypical population (e.g. Sasson et al., 2008, 2011). Children’s longest singular look was also measured; either ‘Longest look to novel or familiar stimuli’ at referent selection to examine children’s attention to novel stimuli during encoding, or ‘Longest look to target or foil’ to decipher whether children allocated their longest look to the requested referent during retention phases. It offered a categorical indication of children’s attentional preference based on a singular look, which can be comparable to behavioural accuracy (Ambridge & Rowland, 2013). The longest look variable also accounts for the likelihood that children’s attention will decrease over the trial (Bailey & Plunkett, 2002; Ma et al., 2011; Schafer & Plunkett, 1998).

Table 2

Looking measures recorded during the experimental tasks

Looking Measure	Coding Scheme	Experimental Phase	Outcome
Proportion of time spent looking at the target stimuli	0 to 1	Referent selection, 5-minute retention, 24-hour retention	Proportion of time spent looking at the target stimuli, compared to the proportion of time spent looking at the distractor stimuli.
Number of looks to the target stimuli	0-12	Referent selection, 5-minute retention, 24-hour retention	Every new look towards the target stimuli was counted as a new look.
Longest look to novel or familiar stimuli	0 (familiar) or 1 (novel)	Referent selection	If the longest look was to the novel stimuli, or familiar stimuli.
Longest look to target or foil stimuli	0 (foil) or 1 (target)	5-minute retention, 24-hour retention	If the longest look was to the target stimuli, or the foil stimuli.

4.5. Results and Discussions

Due to the variety of measures and the comprehensive nature of our analyses, we present and discuss our findings in two sections. The first section analyses and interprets children's word learning accuracy and response time data. The second section compares the groups' visual attention at each stage of the word learning task and assesses whether looking variability predicted word learning outcomes.

4.5.1. Do autistic and neurotypical children differ in accuracy and response times when learning names for liked and disliked stimuli?

Accuracy and response time data were analysed via mixed-effects models using the glmer and lmer functions from the lme4 package in R (Bates et al., 2015). Population was contrast coded as -0.5 (neurotypical) and 0.5 (autistic). Condition was coded as -0.5 (disliked) and 0.5 (liked). Trial type was coded as -0.5 (familiar) and 0.5 (novel). By-word referent selection accuracy was coded as -0.5 (incorrect) and 0.5 (correct) when included as a fixed effect in retention accuracy analyses. Total referent selection accuracy for novel trials was coded as 0-4. Number of labels heard at referent selection for each novel word was coded as 1-5 (autistic $M = 2.04$, $SD = 0.79$; neurotypical $M = 1.70$, $SD = 0.80$). By-word 5-minute retention accuracy was coded as 0-2. Trial-level accuracy as a dependent measure was coded as 1 (correct) or 0 (incorrect) for all analyses.

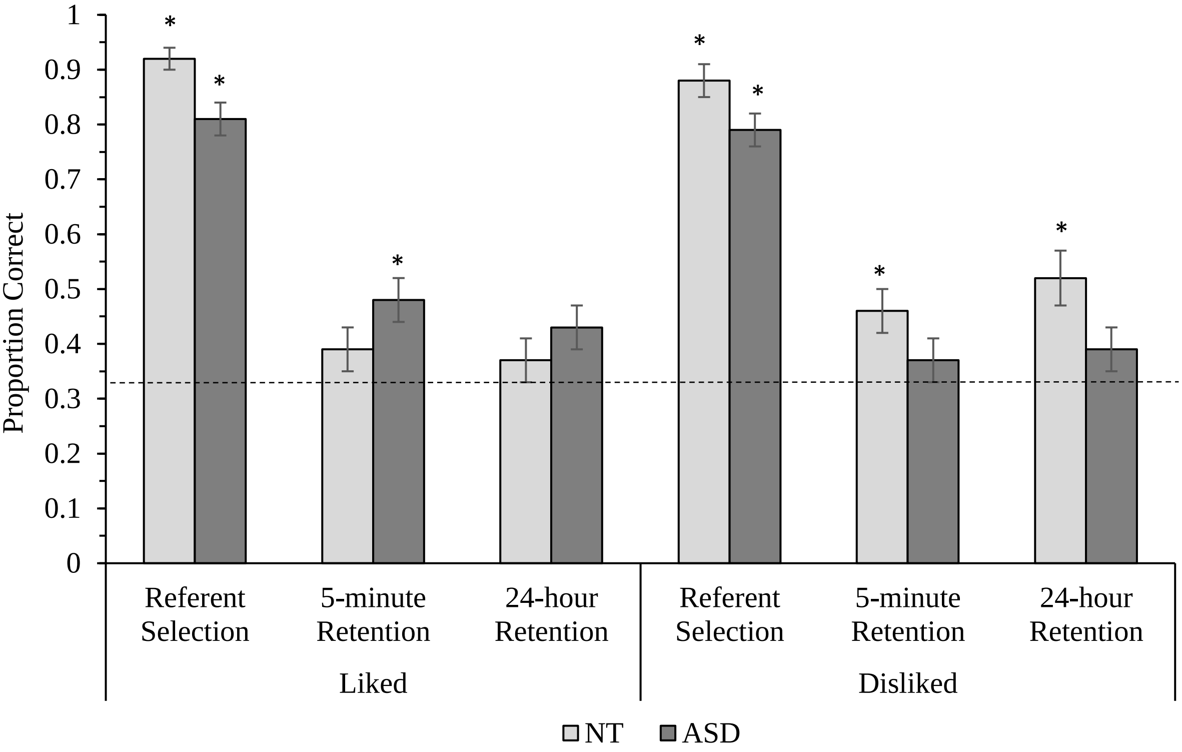
The likelihood of children responding correctly by chance on each trial was 33%. All models were built up sequentially, adding in one fixed effect at a time and comparing each model with the previous best-fitting model using log-likelihood tests. Each analysis started with a baseline model containing by-participant and by-word random intercepts, with a random slope of condition x trial type per participant for referent selection, or condition per participant for retention phases. If some models in a sequence failed to converge, the random effects were simplified until all models in the sequence successfully converged. Only final models are reported; please refer to Appendix D for full details of the model building sequences and analyses of individual differences.

Referent selection accuracy

Referent selection accuracy was analysed via generalised linear mixed-effects models testing the effects of population, condition, and trial type. This analysis contained 542 data points. Two trials for one neurotypical participant were removed due to a technical issue. Descriptive statistics for referent selection accuracy are presented in Figure 3.

Figure 3

*Mean referent selection, 5-minute retention, and 24-hour retention trial accuracy for neurotypical (NT) and autistic children (ASD), error bars show ± 1 SE. Stars above columns indicate where performance was significantly different from chance, indicated by the dotted lines (*p <.05)*



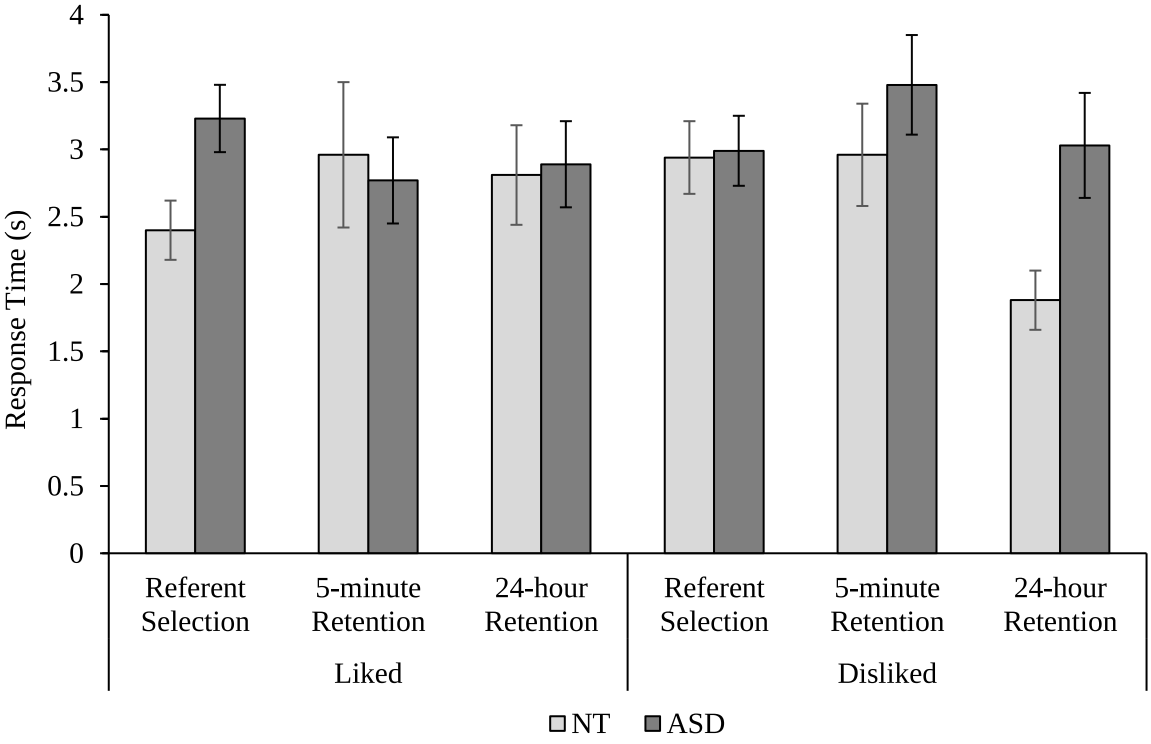
The final model included fixed effects of trial type (*z* = -3.13, *p* =.002) and population (*z* = -2.28, *p* = .022; see Table 3) indicating that children responded with significantly greater accuracy on familiar trials than novel trials, and neurotypical children were more accurate than autistic children. However, it is noteworthy that both groups responded well above chance on novel trials across conditions (neurotypical liked *M* = 0.87, neurotypical disliked *M* = 0.85; autism liked *M* = 0.74, autism disliked *M* = 0.71), demonstrating their effective use of mutual exclusivity.

Referent selection response times

Children's response times for correctly answered referent selection trials were analysed using linear mixed-effects models, testing the effects of population, condition, and trial type. We calculated the average correct response time for each population in each trial type and condition, and removed outliers that were ≥ 3*SD* above the mean for the sub-group (e.g. autistic children in the liked condition responding to novel trials). For autistic children, 216 of 218 (99%) correct responses were included in our analyses. For neurotypical children, 237 of 243 (98%) correct responses were included. With outliers excluded, mean correct response times for each population are reported in Figure 4.

Figure 4

Mean response times on correctly answered referent selection, 5-minute retention, and 24-hour retention trials for neurotypical (NT) and autistic children (ASD), error bars show ± 1 SE



The final model included fixed effects of trial type (*t* = 5.96, *p* <.001) and condition (*t* = 2.06, *p* = .040; see Table 3). Both groups were slower to identify stimuli in the liked condition in comparison to the disliked condition. Children in both populations were slower to generate correct responses for novel trials compared to familiar trials.

Table 3

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) of children's accuracy on referent selection trials, and response times on correctly answered referent selection trials

	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Referent Selection Accuracy	(Intercept)	2.35	0.24	9.94	<.001
	Trial Type	-1.13	0.36	-3.13	.002
	Population	-1.01	0.44	-2.28	.022
		AIC	**BIC**	**logLik**	**deviance**
		437.9	498.0	-204.9	409.9
	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |*t*|)**
Referent Selection Response Times	(Intercept)	3.05	0.32	9.53	<.001
	Trial Type	1.13	0.19	5.96	<.001
	Condition	0.39	0.19	2.06	.040
		AIC	**BIC**	**logLik**	**deviance**
		2003.7	2024.3	-996.9	1993.7

5-minute retention accuracy

Children's retention accuracy after 5-minutes was analysed via generalised linear mixed-effects models testing the effects of population, condition, by-word referent selection accuracy, and number of labels heard at referent selection for the target word. Four trials from the neurotypical group were excluded due to a technical error. The model in these analyses contained 540 data points. The descriptive statistics are reported in Figure 3.

The final model included a fixed effect of by-word referent selection accuracy (*z* = 2.49, *p* = .013; see Table 4), suggesting that both groups were more accurate during 5-minute retention if they were previously more accurate during referent selection.

Table 4

Summary of the fixed effects in the final generalised linear mixed-effects models (log odds) of children's accuracy on 5-minute retention trials

Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
(Intercept)	-0.52	0.15	-3.58	<.001
Referent Selection Accuracy	0.61	0.24	2.49	.013
	AIC	**BIC**	**logLik**	**deviance**
	731.0	756.8	-359.5	719.0

5-minute retention response times

Children's response times for correctly answered 5-minute retention trials were analysed using linear mixed-effects models. Outliers were identified and removed in the same way as described for referent selection trials. The models in these analyses included 114 of 115 (99%) correct responses from autistic children, and 113 of 114 (99%) correct responses from neurotypical children. With outliers excluded, mean correct response times for each population are reported in Figure 4.

The inclusion of fixed effects (population and condition) did not improve model fit.

24-hour retention accuracy

Children's retention accuracy after 24-hours was analysed via generalised linear mixed-effects models testing the effects of population, condition, by-word referent selection accuracy, total referent selection accuracy for novel trials, number of labels heard at referent selection for the target word, and by-word 5-minute retention accuracy (all coded as described previously). Two neurotypical children in the disliked condition, and two participants in the liked condition (one autistic participant, one neurotypical participant) did not complete the 24-hour retention trials due to absence. Due to a technical error, eight additional trials were excluded (four from the autistic group and four from the neurotypical group). The model in this analysis contained 504 data points. Descriptive statistics for 24-hour retention accuracy are presented in Figure 3.

The final model contained fixed effects of by-word referent selection accuracy ($z = 1.86$, $p = .063$; marginally significant) and 5-minute retention accuracy ($z = 2.57$, $p = .010$; see Table 5). Children were more likely to respond accurately at 24-hour retention if they were previously more accurate at referent selection and 5-minute retention.

Table 5

Summary of the fixed effects in the final generalised linear mixed-effects models (log odds) of children's accuracy on 24-hour retention trials

Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
(Intercept)	-0.84	0.20	-4.28	<.001
Referent Selection Accuracy	0.50	0.27	1.86	.063
5-minute Retention Accuracy	0.38	0.15	2.57	.010
	AIC	**BIC**	**logLik**	**deviance**
	668.2	697.7	-327.1	654.2

24-hour retention reaction times

Children's response times for correctly answered 24-hour retention trials were analysed using linear mixed-effects models. Outliers were identified and removed in the same way as for previous analyses. For autistic children, 104 of 106 (98%) correct responses were included in our analyses. For neurotypical children, 105 of 106 (99%) correct responses were included. With outliers excluded, mean correct response times for each population are reported in Figure 4.

The final model was the baseline model – variation in children's correct response times on 24-hour retention trials was not predicted by any fixed effects (population or condition).

4.5.2. Discussion: Accuracy and Response Times

These analyses investigated whether autistic and neurotypical children differ in their ability to identify and retain the meanings of novel words associated with liked and disliked stimuli. Crucially, we examined children's accuracy and response speed across three distinct stages of word learning: referent selection, 5-minute retention, and 24-hour retention after a period of sleep. In comparison to neurotypical controls matched on receptive vocabulary, autistic children were *less* accurate to spontaneously identify the meaning of novel words using ME. Despite this, both groups recalled word-referent mappings after 5 minutes and 24 hours with comparable accuracy. Whether stimuli were liked or disliked did not influence word learning accuracy in either group. Furthermore, autistic and neurotypical children's response times did not significantly differ at any of the three word learning phases.

Our results revealed that autistic children were less accurate than neurotypical peers at utilising ME-based referent selection to identify the meanings of novel words across conditions. Some research suggests that autistic children can perform ME-based fast mapping with comparable accuracy to language matched neurotypical children (e.g. Carter & Hartley, 2021; Preissler & Carey, 2005; Rothwell et al., accepted, see Study 2), whilst other studies show population differences in this process (Hartley et al., 2019; Mathée-Scott, et al., 2021). However, even in studies showing reduced ME accuracy for autistic children, they often still demonstrated above-chance use of ME (e.g. Hartley et al., 2019). These population distinctions may be explained by differences in attentional mechanisms. Neurotypical children can easily and flexibly shift their attention between stimuli and engage with a multitude of environmental features (Landry & Bryson, 2004). However, autistic children often show more inflexible attentional mechanisms, perseverating on specific topics or features within a learning context (American Psychiatric Association, 2013; Sasson et al., 2008, 2011; Venker, 2018). As such, autistic children's differences in attentional mechanisms may disrupt ME in word learning, as they are more likely to fixate on stimuli that are not the intended target. Despite this finding, our two groups' retention accuracy did not significantly differ, so it is apparent that the autistic children still successfully learned through ME-based referent selection.

After a 5-minute delay, both groups retained novel words with similar accuracy. Congruent with the findings of Rothwell et al. (accepted, see Study 2), we discovered that referent selection accuracy predicted 5-minute retention accuracy. This suggests that the mechanisms underpinning referent selection and retention after a short delay may be less distinct than other studies have proposed (e.g. Carter & Hartley, 2021; Hartley et al., 2019, 2020). Kucker and Samuelson (2012) demonstrated that 2-year-olds can retain words following learning via ME if they play with the objects before referent selection. As the present study allowed children to explore the objects before the word learning task to identify their liked and disliked stimuli, it is possible that exposure to stimuli prior to referent selection could have influenced this finding. However, as our findings are both supported and refuted by other studies, we would suggest that further research is required to investigate under what conditions these mechanisms inter-relate.

Autistic children also retained novel words after 24-hours with comparable accuracy to neurotypical children, suggesting that both short- and long-term word learning mechanisms are fundamentally unimpaired in this population. Interestingly, 24-hour retention accuracy was predicted by referent selection accuracy and 5-minute retention accuracy (see also Rothwell et al., Study 1, Study 2, accepted, for similar results). The dynamic associative account of word learning posits that both 'fast' and 'slow' word learning mechanisms combine to support accurate and robust word learning (McMurray et al., 2012). This theory proposes that referent selection and learning interrelate through children maintaining correct word-referent associations following identification and eliminating incorrect referents. Whilst some studies suggest that these mechanisms are somewhat distinct (e.g. Horst & Samuelson, 2008), the present study highlights that fast- and slow- word learning stages are inter-related. Given autistic children's lower response accuracy at referent selection, our data may indicate that they retained a higher proportion of words that were successfully disambiguated at referent selection, and thus the mechanisms may be more strongly inter-related in this population. It is therefore essential that future work investigates how we can improve encoding at referent selection to scaffold autistic children's vocabulary development.

Prior research investigating the role of sleep on novel word learning focuses primarily on linguistically able autistic children (e.g. Fletcher et al., 2020; Henderson et al., 2014). However, here we discovered that language delayed autistic children performed with comparable accuracy to their neurotypical peers at both 5 minutes and 24 hours - sleep consolidation did not significantly affect the retention accuracy of either population. By contrast, in a similar word learning task, Rothwell et al. (Study 1, Study 2, accepted) discovered that autistic children received a greater benefit of sleep consolidation than neurotypical children when recalling novel names of object stimuli in comparison to animal stimuli. In Rothwell and colleagues' previous studies, the presence of interesting animal stimuli initially aided more robust encoding for retrieval after 5 minutes but sleep then helped to encode the less interesting object stimuli representations for better retrieval after 24 hours. In the present study, it is possible that the absence of animals reduced autistic children's overall investment in the task during referent selection, possibly explaining why their 24-hour retention did not receive the same benefits of sleep as in Rothwell et al. (accepted, Study 2, Study 3). However, this suggestion is speculative and, given the contrast in findings to previous studies, further research is required to understand the cause of these findings.

Investigating how quickly children generated correct responses highlights the pace of children's information processing throughout word learning. During referent selection, children were significantly slower to generate correct responses during novel trials compared to familiar trials. As familiar trials required children to simply select referents they already knew the names of, this finding was unsurprising. To decipher a correct novel referent, children had to employ ME, attending to an array of stimuli, eliminating familiar competitors, and selecting the correct novel referent (Halberda, 2006). As novel referent selection requires greater cognitive demands, it is unsurprising that children in both populations were slower to make their selections on the novel trials (Bion et al., 2013; Rothwell et al., accepted, see Study 2). However, children in both populations were slower to identify liked stimuli at referent selection. It is probable that such stimuli captured children's attention due to their heightened interest, and they therefore chose to spend longer studying items in the array before identifying referents. Interestingly, unlike findings from previous research (e.g. Hilton & Westermann, 2017; Rothwell et al., Study 1, Study 2) longer time spent studying stimuli at encoding did not benefit subsequent retention. However, subsequent analyses of visual attention data in part two of this study (section 4.5.3) will assess this hypothesis.

Autistic children were not significantly slower than neurotypical children to generate correct responses during any of the word learning stages. Limited research thus far has investigated autistic children's response times during referent selection and retention, so it is unclear how response times are affected by experimental features. Some previous research suggests that autistic children are slower to generate correct responses than neurotypical children when recalling word meanings at 5-minute retention (e.g. Hartley et al., 2020), whereas others find this difference depends on certain stimulus types and experimental stages (e.g. Rothwell et al., accepted, see Study 2). In Rothwell et al. (Study 1, Study 2, accepted), we found that autistic children were significantly slower to generate correct responses when categorically interesting animal stimuli were present, likely due to increased attentional allocation to such stimuli. These longer response times did, however, have the potential to aid word learning in autistic individuals. By contrast, in the present study, autistic children were not significantly slower than neurotypical children to generate correct responses at referent selection. However, autistic children were significantly less accurate at referent selection than neurotypical children. Taken together, these findings highlight the heterogeneity of autistic children's response times across studies, tasks, and samples. As such, further research is required to understand how environmental conditions influence the speed of information processing during word learning, and its relationship to accuracy.

Overall, these analyses revealed that stimulus condition determined response speed at referent selection, but not accuracy at referent selection or retention in either population. These findings suggest that children in both populations did not demonstrate a strong enough preference for their chosen novel object stimuli to afford differences in novel word retention. Although children may have identified idiosyncratic preferences for the novel objects, it may be that categorical interests (e.g. animals) have the stronger influence on word learning outcomes (e.g. Ackermann et al., 2020; Rothwell et al., accepted, see Study 2). If this is the case, it is possible that children spent similar amounts of time studying liked and disliked novel stimuli across the word learning stages, perhaps explaining why retention was not affected by condition. Additionally, these findings do not preclude the possibility that individual differences in attention to specific objects within or across conditions impacted word learning accuracy. Therefore, the following gaze analyses examining children's in-trial attention, and how this predicted performance at each word learning stage, aim to address these gaps in knowledge.

4.5.3. Does autistic and neurotypical children's visual attention differ while learning names for liked and disliked stimuli, and does variability in visual attention predict their learning accuracy?

In the following analyses we examine whether autistic and neurotypical children differed in their looking behaviour during each stage of the word learning task. To elucidate the relationship between visual attention and word learning performance, we also investigated how variability in children's in-trial looking behaviour, and looking behaviour during referent selection, predicted their response accuracy.

All models were conducted using the glmer and lmer functions from the lme4 package in R (Bates et al., 2015). Population, condition, and trial type were coded as described previously. Proportion of time spent looking at the target object on each trial was scored between 0 and 1. Number of looks to the target object on each trial ranged from 0 to 12, with every new look towards the target stimuli being counted. Longest look to novel or familiar stimuli was coded as 0 (familiar) or 1 (novel) for analyses examining between-group differences in this measure, and -0.5 (longest look to familiar object) or 0.5 (longest look to novel object) when included as a fixed effect in analyses predicting accuracy. This variable was only included in referent selection analyses, as the retention phases only involved novel objects. Longest look to target or foil stimuli was coded as 0 (foil) or 1 (target) for analyses investigating between-group differences in this measure at retention, and -0.5 (longest look to foil stimuli) and 0.5 (longest look to target stimuli) when included as a fixed effect in analyses predicting accuracy.

All analyses were undertaken following the same modelling procedures described for the accuracy and response time data. Each analysis started with a baseline model containing by-participant and by-word random intercepts, with a random slope of condition x trial type per participant for referent selection, or condition per participant for retention phases. As only final models are reported, please refer to Appendix D for full details of the model building sequences.

Referent selection

Linear mixed-effects models (unless otherwise specified) tested whether effects of population, condition, and trial type predicted variability in each visual attention measure during referent selection (see Table 6 for descriptive statistics). Generalised linear mixed-effects models tested whether children’s in-trial visual attention predicted their behavioural response accuracy at referent selection. These models contained 542 data points (exclusions for both model types were the same as described for the previous accuracy analyses).

Table 6

Mean values for the individual visual attention measures at referent selection, split by condition, trial type, and population

Looking Measure	Condition	Trial Type	ASD	NT
Proportion of time spent looking at target stimuli (0-1)	Liked	Familiar	0.64	0.67
		Novel	0.49	0.55
	Disliked	Familiar	0.67	0.67
		Novel	0.45	0.52
Number of looks towards target stimuli (0-8)	Liked	Familiar	1.91	1.47
		Novel	2.34	1.87
	Disliked	Familiar	1.69	1.38
		Novel	1.78	1.57
Longest look to novel or familiar stimuli (0/1)	Liked	Familiar	0.10	0.13
		Novel	0.60	0.63
	Disliked	Familiar	0.15	0.12
		Novel	0.49	0.58

Proportion of time spent looking at the target

Proportion of time spent looking at the target referent was predicted by a fixed effect of trial type (t = -6.11, p <.001; see Table 7). Children in both groups looked longer at the target during familiar trials than novel trials.

Referent selection accuracy was predicted by a fixed effect of the looking measure (z = 8.60, p <.001; see Table 7). Across populations and conditions, as children's proportion of looking towards the target increased, so too did their referent selection accuracy.

Table 7

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) for proportion of time spent looking at the target stimuli during referent selection

	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
Between-group Differences	(Intercept)	0.58	0.02	34.63	<.001
	Trial Type	-0.16	0.03	-6.11	<.001
		AIC	**BIC**	**logLik**	**deviance**
		68.7	90.2	-29.4	58.7
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	-2.57	0.53	-4.86	<.001
	Proportion of time spent looking at the target	12.12	1.41	8.60	<.001
		AIC	**BIC**	**logLik**	**deviance**
		245.9	301.8	-110.0	219.9

Number of looks to the target

Number of looks to target stimuli was predicted by fixed effects of population (t = 2.29, p = .029), trial type (t = 3.17, p = .002), and condition (t = 3.38, p <.001; see Table 8). Across conditions, autistic children made more looks to target stimuli than neurotypical children, and children in both groups made more looks towards the target during novel trials than familiar trials, and more looks towards the target in the liked condition compared to the disliked condition.

Referent selection accuracy was predicted by a looking measure x population interaction (z = -3.30, p <.001; see Table 8). This interaction was deconstructed by testing the looking measure effect on autistic and neurotypical children separately. Across conditions, children who made more frequent looks towards the target during referent selection responded more accurately, but this effect was stronger for the neurotypical group (*z* = 4.65, *p* <.001) than the autistic group (*z* = 3.41, *p* <.001).

Table 8

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) for number of looks at the target stimuli during referent selection

	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
Between-group Differences	(Intercept)	1.75	0.08	22.32	<.001
	Population	0.36	0.16	2.29	.029
	Trial Type	0.27	0.09	3.17	.002
	Condition	0.29	0.09	3.38	<.001
		AIC	**BIC**	**logLik**	**deviance**
		1595.7	1621.4	-791.8	1583.7
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	0.40	0.38	1.04	.30
	Number of looks to target	1.59	0.31	5.08	<.001
	Population	1.13	0.76	1.49	.14
	Number of looks x Population	-2.06	0.63	-3.30	<.001
		AIC	**BIC**	**logLik**	**deviance**
		393.1	457.6	-181.6	363.1

Longest look to novel or familiar stimuli

This variable was analysed via generalised linear mixed effects models. Whether children's longest look was towards novel or familiar stimuli was predicted by a fixed effect of trial type (*z* = 8.17, *p* <.001; see Table 9). Both autistic and neurotypical children looked longer at the novel object on novel trials than familiar trials.

Referent selection accuracy was predicted by a looking measure x trial type interaction (*z* = 7.80, *p* <.001; see Table 9). This interaction was deconstructed by testing the looking measure effect on familiar and novel trials separately. Accuracy on novel trials significantly increased when children's longest look was towards the novel stimuli (*z* = 5.74, *p* <.001). However, on familiar trials, accuracy significantly decreased when children's longest look was towards the novel object (*z* = -5.67, *p* <.001).

Table 9

Summaries of the fixed effects in the final generalised linear mixed-effects models (log odds) for the longest look to novel or familiar stimuli during referent selection

	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Between-group	(Intercept)	-0.88	0.14	-6.12	<.001
Differences	Trial Type	2.45	0.30	8.17	<.001
		AIC	**BIC**	**logLik**	**deviance**
		593.3	649.1	-283.6	567.3
	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Predicting	(Intercept)	2.08	0.27	7.76	<.001
Accuracy	Longest look novel or familiar	-0.27	0.36	-0.75	.45
	Trial Type	-0.23	0.43	-0.54	.59
	Longest look x Trial Type	5.58	0.72	7.80	<.001
		AIC	**BIC**	**logLik**	**deviance**
		373.0	437.5	-171.5	343.0

5-minute retention

Linear mixed-effects models (unless otherwise specified) tested whether effects of population and condition predicted variability in each looking measure at 5-minute retention (see Table 10 for descriptive statistics). Generalised linear mixed-effects models tested whether children's in-trial looking, and looking at referent selection, predicted their behavioural response accuracy at 5-minute retention. All models contained 542 data points, with two trials excluded due to a computer error, as described in the accuracy analyses.

Table 10

Mean values for the individual visual attention measures at 5-minute retention, split by condition and population

Looking Measure	Condition	ASD	NT
Proportion of time spent looking at the target stimuli (0-1)	Liked	0.40	0.39
	Disliked	0.33	0.40
Number of looks towards the target stimuli (0-12)	Liked	1.74	1.31
	Disliked	1.34	1.32
Longest look to target or foil stimuli (0/1)	Liked	0.43	0.37
	Disliked	0.34	0.41

Proportion of time spent looking at the target stimuli

This looking measure did not significantly differ between populations or conditions; the inclusion of fixed effects did not improve fit in comparison with the baseline model.

Children's 5-minute retention accuracy was predicted by a fixed effect of this looking measure ($z = 12.03$, $p <.001$; see Table 11). Across groups and conditions, children who looked proportionately more towards the target during referent selection responded more accurately.

Variability in proportion of time spent looking at the target stimuli during referent selection predicted 5-minute retention accuracy ($z = 1.79$, $p = .074$; see Table 11). Children who looked proportionately longer at target stimuli during referent selection responded more accurately at 5-minute retention across groups and conditions. However, given that the model comparison (see Appendix D) and fixed effect within the model approached significance, this result should be treated cautiously.

Table 11

Summaries of the fixed effects in the final generalised linear mixed-effects model (log odds) of 5-minute retention accuracy, predicted by proportion of time spent looking at the target stimuli

	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	-3.37	0.29	-11.57	<.001
	Proportion	7.84	0.65	12.03	<.001
		AIC	**BIC**	**logLik**	**deviance**
		406.9	432.7	-197.5	394.9
	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Referent Selection Looking Predicting Accuracy	(Intercept)	-0.67	0.23	-2.98	.003
	Proportion of time looking at target during referent selection	0.66	0.37	1.79	.074
		AIC	**BIC**	**logLik**	**deviance**
		735.8	761.6	-361.9	723.8

Number of looks to the target stimuli

This looking measure was predicted by a population x condition interaction which approached significance (t = 1.86, p = .071; see Table 12). This interaction was deconstructed by testing the effect of condition on autistic and neurotypical participants separately, and the effect of population on liked and disliked conditions separately. Autistic children made significantly more looks in the liked condition compared to the disliked condition (t = 3.02, p = .003), but neurotypical children did not significantly differ in their looking between conditions (t = -0.08, p = .93). The populations differed in their number of looks across conditions in the liked condition (t = 2.09, p = .045) but not the disliked condition (t = 0.07, p = .95).

Children's 5-minute retention accuracy was predicted by a looking measure x population interaction (z = -1.84, p = .065; see Table 12). This interaction was deconstructed by testing the looking measure effect on autistic and neurotypical children separately. Across conditions, children who made more frequent looks towards the target during referent selection responded more accurately, but this effect was stronger for the neurotypical group ($z = 5.70$, $p < .001$) than the autistic group ($z = 4.84$, $p < .001$).

Variability in number of looks towards the target stimuli during referent selection predicted 5-minute retention accuracy ($z = 1.85$, $p = .064$; see Table 12). Children who looked more frequently at target stimuli during referent selection responded more accurately at 5-minute retention across groups and conditions.

However, given that the models for number of looks at 5-minute retention provided a borderline significant improvement in fit compared to the baseline model (see Appendix D), alongside approaching significance fixed effects and interaction effects, we would suggest these interpretations should be taken cautiously.

Table 12

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) for 5-minute retention, predicted by number of looks to the target stimuli

	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
Between-Group Differences	(Intercept)	1.43	0.09	15.06	<.001
	Population	0.22	0.18	1.26	.22
	Condition	0.20	0.11	1.74	.09
	Population x Condition	0.42	0.23	1.86	.071
		AIC	**BIC**	**logLik**	**deviance**
		1842.8	1881.4	-912.4	1824.8
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	-1.49	0.21	-7.05	<.001
	Number of looks	0.84	0.11	7.58	<.001
	Population	0.16	0.40	0.41	.68
	Number of looks x Population	-0.41	0.22	-1.84	.065
		AIC	**BIC**	**logLik**	**deviance**
		668.7	703.1	-326.4	652.7
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Referent Selection Looking Predicting Accuracy	(Intercept)	-0.63	0.20	-3.13	.002
	Number of looks to target at referent selection	0.15	0.08	1.85	.064
		AIC	**BIC**	**logLik**	**deviance**
		735.6	761.4	-361.8	723.6

Longest look to target or foil

This looking measure did not significantly differ between populations or conditions; the inclusion of fixed effects did not improve fit in comparison with the baseline model.

Children’s 5-minute retention accuracy was predicted by this looking measure (z = 12.64, p <.001; see Table 13). Looking longest at the target object (rather than a foil) was associated with significantly greater 5-minute retention accuracy across populations and conditions.

Whether children looked longest at the target or a foil during referent selection did not predict 5-minute retention accuracy.

Table 13

Summary of the fixed effects in the final generalised linear mixed-effects model (log odds) of 5-minute retention accuracy, predicted by longest look to the target/foil stimuli during 5-minute retention

Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
(Intercept)	-0.06	0.15	-0.43	.67
Longest look target	2.96	0.23	12.64	<.001
	AIC	**BIC**	**logLik**	**deviance**
	547.2	573.0	-267.6	535.2

24-hour retention

Linear mixed-effects models (unless otherwise specified) tested whether effects of population and condition predicted variability in each looking measure at 24-hour retention (see Table 14 for descriptive statistics). Generalised linear mixed-effects models tested whether children's in-trial looking, and looking at referent selection, predicted their behavioural response accuracy at 24-hour retention. As described in the accuracy analyses, for both model types, data from one neurotypical and one autistic child in the liked condition, and two neurotypical children in the disliked condition, were excluded from this phase due to absence. Four additional trials from autistic participants, and two trials from a neurotypical participant, were excluded due to computer error. All models contained 506 data points.

Table 14

Mean values for the individual visual attention measures at 24-hour retention, split by condition and population

Looking Measure	Condition	ASD	NT
Proportion of time spent looking at the target stimuli (0-1)	Liked	0.38	0.36
	Disliked	0.39	0.46
Number of looks towards the target stimuli (0-11)	Liked	1.63	1.11
	Disliked	1.40	1.24
Longest look to target or foil stimuli (0/1)	Liked	0.40	0.37
	Disliked	0.38	0.51

Proportion of time spent looking at the target stimuli

This looking measure did not significantly differ between populations or conditions; the inclusion of fixed effects did not improve fit in comparison with the baseline model.

Children's 24-hour retention accuracy was predicted by a fixed effect of this looking measure (z = 11.36, p <.001; see Table 15). Across groups and conditions, as children's proportion of looking at the target object increased, so did their 24-hour retention accuracy.

Variability in proportion of time spent looking at the target stimuli during referent selection predicted 24-hour retention accuracy (z = 1.79, p = .073; see Table 15). Children who looked proportionately longer at target stimuli during referent selection responded more accurately at 24-hour retention across groups and conditions. However, the model comparison (see Appendix D) and fixed effect within the model approached significance, so this interpretation should be treated cautiously.

Table 15

Summaries of the fixed effects in the final generalised linear mixed-effects models (log odds) for 24-hour retention accuracy, predicted by proportion of time spent looking at the target stimuli

	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Predicting Accuracy	(Intercept)	-3.95	0.36	-11.11	<.001
	Prop. of time looking to target	8.71	0.77	11.36	<.001
		AIC	**BIC**	**logLik**	**deviance**
		341.7	367.0	-164.8	329.7
	Fixed effects	**Estimated coefficient**	**Std. error**	*z*	**Pr(> |z|)**
Referent Selection Looking Predicting Accuracy	(Intercept)	-0.73	0.25	-2.91	.004
	Prop. of time looking at target during referent selection	0.74	0.41	1.79	.073
		AIC	**BIC**	**logLik**	**deviance**
		675.7	701.1	-331.9	663.7

Number of looks to the target stimuli

Number of looks to the target was predicted by a fixed effect of population ($t = 2.27$, $p = .030$; see Table 16). Across conditions, autistic children made significantly more looks towards the target than neurotypical children.

Children’s 24-hour retention accuracy was predicted by a fixed effect of the looking measure ($z = 7.55$, $p < .001$; see Table 16). More frequent looks to the target object were associated with significantly higher 24-hour retention accuracy across populations and conditions.

Variability in number of looks towards the target stimuli during referent selection did not predict 24-hour retention accuracy.

Table 16

Summaries of the fixed effects in the final generalised and linear mixed-effects models (log odds) for number of looks towards the target stimuli during 24-hour retention

	Fixed effects	**Estimated coefficient**	**Std. error**	***t***	**Pr(> |t|)**
Between-group	(Intercept)	1.33	0.08	16.18	<.001
Differences	Population	0.37	0.16	2.27	.030
		AIC	**BIC**	**logLik**	**deviance**
		1533.1	1550.0	-762.5	1525.1
	Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
Predicting	Number of looks	-1.61	0.22	-7.16	<.001
Accuracy		0.90	0.12	7.55	<.001
		AIC	**BIC**	**logLik**	**deviance**
		604.3	629.6	-296.1	592.3

Longest look to target or foil stimuli

Generalised linear mixed-effects models revealed that this looking measure did not significantly differ between populations or conditions; the inclusion of fixed effects did not improve fit in comparison with the baseline model.

Children's 24-hour retention accuracy was predicted by a fixed effect of the looking measure ($z = 12.72$, $p < .001$; see Table 17). Looking longest at the target was associated with significantly higher 24-hour retention accuracy.

Whether children looked longest at the target or a foil during referent selection did not predict 24-hour retention accuracy.

Table 17

Summary of the fixed effects in the final generalised linear mixed-effects model (log odds) of 24-hour retention accuracy, predicted by the longest look to the target/foil stimuli during 24-hour retention

Fixed effects	**Estimated coefficient**	**Std. error**	***z***	**Pr(> |z|)**
(Intercept)	-0.16	0.15	-1.11	.27
Longest look target	3.24	0.25	12.72	<.001
	AIC	**BIC**	**logLik**	**deviance**
	474.8	500.2	-231.4	462.8

4.5.4. Discussion: Looking during word learning and relationships with accuracy

We investigated how visual attention of autistic and neurotypical children differs across word learning stages and explored how looking measures predicted behavioural response accuracy. During referent selection, autistic children looked more frequently towards target stimuli, but responded less accurately than neurotypical peers. Both groups made more looks towards stimuli in the liked condition compared to the disliked condition during referent selection. During 5-minute retention, both populations allocated their visual attention similarly across conditions. However, after a 24-hour delay, autistic children made more looks towards the target than neurotypical children. Children’s in-trial visual attention predicted accuracy at all three stages across groups and conditions. Children’s looking during referent selection predicted their 5 minute and 24-hour retention accuracy. In terms of explicit behavioural responses, five-minute retention accuracy also predicted 24-hour retention accuracy, so visual attention at encoding may have significant downstream consequences for both short- and long-term consolidation of new vocabulary.

Children allocated proportionally more of their looking time towards targets on familiar trials, compared to novel trials, at referent selection. This perhaps reflects children’s certainty during the identification of familiar referents in comparison to more unusual, novel objects. In the case of identifying familiar referents, children could more easily identify and fixate upon these stimuli as they were stimuli for which they had pre-existing labels. However, in novel trials, children had to attend to the array of stimuli to rule out the familiar objects and select the correct novel object as the intended referent (de Marchena et al., 2011; Rothwell et al., accepted, see Study 2; Zosh et al., 2013). Given that identifying an unfamiliar referent requires greater attentional distribution across the visual scene, as well as heightened cognitive processing whilst using ME, it is unsurprising that both groups looked more divisively at stimuli on novel trials.

At referent selection, autistic children made more looks towards the target stimuli than neurotypical children, and both populations made more looks in novel trials (vs. familiar trials) and in the liked condition (vs. disliked condition). Increased frequency of looking also predicted greater accuracy in both groups, although this effect was stronger for neurotypical children. During referent selection, children needed to check back and forth to correctly identify referents, engaging in ME and ruling out competitor stimuli before selecting the correct referent (Halberda, 2003; Hartley et al., 2020; Lewis et al., 2020). Given the delayed linguistic profile of our autistic sample coupled with their lower referent selection accuracy (for both familiar and novel trials), it is possible that they required more looks towards target stimuli to decipher the correct referent due to uncertainties associated with comprehension. Since both groups made more looks towards the target in the liked condition, object preference could have resulted from perceptual features or salient properties of stimuli, thus making children more likely to look at their liked objects. This aligns with previous evidence that both autistic and neurotypical children attend to stimuli relating to their interests and preferences more intensively (Ackermann et al., 2020; Sasson et al., 2008, 2011). However, it is important to note that the conditional effects present in children's looking behaviour did not influence their response accuracy at referent selection.

Our analyses revealed that, across groups and conditions, 5-minute retention accuracy was predicted by number of looks towards the target during referent selection. Both 5 minute and 24-hour retention accuracy were predicted by proportion of time spent looking at the target during referent selection. These data corroborate the findings of our behavioural analyses that greater referent selection accuracy afforded superior 5 minute and 24-hour retention accuracy, and 5-minute retention accuracy predicted increased 24-hour retention accuracy. Together, these findings substantiate the claims of prior research that children's encoding during novel word disambiguation predicts the likelihood of longer-term learning (Bion et al., 2013; McMurray et al., 2012; Rothwell et al., accepted, see Study 2). They indicate the strong relationship between these stages and measures and suggest that visual attention at encoding may have significant downstream consequences for long-term consolidation of new vocabulary. These combinations of results highlight that enhancing and reinforcing children's identification of meaning may result in greater likelihood of novel word retention (see Rothwell et al., Study 2 for similar results).

All three of our looking measures significantly predicted response accuracy at referent selection, 5-minute retention, and 24-hour retention. Indeed, this finding shows that visual attention is an important predictor of learning accuracy. Moreover, across all three experimental phases, both longest look and proportional looking did not significantly differ across groups or conditions. Research utilising looking measures such as proportional looking to indicate accuracy also suggest that neurotypical and autistic children perform comparably well on word learning tasks (e.g. Akechi et al., 2013; Rothwell et al., Study 2; Swensen et al., 2007; Venker, 2019). Although we did not find differences in measures of proportional looking, group differences in behavioural response accuracy were identified at referent selection. Therefore, these findings highlight the benefit of using multiple methods to investigate word learning in autistic and neurotypical children. Whilst looking data reveals how children allocate their attention across the visual scene, response accuracy demonstrates children's explicit choices following attention to audio-visual stimuli.

4.6. General Discussion

Here, we investigated whether autistic and neurotypical children matched on receptive language abilities differ in accuracy and/or visual attention when learning words associated with liked and disliked stimuli. Advancing beyond previous research, the study examined how explicit measures of word learning accuracy and visual attention inter-relate. Our data reveal that autistic children were *less* accurate to spontaneously identify the meanings of novel words using ME than neurotypical peers. However, both groups retained word-referent mappings after 5 minutes and 24 hours with comparable accuracy. The extent to which children liked, or disliked, object stimuli did not influence the word learning accuracy of either group. Both autistic and neurotypical children did not significantly differ in the time they took to generate correct responses across all three experimental stages. Analyses of in-trial visual attention revealed conditional effects on looking behaviour at referent selection for both populations, and 5-minute retention for autistic children. Despite this, no between-condition differences in accuracy were detected at either retention test. We also discovered that children's 5 minute and 24-hour retention accuracy was predicted by visual attention at referent selection, indicating a relationship between fast mapping and retention mechanisms. Population differences in visual attention were observed at 24-hour retention, with autistic children making more looks towards targets, although no differences in response accuracy were identified.

In terms of the relationship between children's idiosyncratic preferences and word learning, our studies highlight two potential outcomes. One possibility is that stimulus preferences do not influence word learning in autistic or neurotypical populations. Whilst this may be the case, other studies (e.g. Ackermann et al., 2020; Rothwell et al., Study 1, Study 2, accepted) have demonstrated that interests in specific categories of stimuli affect word learning in neurotypical and autistic children. Reflecting upon this, we suggest it is more likely that children's self-selected liked and disliked stimuli did not elicit strong enough preferences in the present study to afford conditional differences in word learning. Indeed, stimuli in the present study comprised random, generic experimental stimuli. Although children selected four stimuli as being their most liked stimuli out of the set of twelve, this does not necessarily mean that they had a particular interest in these stimuli (they may have just been their favourite of the options given). Thus, a second – and perhaps more likely – possibility is that it is specifically long-standing preferences, or stimuli that relate to particular categories of interest, that impact word learning to a greater extent in these populations.

The present study acknowledges that word learning is an integrated, multi-stage process by investigating referent selection *and* retention in autistic and neurotypical children (Hartley et al., 2019, 2020; Horst & Samuelson, 2008). Although our language delayed autistic participants responded with lower accuracy at referent selection compared to language matched neurotypical peers, both groups retained novel words with comparable accuracy after 5 minutes and 24 hours (for similar results, see: Carter & Hartley, 2021; Haebig et al., 2017; Hartley et al., 2019, 2020; Rothwell et al., accepted, see Study 2). The discrepancy between referent selection and retention accuracy adds to a growing body of evidence showing that accurate referent selection does not necessitate accurate retention (e.g. Gliga et al., 2022; Hartley et al., 2019; Horst & Samuelson, 2008). Whilst our autistic population appeared to be less accurate to identify the meanings of words, their comparable retention accuracy suggests that they still encoded and retained novel words just as well as their neurotypical counterparts. As such, these findings highlight the requirement to study word learning as integrated multi-processes that take place in both the short and long term and raise caution to drawing conclusions about ‘word learning’ from referent selection data alone (Hartley et al., 2019; McMurray et al., 2012). Overall, our data reveal that autism does not impair fundamental mechanisms that underpin word learning. When expectations are based on receptive language ability, autistic children can retain novel words as effectively as neurotypical peers. Consequently, delays in autistic children’s naturalistic language learning and vocabulary development must have an alternative cause.

This study’s novel combination of behavioural accuracy and looking measures demonstrates the importance of integrating multiple methodologies to gain a comprehensive insight into children’s word learning. Our looking data revealed a conditional effect that was not present in children’s behavioural response data. Both autistic and neurotypical children looked more frequently at stimuli in the liked condition compared to the disliked condition at referent selection, and this finding was also apparent for autistic children at 5-minute retention. However, behavioural responses suggested no conditional differences in word identification across the three experimental stages. Unlike previous studies (Ackermann et al., 2020; Rothwell et al., Study 1, Study 2, accepted), the present study only included one stimulus category – novel objects – rather than contrasting stimulus categories with varying preference levels (e.g. animals vs. objects). As such, whilst children may look more frequently at items they favour, this does not appear to be strong enough to influence their word learning when liked and disliked referents belong to the same category. Indeed, this may be because the visual distinction between the two conditions was not especially salient – all stimuli varied in texture and were brightly coloured. Alternatively, idiosyncratic preferences to generic experimental stimuli may not be as strong as long-standing categorical or preferential interests, eliciting fewer differences in attentional mechanisms and, consequently, learning.

Comparing looking and behavioural measures revealed that these outcomes can both compliment and contradict one another. Although increased looking predicted accuracy at each word learning stage, population differences in looking allocation did not always result in population differences in learning accuracy (or vice versa). Our analyses for number of looks towards the target object revealed that autistic children made more frequent looks across conditions at both referent selection and 24-hour retention, and in the liked condition at 5-minute retention. However, these more frequent looks for autistic children did not yield equivalent conditional differences in behavioural responses at 5-minute retention, or population differences in response accuracy at 24-hour retention. As such, differences in visual attention may not always directly translate to greater or poorer performance on explicit behavioural responses. It is therefore possible that looking time paradigms, or specific looking measures, may only represent a partial picture, and multiple methodologies may yield a more comprehensive insight into children's word learning.

Our finding that superior retention accuracy was predicted by greater visual allocation towards targets compared to distractors at referent selection stipulates a direct relationship between fast mapping and slow learning mechanisms. Whilst previous research with neurotypical children proposes that these word learning mechanisms are distinct (e.g. Horst & Samuelson, 2008), our findings indicate a link between visual attention during referent selection and retention (for similar results see also Rothwell et al., Study 2). Specifically, our data suggest that encoding during referent selection is important for 5-minute retention and 24-hour retention. These findings highlight the importance of optimising input at referent selection to support more accurate retention. Even though our autistic children made less accurate responses during referent selection than neurotypical peers, clearly these responses were not detrimental to their overall retention. Considering this, it is possible that autistic children in fact retained a higher proportion of correctly mapped words than neurotypical children. Overall, we demonstrate here that autism does not detrimentally impair word learning mechanisms (e.g. Hartley et al., 2020; McGregor et al., 2013) or exploration of visual stimuli (Foti et al., 2015; Roser et al., 2015). As such, further work should advance our understanding of how increasing visual attention during encoding can lead to more robust novel word representations, and how the relationship between learning mechanisms and visual attention differs between short- and long-term retention.

The array of findings demonstrated in the present study can inform interventions that support autistic children's word learning. Our study adds to a growing body of literature that touch-screen technology is a motivating and effective tool to enhance the word learning of autistic children (see also Hartley et al., 2020, Rothwell et al., accepted, see Study 2). Moreover, presenting a limited number of referents for children to actively decipher correct novel-word referents may be an effective way to facilitate autistic children's word learning. Consequently, these principles could be utilised in naturalistic learning contexts, as well as clinical and educational interventions, to help scaffold vocabulary acquisition in autistic children with delayed language development. Specifically, given that looking behaviour at referent selection predicted retention accuracy, optimising the environment to promote curiosity and exploration during identification of meaning may be particularly important for novel word learning.

Naturally, we must address the limitations of this study. We acknowledge that, unfortunately, our findings were obtained from a single study with modest sample sizes. Our participant recruitment was hindered by school closures and lockdown restrictions due to the COVID-19 pandemic which occurred whilst the study was underway. Therefore, future research should seek to replicate our findings with larger samples. Given that our study compared word learning abilities across populations when delays in language development were controlled for, our groups were matched on receptive vocabulary, but not chronological age (the autistic sample was significantly older than the neurotypical sample). Prior research suggests that matching on chronological age would likely have resulted in our autistic participants achieving significantly lower accuracy on our word learning tasks than neurotypical peers (Charman et al., 2003; Mathée-Scott et al., 2021; Tek et al., 2008). Moreover, as our study did not emulate a naturalistic environment, we cannot directly state whether our samples' performance levels would be comparable in naturalistic word learning contexts. Thus, future research should explore autistic children's word learning in a more demanding, naturalistic environment. For example, presenting children with more salient distractors, more complex and numerous stimuli, and faster paced learning environments with restricted processing times, would allow us to examine how performance is influenced by more naturalistic challenges.

An additional consideration surrounds the effectiveness of the conditional manipulation in the present study. We only identified significant conditional effects on number of looks and response times at referent selection (both populations), and number of looks at 5-minute retention (autistic group only). We did not observe conditional effects in behavioural accuracy, or any of the other looking or response time measures across the experimental stages. Although the identified conditional effects provide some indication for children having preferences for stimuli they identified as being 'liked', we are hesitant to confidently conclude that our stimuli sets differed sufficiently in terms of interest levels to impact visual attention and subsequent word learning. Indeed, whilst children could have liked their chosen stimuli more than the other stimuli available, we cannot necessarily say that they liked them more than the familiar stimuli in the array. As such, future research could employ other pre-study measures, such as questionnaires (e.g. Rothwell et al., Study 1, Study 2, accepted) or visual attention tasks, to decipher stimuli that are more likely to elicit stronger attentional biases.

In summary, this study provides valuable insight into how autistic and neurotypical children identify *and* retain meanings of novel words associated with self-identified liked and disliked stimuli. Crucially, children's idiosyncratic preferences for novel objects commonly employed in experimental tasks do not substantially influence attention to the extent that their word learning accuracy is significantly impacted. Despite our autistic participants' delayed language development and poorer referent selection accuracy, they responded at least as accurately as vocabulary-matched neurotypical children on retention measures following a short delay of 5 minutes, and a delay of 24 hours following overnight sleep. Consequently, we conclude that autism does not disrupt the fundamental word learning mechanisms. Whilst neither group significantly differed in their response times across the three experimental stages, they differed in their visual allocation towards stimuli at referent selection and 24-hour retention. However, these differences in visual exploration of stimuli did not detrimentally impact autistic children's word learning accuracy beyond referent selection. Our findings also indicate the importance of incorporating mixed methodologies to investigate word learning and demonstrate that examining looking time alone could lead to erroneous or incomplete conclusions.

Chapter 5: General Discussion

5.1. Overview of the Thesis

This thesis has explored how attentional biases influence autistic and neurotypical children's identification and retention of novel word meanings. Study 1 utilised accuracy and response time data to examine how the presence of categorically interesting animal stimuli influences word learning when animals are both targets and distractors, compared to object stimuli. Study 2 employed a novel combination of visual and behavioural data to explore how word learning is influenced by novel animal targets compared to novel object targets, when less interesting object distractors are present. Study 3 combined visual and behavioural responses, investigating how children's self-identified liked and disliked novel objects influenced word learning when presented alongside categorically similar object stimuli. Overall, these studies investigate word learning as a multi-stage system including referent selection via ME and retention after delays of 5 minutes and 24 hours. This approach allows us to explore how word learning mechanisms are both similar and different across populations, and how idiosyncratic and categorical preferences influence children's attention and subsequent word acquisition across the learning stages. Ultimately, the information obtained here can help to inform optimal learning conditions for autistic and neurotypical children, which can be utilised to scaffold interventions for children with language delays. Considerations for effective experimental design to implement in future research are also highlighted.

This general discussion will first recap the main findings and conclusions from each of the papers presented in this thesis. Following this, we discuss how these papers inform the main topics of this thesis: how category interests and idiosyncratic preferences influence word learning accuracy in autism and neurotypical development, processing speed during word learning, and attentional allocation. Next, a reflection of some of the limitations of this thesis will be provided, alongside suggestions for future research, before final conclusions are considered.

All studies in this thesis examined novel word learning across three distinct stages – referent selection, 5-minute retention, and 24-hour retention. These studies compare autistic and neurotypical children matched on receptive vocabulary abilities to account for language delays present in the autistic population. Study 1 investigated how children's animal interests affect word learning accuracy and response speed when learning novel words associated with animals (high interest stimuli) and objects (neutral interest stimuli). For this study, children learnt novel words via ME-based referent selection in the presence of high interest animal distractors. At referent selection, autistic children responded at similar speed for both novel and familiar trials, whereas neurotypical children were significantly faster to respond correctly during familiar trials compared to novel trials. This suggests that autistic children may have spent longer studying animal distractor stimuli across trial types, whereas neurotypical children spent longer disambiguating novel stimuli. Both populations recalled novel words after 5 minutes and 24 hours with similar speeds, revealing that response times at retention were unaffected by category interests. Both groups identified stimuli at referent selection, and recalled them after five minutes, with comparable accuracy. After 24-hours, in comparison to neurotypical children, autistic children demonstrated greater novel word retention for object stimuli than neurotypical children. However, for autistic children, greater interest in animals predicted *poorer* retention after a 5-minute delay, but *superior* retention accuracy after 24-hours. These findings highlight that for the autistic group, heightened attention to non-target competitors during mutual exclusivity, alongside perceptual contrasts between target and distractor stimuli, ultimately led to more robust encoding of novel word-referent mappings.

Study 2 utilised a novel combination of accuracy and looking measures, alongside response time, to investigate whether autistic and neurotypical children matched on receptive vocabulary differ in accuracy and visual attention when learning words associated with animals (high interest stimuli) and objects (neutral interest stimuli). In contrast to Study 1, here children learnt novel words via ME in the presence of familiar object distractor stimuli. Overall, both groups identified meanings of novel words associated with unfamiliar animals and objects with comparable accuracy. After 5 minutes, autistic children retained novel animal names with greater accuracy than neurotypical children. Autistic children also showed a greater increase in their accuracy between 5 minute and 24-hour retention tests and outperformed neurotypical children after a night's sleep. However, autistic children demonstrated slower response times than neurotypical children at each word learning stage, particularly for animal stimuli. Given the greater retention accuracy of autistic children, we anticipate that differences in response time may reflect a speed-accuracy trade-off. At referent selection, autistic children took significantly longer to produce comparably accurate results to neurotypical peers and spent longer examining novel animal stimuli. As such, they may have had more opportunities to encode audio-visual aspects of stimuli, leading to their more accurate retention.

In Study 2, looking time data revealed that autistic children also looked at targets significantly more often than neurotypical children at each word learning stage. Crucially, 24-hour retention accuracy of both populations was predicted by visual attention and auditory input at referent selection, indicating a relationship between fast and slow word learning mechanisms. Overall, these findings demonstrate that differences in visual attention do not necessarily lead to diminished word learning accuracy. Moreover, it could be that as autistic children made more looks towards the target objects, this afforded greater encoding and thus resulted in more accurate novel word retrieval following sleep consolidation. Despite this, the observed differences in visual attention and response times signal potentially reduced pace of processing audio-visual input in autism. Due to the input in naturalistic word learning contexts being quicker and less constrained than under experimental conditions, it is possible that this slower intake of input could impair naturalistic language acquisition.

Study 3 investigated whether autistic and neurotypical children matched on receptive language abilities differ in accuracy, visual attention, and response times during referent selection and novel word retention. Children initially completed an object preference task where they identified novel objects that they liked and disliked. They then learnt novel words for their unique sets of liked and disliked stimuli in the presence of familiar objects, and retention was tested after 5 minutes and 24 hours. Here, autistic children were *less* accurate when spontaneously disambiguating the meanings of novel words using ME than neurotypical peers. However, the groups retained word-referent mappings after 5 minutes and 24 hours with comparable accuracy. Whether stimuli were liked or disliked did not significantly influence the word learning accuracy of either group, and both groups took similar amounts of time to generate correct responses across all three word learning stages.

Analyses of in-trial visual attention in Study 3 revealed that looking behaviour was both complimentary and contradictory of explicit behavioural accuracy measures. During referent selection, population differences were present in both visual attention and explicit behavioural accuracy measures. Conditional differences in looking behaviour were observed during referent selection and 5-minute retention, but these did not influence conditional differences in children’s behavioural accuracy on any of the word learning stages. On the contrary, population differences in looking patterns were discovered at both 5 minute and 24-hour retention, although no differences in behavioural response accuracy were identified. These findings suggest that differences in visual attention in autistic children do not necessarily lead to diminished word learning accuracy. Corroborating the findings of Study 2, children's retention accuracy was predicted by visual attention at referent selection, indicating a relationship between fast mapping and retention mechanisms. Indeed, the relationship between identification of meaning and longer-term learning across studies highlight that if we can increase autistic children intake from input during the initial novel word identification, word learning outcomes may be optimised.

5.2. How do Children's Interests Influence their Word Learning?

Strikingly, Study 1 highlighted the importance of category interests for autistic children's word learning. Autistic children with greater animal interests had more fragile novel word representations after a 5-minute delay, but more accurate word-referent mappings after a 24-hour delay. Autistic children were also significantly more accurate than neurotypical children to recall novel object names after 24-hours. These findings suggest that the presence of animal distractor stimuli may have allowed children to encode more information from the visual scene due to their heightened interests. Additionally, the greater perceptual contrasts between categorically distinct novel objects and familiar animals may have facilitated encoding of more robust novel word-object representations for autistic children, that were consolidated by sleep and more easily retrieved after 24 hours. Curiously, we found no effect of animal interests for the neurotypical group, despite there being no significant differences between the groups on our measure of animal interests.

These findings are supported by previous research reporting that increased attentional allocation aids learning in both neurotypical and autistic children (e.g. Akechi et al., 2013; Axelsson et al., 2012; Begus & Southgate, 2012; Hilton et al., 2019; Parish-Morris et al., 2007). It is possible that our findings were only apparent in the autistic group due to their particularly circumscribed interests, making autistic children more likely to display perseverative attentional allocation towards stimuli that appeal to their interests or that are perceptually salient (Richler et al., 2007; Sasson et al., 2008, 2011; Venker et al., 2021, 2022).

Consequently, autistic children may have more intensively attended to the visual scene during encoding, attending to categorical distinctions which lead to more successful retention. Moreover, individual animal interests did not significantly influence the 24-hour retention of either neurotypical or autistic children in Study 2, where only novel stimuli were animals (familiar stimuli were objects). Overall, it appears that the animal distractors in Study 1 benefited children's learning of non-interest stimuli belonging to a different category, whereas learning new same-category names in Study 2 may not have received the same benefit. However, the findings of Studies 1 and 2 are somewhat converse in relation to long-term consolidation, a point that will be addressed in section 5.4 where we discuss how sleep influenced learning and mediated interest effects.

In Study 3, selecting liked and disliked stimuli from a range of novel objects may not have had a strong enough influence on attention to yield conditional effects on learning for either population. Indeed, this contrasts with the findings of Studies 1 and 2, whereby the presence of animal stimuli in those studies influenced the retention accuracy of autistic children. We therefore suggest that it is pre-existing category interests, rather than idiosyncratic preferences to specific novel object stimuli, which has the stronger and more reliable influence on autistic children's word learning. It is likely that autistic children's predisposition to more intensive and inflexible attentional mechanisms towards preferred stimuli was a benefit in terms of animal interests, and how much information they encoded about the novel stimuli (Elsabbagh et al., 2009, 2013; Honey et al., 2012; Venker et al., 2018).

In view of these findings, it may be that autistic children were exhibiting a 'Goldilocks effect', whereby they allocate auditory and visual attention to stimuli that are neither too simple or familiar, nor too complex or unfamiliar, to avoid wasting cognitive resources (Kidd et al., 2012, 2014). In terms of category knowledge, when children know very few or very many existing category members, their interest in the category may fluctuate. Either they may have minimal exemplars to integrate with the new category members, or they alternatively consider the category to be saturated and do not feel the need to acquire any more category members (Ackermann et al., 2020). When autistic children had particularly strong representations for interesting animals in Study 1, they could disengage from familiar animal stimuli as they belonged to a well-established category so did not need to encode further information. This afforded particularly strong long-term retention in the object condition for autistic children because it resulted in greater focus on object stimuli due to the differences in curiosity at encoding. Following this, autistic children were more able to integrate novel labels into the less established object category. By contrast, in Study 2, familiar object distractors did not appeal to children's categorical interests. Therefore, novel animal stimuli initially had more of an influential role on autistic children's learning due to greater interest in the category at encoding. As such, in Study 2, novel animal names were more easily integrated into pre-existing lexical categories for 5-minute retention, until sleep consolidated information within both categories. In Study 3, all stimuli were categorically less interesting than the animal stimuli present in Studies 1 and 2 (only objects were presented). Such stimuli that did not particularly appeal to children's innate interests thus did not afford as effective integration into the lexicon. Consequently, autistic children did not demonstrate superior retention when all target stimuli belonged to the same category and distractors did not appeal to pre-existing interests.

5.3. Word Learning Mechanisms

All studies in this thesis communicate a key message – that fast, in-the-moment referent identification and slower associative learning processes are inter-related. Previous research has suggested that word learning in children is underpinned by separable referent selection and retention mechanisms (e.g. Kucker et al., 2015; McMurray et al., 2012). However, the present studies suggest that this is not always the case for either neurotypical or autistic children, with predictive relationships between referent selection, 5-minute retention and 24-hour retention accuracy. This aligns with research positing that, for neurotypical children, referent selection accuracy predicts retention (but not necessarily for late talkers; Kucker & Seidler, 2023). Similar findings have also been demonstrated with children at increased risk for autism (Gliga et al., 2022). Referent selection in the present study required children to actively select stimuli through a touch response. It is possible that this active choice was particularly beneficial for children due to heightened engagement with the task (Blewitt & Langan, 2016; So Yeon et al., 2018), in turn reinforcing the mapping for increased retention. Interestingly, in Study 1, autistic children were found to receive a greater benefit of referent selection accuracy on 5-minute retention than neurotypical children. Given that autistic children's attentional mechanisms are often less flexible than those of neurotypical peers (Hutman et al., 2012; Landry & Bryson, 2004), they may gain more from active learning than neurotypical children in some circumstances – for example, when distractor stimuli appeal to a category of interest (Bosseler & Massaro, 2003; Fantasia et al., 2020; Zwaigenbaum et al., 2015). These findings reveal the importance of examining word learning as a multi-stage system – word learning mechanisms appear to be unimpaired in autism at both short- and long-term retention.

Critically, in line with an array of previous research (de Marchena et al., 2011; Hartley et al., 2019, 2020; Parish-Morris et al., 2007; Preissler & Carey, 2005), all studies in this thesis demonstrate that autistic children can engage in ME to accurately decipher novel word referents. It is important to acknowledge, however, that whilst Studies 1 and 2 suggested both populations could fast map with comparable accuracy, in Study 3, autistic children were less accurate than neurotypical children at referent selection. Despite this, autistic children still used ME to identify novel referents significantly above chance ($M_{accuracy} = 0.72$, $M_{chance} =$

0.33), and retained novel words with comparable accuracy to neurotypical peers. As such, these findings add to a growing body of research demonstrating that fundamental word learning mechanisms are not qualitatively atypical in autism.

Consequently, our studies demonstrate that autistic children can retain novel labels at least as accurately as neurotypical peers when expectations are based on receptive vocabulary. Our findings suggest that if the mechanisms underpinning word learning are not impaired in autism, then there must be alternative causes for the language delay experienced by many autistic children. To correctly identify a new word-referent mapping, children must focus on auditory and visual input, which is narrowed by intrinsic or extrinsic sources of information (Hollich et al., 2000). Studies often attribute children's word learning problems to difficulties understanding and applying social pragmatic cues in learning contexts (e.g. Baron-Cohen et al., 1997; McDuffie et al., 2006; Parish-Morris et al., 2007). However, in addition to social cues, word learning relies on input from other sources. Autistic children's attentional mechanisms often differ to those of neurotypical children, with autistic children displaying different preferences to stimuli (Pruden et al., 2006, Parish-Morris et al., 2007), as well as atypical exploration of stimuli (Sasson et al., 2011; Venker et al., 2022). In view of our findings in Studies 1 and 2 that category interests were beneficial in the service of word learning for autistic children, we anticipate that differences in attentional mechanisms that may be responsible for language delays. This point will be revisited in section 5.6 within a discussion of what this means for learning more broadly.

In support of this theory, Studies 2 and 3 extended our approach to examining word learning by introducing mixed methodologies, combining both visual attention and explicit behavioural accuracy measures. This novel combination afforded greater insight into how the three word learning mechanisms inter-relate in terms of visual attention during learning. Investigating children's looking behaviour enabled us to identify whether attention differed across groups and determine how this influenced word learning. Both studies found that visual engagement with stimuli during referent selection predicted retention. This aligns with research by Hilton and colleagues (Hilton & Westermann, 2017; Hilton et al., 2019) who found that shy children who were less likely to attend to a novel target during referent selection retained fewer novel labels than non-shy children. Similarly, research by Bion et al. (2013) demonstrated a positive relationship between attention during referent selection and retention in a looking-based task with neurotypical toddlers. Overall, these findings add to a growing body of literature demonstrating that focusing more on a target during encoding leads to greater retention success. Paying attention to the right thing at the right time ultimately leads to more successful novel word learning. However, population differences in attention allocation do not necessarily result in diminished word learning accuracy.

5.4. Sleep and Word Learning

Our findings highlight the importance of sleep for novel word learning. Sleep consolidates novel declarative memories, such as novel words, by reactivating recently encoded novel word-referent relationships (Diekelmann & Born, 2010; Robertson, 2009). Prior research has suggested that sleep consolidates word learning in both neurotypical children (Brown et al., 2012; Henderson et al., 2012; Ma et al., 2022) and autistic children without language delay (Fletcher et al., 2020; Henderson et al., 2014). However, to our knowledge, no studies prior to this thesis have investigated the effect of sleep on language delayed autistic children. We expected sleep to potentially be less beneficial for our sample of autistic children due to the presence of language deficits and likely differences in sleep quality (Devnani & Hedge, 2015). However, all three studies demonstrate that sleep consolidation was at least as beneficial for language delayed autistic children as neurotypical children matched on receptive vocabulary.

In Study 1, autistic children's retention accuracy *decreased* by 7% in the animal condition after sleep ($M_{5\text{-minute}} = 0.46$ $M_{24\text{-hour}} = 0.43$) but *increased* by 17% in the object condition ($M_{5\text{-minute}} = 0.53$, $M_{24\text{-hour}} = 0.62$). Autistic children retained a similar proportion of novel objects and novel animals after 5 minutes, but more novel objects than animals after a 24-hour delay. By contrast, following sleep, in Study 2 autistic children's retention accuracy increased by 13% in the animal condition ($M_{5\text{-minute}} = 0.55$ $M_{24\text{-hour}} = 0.62$) and 54% in the object condition ($M_{5\text{-minute}} = 0.37$, $M_{24\text{-hour}} = 0.57$). Autistic children retained more novel animal names than novel object names after a 5-minute delay but performed similarly across conditions after a 24-hour delay. The effects apparent in Study 1 are likely due to the attentional salience of target stimuli, particularly given that object label retention was more accurate in Study 1 than Study 2. In Study 2, target animals were interesting and perceptually salient, as they differed categorically to the other stimuli in the array. In Study 1, target animals were interesting, but not perceptually salient - target objects were more attentionally salient here as they comprised only a third of the array at referent selection. As such, the greater perceptual distinctions may have afforded more robust encoding of novel word-referent mappings for novel objects in Study 1, which were consolidated by a period of sleep leading to increased performance after 24 hours.

In Study 2, autistic children were more accurate to identify novel animal stimuli than novel object stimuli after a 5-minute delay. However, following a period of overnight sleep, this conditional effect diminished, with autistic children outperforming neurotypical children in both conditions after 24 hours. Whilst autistic children's representations of novel object words were initially more fragile at 5-minute retention in Study 2, sleep consolidated these more vulnerable representations into memory networks, improving recall of novel words after a 24-hour delay (Gais & Born, 2004). These findings align with previous research that the benefits of sleep consolidation are greater for weak memory traces than for strong memory traces (e.g. Diekelmann et al., 2009; Gais & Born, 2004).

Similarly, sleep protected neurotypical children's novel word consolidation from decay. In Studies 1 and 2, neurotypical children's performance was similar at 5-minute retention and 24-hour retention across conditions. Although autistic children outperformed neurotypical children in the object condition in Study 1, and across conditions in Study 2, neurotypical children's retention performance did not worsen in either condition or study. This suggests that sleep provided the anticipated protective effect on new declarative memories (Axelsson et al., 2018; Horváth et al., 2016; Williams & Horst, 2014), but potentially to a lesser extent than for autistic children. It is possible that the benefits of sleep for neurotypical children were smaller due to their significantly lower chronological ages than the autistic group. Whereas our sample of autistic children were on average over 7 and a half years old, the neurotypical group averaged a little over 4 years old. We originally noted in Study 2 that chronologically older children have longer sleep cycles, which can be directly related to memory consolidation cycles (Hill et al., 2007; Horváth & Plunkett, 2016; Montgomery-Downs et al., 2006). It is therefore possible that the younger neurotypical children did not receive as great of a benefit from sleep due to more immature sleep cycles and memory mechanisms (a point which will be revisited in section 5.7).

However, there is a range of prior literature demonstrating that overnight sleep benefits neurotypical children's retention at younger ages. For example, Axelsson et al. (2018) found that 2-year-olds' novel word retention benefits from napping after 4 hours and sleeping overnight. Similarly, Williams and Horst (2014) taught 3-year-olds novel words via storybook reading and found that sleep was beneficial for retention following delays of 2.5 hours, 24 hours, and 7 days. However, in our Studies 1 and 2 (Rothwell et al.), sleep did not particularly improve the novel word retention of neurotypical children. Moreover, in Study 3

(Rothwell et al.) we did not find that sleep consolidated novel words for either autistic or neurotypical children. In Studies 1 and 2, the benefits of sleep for autistic children were perhaps aided by the presence of categorically interesting animal stimuli at encoding, meaning that the less categorically interesting object stimuli further benefitted from sleep consolidation due to initially more fragile representations. However, the presence of solely object stimuli in Study 3 meant that autistic children did not benefit from categorical distinctiveness or pre-existing interests in animal stimuli (unlike Studies 1 and 2). Therefore, encoded word-referent representations may have been less robust in Study 3, perhaps meaning that the same advantage of sleep consolidation on novel word-referent mappings was not achieved. Given that the findings of our studies contrast with previous literature on the benefits of sleep for neurotypical children's novel word retention, this research area requires further investigation.

5.5. Processing Speed

Our third dependent measure, response time, speaks to how quickly children generated correct responses during word learning processes. In Study 1, autistic children responded at a similar pace for both novel and familiar referent selection trials, whereas neurotypical children were significantly slower to respond correctly during novel trials compared to familiar trials. Recall that familiar stimuli in Study 1 were animals, suggesting that autistic children spent similar time exploring stimuli in both trial types – perhaps due to the interesting animal distractors. By contrast, neurotypical children appeared to spend comparatively less time exploring stimuli in familiar trials relative to novel trials. Similarly, in Study 2, autistic children were particularly slow to identify novel animal stimuli. In both studies, the longer response times of autistic children when identifying animal stimuli may reflect their greater attentional allocation to stimuli that appeal to their interests. Previous research proposes that autistic children perseverate on stimuli that are particularly salient to them (Sasson et al., 2008, 2011), and can have difficulties disengaging from stimuli that appeal to their interests (Akechi et al., 201; Elsabbagh et al., 2009). Indeed, our findings suggest that categorically interesting animal stimuli attracted autistic children's focus more so than stimuli that belonged to categories that were not of particular interest.

Across all word learning stages in Study 3, there were no observed differences in response times between populations. However, we did find that both autistic and neurotypical children took longer to generate correct responses in the 'liked' stimuli condition at referent selection. Whilst this result indicates that children focused longer on stimuli that captured their interests, it contrasts with our previous studies in that no population difference was detected here. Unlike Studies 1 and 2, in Study 3 object stimuli served as both distractor and target referents (no animal stimuli were used in this study). Consequently, it appears that the presence of categorically similar stimuli across the array elicited more consistent response times across populations. This highlights the importance of category interests for autistic children – when stimuli did not belong to categories of interest, autistic children's response speeds were comparable to those of neurotypical peers. Indeed, preferences for novel object stimuli were not strong enough to elicit differences in visual attentional allocation between populations.

Interestingly, our response time findings in Study 3 differed from those reported by Hartley et al. (2020), where similar samples of autistic children were slower to identify novel object stimuli than receptive vocabulary matched neurotypical peers. Akin to Rothwell et al. (Study 3), Hartley et al. (2020) presented only generic object stimuli at both referent selection and retention. However, in Hartley and colleagues' study, children learnt novel words through cross-situational learning, whilst our study utilised ME-based referent selection. Cross-situational learning requires children to recall prior exposures in order to disambiguate word meanings, whereas arrays that afford the use of ME present all the necessary information within a single naming event. As such, cross-situational methods are more cognitively demanding than ME, requiring greater use of working memory mechanisms to store information about previous exposures (McGregor et al., 2022; Neveu & Kaushanskaya, 2023; Vlach & Sandhofer, 2014). Given that autistic children demonstrate difficulties in executive function, including working memory (Joseph et al., 2005; Lopez et al., 2005; Robinson et al., 2009), this may account for the slower response times of autistic children in Hartley et al's (2020) study. This methodological difference between Rothwell et al. (Study 3) and Hartley et al. (2020) may explain why the same population differences in response time were not identified across the studies, despite including comparable stimuli sets and participants with similar characteristics. Investigating the role of interests on autistic and neurotypical children's word learning in a cross-situational learning paradigm would be an informative objective for future research to examine how interests influence referent selection mechanisms besides fast mapping.

Differences in visual attention may explain variation in response times across populations and conditions. In Study 2, population differences between multiple looking measures were found across word learning stages. Similarly, population differences in response time were discovered at referent selection and retention in Study 2. By contrast, population differences in only a single looking measure – number of looks - were identified in Study 3. Despite this, no significant population differences in response speed were identified at retention in Study 3. These combinations of results highlight that not all differences in looking behaviour directly relate to differences in processing times. It is possible that population differences in response time only appear when there are population differences in *multiple* looking measures (as identified in Study 2). Moreover, the relationship between conditional effects on looking measures and response times were also varied. Our looking measures in Study 2 identified a significant conditional effect at referent selection that was also present in our referent selection response time data. However, the response time data demonstrated a population x condition interaction that was not identified in our looking data. Although, examination of visual attention at 24 hours did not identify the same significant conditional effects present in our response time data at 24-hour retention in Study 2. By contrast, conditional differences in number of looks and response time were congruent at referent selection in Study 3. However, conditional differences in number of looks were identified for the autistic group at 5-minute retention in Study 3, but this did not translate to conditional differences in response times for autistic children. Thereby, we highlight the usefulness of gathering data from a variety of measures to build a comprehensive picture of children's processing speed. However, it is crucial to note that because our studies employed a unique combination of touchscreen and looking time methods, trial length was not controlled – children's progression through the studies was dependent on their response speeds. Investigating how children's looking allocation is affected when trial length is consistent would illuminate the extent to which visual attention is reflective of processing time.

5.6. Practical Applications

This thesis has many practical applications, both in terms of experimental design and the potential to inform understanding of word learning in naturalistic settings. In Studies 1 and 2, autistic children were more accurate to recall novel stimuli at retention compared to neurotypical children. Both studies also found that autistic children had slower response times at referent selection in comparison to neurotypical children. As such, these findings could be the result of autistic children demonstrating a speed-accuracy trade-off. Slower responses meant autistic children had longer to encode perceptual features and hear more repetitions of the novel label associated with target stimuli, which could have scaffolded their more accurate retention. If processing times were restricted, perhaps autistic children would have performed less accurately, or conversely, neurotypical children may have responded more accurately had they spent longer examining stimuli at referent selection.

Interestingly, Study 3 was the only study to identify no significant population differences in response speed during referent selection. It was also the only study in this thesis to identify significantly poorer referent selection in the autistic group compared to the neurotypical group. Had autistic children taken longer to generate correct responses, they may have achieved accuracy comparable to neurotypical peers. Given autistic children's comparable retention accuracy in Study 3, it is possible that they successfully retained a higher proportion of novel word-referent mappings successfully encoded at referent selection.

Taken together, the findings of our studies suggest that autistic children are slower to generate correct responses for stimuli that belong to a category of interest when learning occurs through ME-based referent selection. These studies also highlight the requirement of future research to investigate how accuracy and response speed interrelate. Insight into the relationship between these two variables would advance our understanding of how autistic children can learn in the natural environment. Indeed, slower processing of information in natural language environments may mean that autistic children have limited opportunities to acquire informants from their surroundings, potentially impairing their novel word acquisition (McMurray et al., 2012; Yu & Smith, 2012). These findings state that allowing autistic children longer processing times during word learning could be beneficial for language acquisition. Consequently, a consideration when designing learning environments and interventions would be to increase processing time to maximise learning output.

Our findings that language delayed autistic children were as accurate at learning novel words as vocabulary-matched neurotypical peers suggest that, under the right conditions, word learning can be optimised for children across an array of abilities. Our study set-up likely appealed to the strengths of autistic children - distractions were minimised and potential referents were limited from a choice of three. Research with neurotypical children found that children who learn novel words via ME in the presence of only two competitors successfully retained novel word-referent pairings, whereas those who learned alongside three or four competitors did not retain the novel labels (Horst et al., 2010). In the present studies there were only two competitor stimuli present during training, which supported successful retention in both neurotypical and autistic children. Whilst future research is necessary to investigate how number of competitors influences autistic children's word learning, here we highlight that limiting competition in naturalistic learning environments could increase their learning efficiency.

Children's performance in the present studies could partly be due to their mode of delivery. Studies investigating the effectiveness of touch-screen technology as a learning medium have yielded mixed results (e.g. Allen et al., 2015; Wainwright et al., 2020). Allen et al. (2015) state that it is not the method of delivery, but the content presented, that has the greater influence on children's symbolic understanding of word-picture-object relations. Wainwright and colleagues (2020) taught autistic and neurotypical children new information via either a paper book or e-book incorporating interactive, touch-screen features. Whilst they did not find any significant differences in learning performance, they discovered that children's visual engagement with stimuli was greater in the e-book condition compared to the paper book condition. In our studies, language delayed autistic children's novel word retention was at least as accurate as neurotypical children's performance. Consequently, we suspect that the use of touch-screen technology was clearly engaging in teaching novel words to autistic children. We therefore suggest that touch-screen technology may be considered for use within interventions to optimise autistic children's word learning. However, it is important to consider a balance between utilising touchscreens and more naturalistic contexts.

Potentially, touchscreens could be initially used to teach meanings of new vocabulary, and then learning can be transitioned to real-world situations to foster extension of knowledge across contexts. Overall, our studies highlight the importance of carefully considering how we craft learning environments to suit the learner, rather than expecting the learner to suit the environment.

Our studies consistently demonstrated that referent selection accuracy, alongside visual and auditory input at referent selection, predicted subsequent retention performance. As such, maximising intake of input at referent selection could lead to greater word learning success. Previous studies indicate that providing reinforcement at referent selection in the form of ostensive naming (Gliga et al., 2022; Horst & Samuelson, 2008), or drawing attention to target objects through illumination (Axelsson et al., 2012), improves children's retention. However, the present studies extend this through increasing attention via manipulation of stimulus categories. Additionally, increasing exposure through extended visual attention to stimuli, and increased number of label repetitions, improved novel word retention in our studies. This aligns with the dynamic associative theory of word learning, which proposes that accumulating information gradually (e.g. through repeated novel word-referent mappings across multiple situations) strengthens long-term learning (McMurray et al., 2012). Given that numerous studies with neurotypical children suggest that repetition increases word learning, this finding is not surprising (Horst & Samuelson, 2008; Vlach & Sandhofer, 2012). However, it does confirm that autistic children receive the same benefit of increased exposure as neurotypical children. It is possible that autistic children might themselves seek additional exposure by studying stimuli for longer and taking more time to respond. This suggests that increased exposure to visual and auditory input may help autistic children to acquire new language.

Our findings from Studies 1 and 2 indicate that autistic children's pre-existing interests in stimulus categories influence their word learning. Overall, heightened attention to the visual scene, coupled with greater perceptual distinctions between familiar animal distractors and novel object targets, afford more robust encoding that is further consolidated by sleep. Moreover, in Study 2, autistic children spent longer attending to novel animal stimuli that captured their interests, and initially retained more novel animal stimuli prior to sleep consolidation. Taken together, these findings suggest that autistic children attended more to information relating to their pre-existing interests, which in turn improved their novel word learning. This has implications for both research and intervention design. We recommend that researchers are mindful when combining stimuli from different categories at referent selection in experimental settings, as prior interests in distractors or targets may impact learning. This also highlights the requirement for interventions to incorporate children's interests to maximise learning. Certainly, we are not proposing that children can only learn information that appeals to their interests. Indeed, our findings in Study 3 that language delayed autistic children's retention accuracy was not poorer than neurotypical children's demonstrates that they can be just as capable of learning names for stimuli that do not align with their categorical interests. Rather, interesting stimuli could be incorporated within interactions involving non-interest stimuli to increase children's attention and engagement in the overall word learning environment and ultimately afford better long-term retention.

Strikingly, we found that population differences in looking behaviour did not lead to poorer word learning accuracy for autistic children. For example, in Study 2, proportion of time spent looking at target stimuli did not significantly differ across groups or conditions at 5-minute retention or 24-hour retention. However, autistic children responded more accurately in the animal condition after 5 minutes, and more accurately than neurotypical children in both conditions at 24-hour retention. Additionally, Studies 2 and 3 found that specific looking measures often predicted the accuracy of one population to a greater extent than the other. Different measures can provide different information about how children learn, and the success with which they acquire novel words. Our data suggest that looking measures illustrate how children engage with stimuli during learning, but not the outcome of learning in terms of accuracy. Conversely, forced choice measures tell us the outcome of learning, but not how they reached this response. Autistic children may engage with stimuli differently, often making more frequent and divisive looks across the visual scene than neurotypical children, but our studies reveal that this does not negatively impact their word learning accuracy. When utilising a single method, the conclusions drawn about children's learning may substantially differ depending on the methodology utilised. As such, it is imperative to select the most appropriate methodology for the task and mechanisms of interest – attention during learning or learning outcome. Undoubtedly, these findings have implications for the existing literature. It is possible that conclusions drawn in previous studies about word learning in both autistic and neurotypical populations may have differed had they included explicit measures of accuracy alongside looking measures (or vice versa). If the most suited methodology is not utilised, conclusions have the potential to be erroneous, or not representative of children's true abilities. We therefore recognise that utilising a combination of distinct methodologies is a strength and paves the way forward for future research.

5.7. Limitations and Future Directions

Of course, the studies that comprise this thesis are not without limitations. Whilst we have highlighted that our study design may appeal to the strengths of autistic children, we appreciate that it did not emulate a naturalistic context. Although this is true for many studies, it means that we cannot be certain whether our findings would replicate under more naturalistic conditions. Natural learning environments are rife with distractors, often presenting a plethora of external cues and potential referents at once. Given that stimuli within naturalistic contexts are often presented at a more rapid rate that is not dependent upon the child's processing speed, they may create a processing bottleneck for autistic children who struggle to attend to all necessary information at a sufficient pace. Consequently, future research should investigate how preferential biases influence word learning in autistic and neurotypical children within real-word learning environments. Though we suggest that learning environments should be designed to suit children's requirements, we appreciate that this is not always possible in day-to-day settings. As such, we must address how word learning mechanisms interplay when stimuli and environmental conditions are not optimised. Future research can therefore extend the present study design but increase complexity and demands. For example, asking children to choose from numerous stimuli, rather than a limited set of three potential referents, would advance understanding of optimal and sub-optimal competition levels.

It is important to note that data collection for this thesis began just before the COVID-19 outbreak occurred. The devastating effects of the global pandemic meant that, due to school and research facility closures, data collection ceased. As such, the sample sizes of our studies are more modest than anticipated. Of course, this means we must be cautious when generalising our findings to the wider population of autistic children. Future research should seek to replicate and extend these findings with larger sample sizes now that conditions permit. Moreover, disruption from the pandemic meant that some of the autistic participants did not have scores for all individual difference measures. Consequently, for autistic participants, we were unable to test the predictive effects of non-verbal intelligence, attention-related behaviours, and autism severity on word learning outcomes in Study 1, the effects of non-verbal intelligence in Study 2, or non-verbal intelligence and expressive vocabulary in Study 3. It is possible that further information can be obtained regarding how these participant characteristics influence children's word learning for different stimulus categories.

We also recognise that acquisition and retrieval of declarative memories may be subject to circadian effects, whereby time of day influences cognitive processes (Lau et al., 2010; Tilley & Warren 1983). Some research suggests that sleep-associated improvements in memory and integration can be explained by better performance in the morning than in the evening (Schmidt et al., 2007). As such, asking participants to learn and recall information at the same time of day would account for these effects (Gais et al., 2006). In our study, different participants were tested at different times of day (24-hours apart) due to practical considerations. Consequently, circadian confounds cannot be ruled out. However, as both populations were tested across a range of times during the day, circadian differences between the groups should be offset. Nonetheless, future research could ensure that all children learn and recall information at a specific time of day to rule out any potential effects of circadian rhythms. Such research has not yet been carried out in language delayed autistic children.

We also reflect upon the effect that sleep timing may have on word learning. Theories suggest that sleep is most effective for memory consolidation if it follows within a few hours of learning to reduce interference in memory traces (Diekelmann et al., 2009; Gais et al., 2006). Studies with neurotypical children also highlight the importance of sleep timing for word learning (e.g. Gómez et al., 2006, Williams & Horst, 2014). A study by Hupbach et al. (2009) found that napping within 4 hours of language exposure helped 15-month-olds to remember the general grammatical pattern 24 hours later. Importantly, learning was not improved if there was a long gap between learning and sleeping. Some studies demonstrate that a story read close to bedtime may elicit retention benefits for children with smaller vocabularies (e.g., James et al., 2020, Walker et al., 2020). However, Henderson et al. (2021) taught children aged 5 to 7 years novel words via hearing a story either at bedtime, or three to five hours before bedtime. By contrast, they discovered that overnight gains in comprehension were larger if the story was read three to five hours before bedtime, rather than at bedtime. It is therefore possible that the consolidatory effects of sleep timing are mediated by vocabulary knowledge (James et al., 2017). As such, future research should consider the influence of sleep timing on novel word learning for autistic children. However, autistic children's sleep characteristics often differ from those of neurotypical peers (Elrod & Hood, 2015), which may introduce practical considerations for future research. For example, sleep onset latency tends to be longer in some autistic individuals (Tse et al., 2020). Consequently, sleep may be induced later than neurotypical children with faster sleep onset latencies, suggesting that the timing of sleep in relation to learning input would be delayed for affected autistic children. Moreover, some studies reveal that autistic children have fewer and shorter naps than neurotypical and developmentally delayed peers (Goodlin-Jones et al., 2008; Schwichtenberg et al., 2011), so requiring autistic children to nap following learning may be more challenging for this population.

Whilst investigating the influence of sleep on novel word learning provides valuable insight into children's word learning mechanisms, it is imperative to consider that the sleep patterns of our samples may have differed. In Studies 1 and 2, autistic children outperformed neurotypical children after a night's sleep (in the object condition for Study 1, and across conditions for Study 2). We highlight that autistic children were chronologically older and thus may have had longer sleep cycles (Hill et al., 2007; Montgomery-Downs et al., 2006). Given that sleep cycle length is directly related to the formation of novel declarative memories such as novel words (Horváth & Plunkett, 2016; Kurdziel et al., 2013; Tamminen et al., 2010), it is possible that their older chronological ages may have influenced this finding. However, in Study 3, samples of children were not matched on chronological age, and we did not detect advantageous benefits of sleep for either neurotypical or autistic children's novel word learning. Therefore, it appears that category interests influence longer term novel word consolidation in autistic children alongside the facilitative effects of sleep. However, future research should seek to replicate this finding with chronologically aged-matched samples to ensure age differences in sleep cycles are accounted for.

Autistic children are particularly vulnerable to sleep disorders, which often include parasomnia, sleep anxiety, and bedtime resistance (Díaz-Román et al., 2018; Souders et al., 2009). Cognitive function and memory consolidation have been linked to sleep duration and quality, particularly in autistic children (Calhoun et al., 2020; Maski et al., 2015). Based on our findings in Studies 1 and 2, we know that sleep helps to consolidate novel word retention in autistic children, particularly when stimuli are categorically interesting. However, we did not gather information about the properties of children's sleep in these studies, so we do not know how particular aspects of sleep may have influenced word learning. As such, future research should involve objective sleep measures, such as polysomnography or actigraphy, to elucidate how differences in sleep quality influence vocabulary acquisition over both the short- and long-term.

Whilst testing retention after 24 hours is a crucial step to examining longer term retention, we acknowledge that memories may still decay beyond the 24-hour period. Previous studies have shown that neurotypical and autistic children can retain novel labels after delays of more than 2 weeks (Wainwright et al., 2020). In contrast, Norbury et al. (2020) suggested that although autistic children were more successful than neurotypical children at mapping phonological forms to novel referents, this advantage was not maintained following a delay of 4 weeks. Given the observed advantage of overnight sleep on novel word retention for autistic children in Studies 1 and 2, an important next step would be to analyse whether these population differences would endure over longer time periods. Exploring whether novel word representations decay when learning relates to specific categories of interest would allow us to investigate the fragility of newly encoded novel words, and further understand how we can optimise language acquisition.

Of course, animals represent a very limited subset of stimuli which children can find particularly interesting. Autistic children's special interests vary widely, with a recent survey by Nowell et al. (2021) indicating that television, objects, music, and toys are among the most frequently reported categories of interests. Within these categories, children may have very niche interests, for example focusing on specific characters in a television programme such as Thomas the Tank Engine or characters from the Marvel Universe (Uljarević et al., 2022). A study incorporating children's 'tailored' special interests would be an excellent next step in investigating the influence of interests on novel word learning. We would expect to see greater conditional differences with these particularly strong special interests, and investigating how special interests may help or hinder learning more broadly would provide valuable insight into how we can more robustly scaffold autistic children's word learning.

5.8. Thesis Conclusions

Overall, this thesis advances theoretical understanding of word learning mechanisms in autistic and neurotypical children in numerous ways. Crucially, we have illustrated the strengths of autistic children and their abilities to learn novel words despite language delay. We highlight that categorical interests, but not idiosyncratic preferences for random stimuli, reliably influence word learning in autistic children. The novel word-referent representations of autistic children are further consolidated by overnight sleep, protecting novel words from decay, and affording longer term retention. We suggest that autistic children's slower response times may be attributed to their heightened interest in stimuli. Indeed, spending longer examining target stimuli may have benefited autistic children's learning by providing increased quantity of visual and auditory input, resulting in more robust encoding of novel word-referent associations in memory. As such, we recommend that autistic children should be afforded opportunities to learn at their own pace whenever possible. Our findings also indicate that interventions would benefit from maximising children's intake of input at referent selection to ensure robust encoding of word-referent relationships. Our novel combination of looking and accuracy measures reveal that population differences in visual attention do not necessitate diminished learning - autistic children achieved similar learning outcomes to neurotypical peers via a slightly different route. However, visual attention and explicit behavioural responses did not always directly relate, so we emphasise the importance of selecting the most appropriate methodology when designing studies. Importantly, this thesis advances understanding of vocabulary acquisition in autism and emphasises that fundamental word learning mechanisms in language delayed autistic children are not impaired. Promisingly, these studies highlight several strategies for developing effective learning contexts to maximise language acquisition. Crucially, rather than constraining autistic children's word learning based on ideations from neurotypical development, our studies demonstrate that we must ensure learning contexts suit the numerous strengths of the autistic learner.

Consolidated Bibliography

Abrahams, B. S., & Geschwind, D. H. (2008). Advances in autism genetics: on the threshold of a new neurobiology. *Nature Reviews Genetics, 9,* 341-355. doi: 10.1038/nrg2346

Ackermann, L., Hepach, R., & Mani, N. (2020). Children learn words easier when they are interested in the category to which the word belongs. *Developmental Science, 23*(3). doi: 10.1111/desc.12915

Adamson, L. B., Bakeman, R., Deckner, D. F., (2004). The development of symbol-infused joint engagement. *Child Development. 75*, 1171–1187. doi: 10.1111/j.1467-8624.2004.00732.x

Adamson, L. B., Bakeman, R., Deckner, D. F., & Romski, M. (2009). Joint engagement and the emergence of language in children with autism and Down syndrome. *Journal of Autism and Developmental Disorders*, *39*, 84-96. doi: 10.1007/s10803-008-0601-7

Akechi, H., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2013). Brief report: Pointing cues facilitate word learning in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *43*(1), 230-235. doi: 10.1007/s10803-012-1555-3

Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2011). Do children with ASD use referential gaze to learn the name of an object? An eye-tracking study. *Research in Autism Spectrum Disorders*, *5*(3), 1230-1242. doi: 10.1016/j.rasd.2011.01.013

Allen, M., Hartley, C., & Cain, K. (2015). Do iPads promote symbolic understanding and word learning in children with autism? *Frontiers in Psychology, 6,* 138–138. doi: 10.3389/fpsyg.2015.00138

Allen, M. L., Hartley, C., & Cain, K. (2016). iPads and the use of "apps" by children with autism spectrum disorder: Do they promote learning?. *Frontiers in psychology*, *7*, 1305. doi: 10.3389/fpsyg.2016.01305

Ambridge, B. & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. Wiley Interdisciplinary Reviews. *Cognitive Science, 4*(2), 149–168. doi: 10.1002/wcs.1215

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

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., … &, Pickles, A. (2007). Patterns of growth in verbal abilities among children with autism spectrum disorder. *Journal of Consulting and Clinical Psychology, 75*, 594–604. doi: 10.1037/0022-006X.75.4.594.

Artis, J., & Arunachalam, S. (2023). Semantic and Syntactic Properties of Words and the Receptive-Expressive Gap in Autistic and Non-Autistic Children. Language Disorder. *Journal of Speech, Language, and Hearing Research, 66*(5), 1771-1791. doi: 10.1044/2023_JSLHR-22-00369.

Arunachalam, S., & Luyster, R. J. (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. *Autism Research, 8*, 810-828. doi: 10.1002/aur.1590.

Arunachalam, S., & Luyster, R. (2018). Lexical Development in Young Children With Autism Spectrum Disorder (ASD): How ASD May Affect Intake From the

Input. *Journal of Speech, Language, and Hearing Research: JSLHR, 61*(11), 2659-2672. doi: 10.1044/2018_JSLHR-L-RSAUT-18-0024

Attwood, T. (2003). *Understanding and managing circumscribed interests.* In M. Prior (Ed.), Learning and behaviour problems in Asperger syndrome (pp. 126–147). New York, NY: Guilford Press.

Axelsson, E., Churchley, K., & Horst, J. (2012). The right thing at the right time: Why ostensive naming facilitates word learning. *Frontiers in Psychology, 3,* 88. doi: 10.3389/fpsyg.2012.00088

Axelsson, E. L., & Horst, J. S. (2014). Contextual repetition facilitates word learning via fast mapping. *Acta Psychologica*, *152*, 95-99. doi: 10.1016/j.actpsy.2014.08.002

Axelsson, E. L., Perry, L. K., Scott, E. J., & Horst, J. S. (2016). Near or far: The effect of spatial distance and vocabulary knowledge on word learning. *Acta Psychologica, 163*, 81-87. doi: 10.1016/j.actpsy.2015.11.006

Axelsson, E. L., Swinton, J., Jiang, I. Y., Parker, E. V., & Horst, J. S. (2021). Prior exposure and toddlers' sleep-related memory for novel words. *Brain Sciences, 11*(10), 1366. doi: 10.3390/brainsci11101366

Axelsson, E. L., Swinton, J., Winiger, A. I., & Horst, J. S. (2018). Napping and toddlers' memory for fast-mapped words. *First Language*, *38*(6), 582-595. doi: 10.1177/0142723718785490

Bailey, T., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, *17*, 1267–1284. doi: 10.1016/S0885-2014(02)00116-8

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development, 62*, 875–90. doi: 10.2307/1131140. (pp. 131-158). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Baldwin, D.A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language, 20*(2), 395-418. doi: 10.1017/S0305000900008345

Baldwin, D., Markman, E., Bill, B., Desjardins, R., Irwin, J., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development,67*(6), 3135-3153. doi: 10.2307/1131771

Baldwin, D. A., & Moses, L. J. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development*, *10*(3), 309-329. doi: 10.1111/1467-9507.00168

Bannard, C., & Tomasello, M. (2012). Can we dissociate contingency learning from social learning in word acquisition by 24-month-olds? *PLoS One, 7*(11), e49881. doi: 10.1371/journal.pone.0049881

Baron-Cohen, S. (1995). The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In *Portions of this paper were presented at the Society for Research in Child Development Conference, New Orleans, Mar 1993; the British Psychological Society, Welsh Branch," Faces" Conference, U Wales Coll of Cardiff, Sep 1993; and the British Society for the Philosophy of Science" Roots of Joint Reference" Conference, U Bristol, Nov 1993.* Lawrence Erlbaum Associates, Inc.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development, 68*, 48–57. doi: 10.2307/1131924.

Baron-Cohen, S., Lombardo, M. V., Auyeung, B., Ashwin, E., Chakrabarti, B., & Knickmeyer, R. (2011). Why are autism spectrum conditions more prevalent in males?. *PLoS biology*, *9*(6), e1001081. doi: 10.1371/journal.pbio.1001081

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014 Institute for Statistics and Mathematics of WU website http://CRAN.R-project.org/package=lme4

Bean Ellawadi, A., & McGregor, K. K. (2016). Children with ASD can use gaze to map new words. *International Journal of Language & Communication Disorders, 51*(2), 212-218. doi: 10.1111/1460-6984.12192

Bedford, R., Gliga, T., Frame, K., Hudry, K., Chandler, S., Johnson, M.H., & Charman, T. (2013). Failure to learn from feedback underlies word learning difficulties in toddlers at risk for autism. *Journal of Child Language, 40(01)*, 29–46. doi: 10.1017/S0305000912000086.

Begus, K., Gliga, T., & Southgate, V. (2014). Infants Learn What They Want to Learn: Responding to Infant Pointing Leads to Superior Learning. *PLoS ONE*, 9(10). doi: 10.1371/journal.pone.0108817

Begus, K., & Southgate, V. (2012). Infant pointing serves an interrogative function. *Developmental Science, 15*(5), 611-617. doi: 10.1111/j.1467-7687.2012.01160.x

Behrmann, M., Thomas, C., & Humphreys, K. (2006). Seeing it differently: visual processing in autism. *Trends in Cognitive Sciences, 10*(6), 258–264. doi: 10.1016/j.tics.2006.05.001

Benjamin, D. P., McDuffie, A. S., Thurman, A. J., Kover, S. T., Mastergeorge, A. M., Hagerman, R. J., & Abbeduto, L. (2015). Effect of speaker gaze on word learning in fragile X syndrome: a comparison with nonsyndromic autism spectrum disorder. *Journal of Speech, Language, and Hearing Research, 58*(2), 383–95. doi: 10.1044/2015_JSLHR-L-14-0136

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America, 109*(9), 3253-3258. doi: 10.1073/pnas.1113380109

Bhat, A. N., Galloway, J. C., & Landa, R. J. (2010). Social and non-social visual attention patterns and associative learning in infants at risk for autism. *Journal of Child Psychology and Psychiatry, 51*(9), 989–997. doi: 10.1111/j.1469-7610.2010.02262.x

Bion, R.A.H., Borovsky, A., & Fernald. A. (2013). Fast Mapping, Slow Learning: Disambiguation of Novel Word-Object Mappings in Relation to Vocabulary Learning at 18, 24, and 30 Months. *Cognition, 126*(1), 39-53. doi: 10.1016/j.cognition.2012.08.008

Blair, C. (2016). Executive function and early childhood education. *Current Opinion in Behavioral Sciences, 10*, 102-107. doi: 10.1016/j.cobeha.2016.05.009

Blewitt, P., & Langan, R. (2016). Learning words during shared book reading: The role of extratextual talk designed to increase child engagement. *Journal of Experimental Child Psychology*, *150*, 404-410. doi: 10.1016/j.jecp.2016.06.009

Bloom, P. (1997). Intentionality and word learning. *Trends in Cognitive Sciences*, *1*(1), 9-12. doi: 10.1016/S1364-6613(97)01006-1

Bloom, P. (2002a). Mindreading, Communication and the Learning of Names for Things. *Mind & Language, 17*, 37-54. doi: 10.1111/1468-0017.00188

Bloom, P. (2002b). *How children learn the meanings of words*. MIT press.

Bloom, P. & Markson, L. (1998). Capacities underlying word learning. *Trends in Cognitive Sciences, 2*(2), 67–73. doi: 10.1016/S1364-6613(98)01121-8

Bölte S., & Poustka, F. (2004). Comparing the profiles of savant and non-savant individuals with autistic disorder. *Intelligence. 32,* 121–131. doi: 10.1016/j.intell.2003.11.002

Booth, A. E., McGregor, K. K., & Rohlfing, K. J. (2008). Socio-pragmatics and attention: Contributions to gesturally guided word learning in toddlers. *Language Learning and Development, 4*(3), 179– 202. doi: 10.1080/15475440802143091

Bosseler, A., Clarke, M., Tavabi, K., Larson, E., Hippe, D., Taulu, S., & Kuhl, P. (2021). Using magnetoencephalography to examine word recognition, lateralization, and future language skills in 14-month-old infants. *Developmental Cognitive Neuroscience*, 47, 100901. doi: 10.1016/j.dcn.2020.100901

Brandone, A. C., Pence, K. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Action speaks louder than words: Young children differentially weight perceptual, social, and linguistic cues to learn verbs. *Child Development*, *78*(4), 1322-1342. doi: 10.1111/j.1467-8624.2007.01068.x

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology, 38*(6), 958–966. doi: 10.1037/0012-1649.38.6.958

Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology*, *130*, 67-78. doi: 10.1016/j.jecp.2014.09.010.

Brown, H., Weighall A., Henderson L. M., Gaskell M. G. (2012). Enhanced recognition and recall of new words in 7- and 12-year-olds following a period of offline consolidation. *Journal of Experimental Child Psychology*. 112, 56–72. doi: 10.1016/j.jecp.2011.11.010

Bryson, S. E., Wainwright-Sharp, J. A., & Smith, I. M. (1990). Autism: A developmental spatial neglect syndrome? In J. Enns (Ed.), *The Development of Attention: Research and Theory* (pp. 405–427). Amsterdam: Elsevier.

Burgess, P., Alderman, N., Forbes, C., Costello, A., M-A.Coates, L., Dawson, D., . . . Channon, S. (2006). The case for the development and use of ecologically valid measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society, 12*(2), 194-209. doi: 10.1017/S1355617706060310

Calhoun, S. L., Pearl, A. M., Fernandez-Mendoza, J., Durica, K. C., Mayes, S. D., & Murray, M. J. (2020). Sleep disturbances increase the impact of working memory deficits on learning problems in adolescents with high-functioning Autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *50*, 1701-1713. doi: 10.1007/s10803-019-03928-y.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17–29.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society of Research in Child Development, 63*, 1–143. doi: 10.2307/1166214.

Carpenter, M., Pennington, B. F., & Rogers, S. J. (2002). Interrelations among social-cognitive skills in young children with autism. *Journal of Autism and Developmental Disorders*, *32*, 91-106. doi: 10.1023/a:1014836521114

Carter, C. K., & Hartley, C. (2021). Are children with autism more likely to retain object names when learning from colour photographs or black-and-white cartoons?. *Journal of Autism and Developmental Disorders*, *51*(9), 3050-3062. doi: 10.1007/s10803-020-04771-2.

Cartmill, E., Armstrong, B. F., Gleitman, L. R., Goldin-Meadows, S., Medinac, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences, 110,* 11278–11283. doi: 10.1073/pnas.1309518110.

Centers for Disease Control and Prevention. (2012). Prevalence of autism spectrum disorders—autism and developmental disabilities monitoring network, 14 sites, United States, 2008. *Morbidity and Mortality Weekly Report. Surveillance Summaries, 61,* 1–19.

Celani, G. (2002). Human Beings, Animals and Inanimate Objects: What Do People with Autism Like? *Autism: The International Journal of Research and Practice, 6*(1), 93-102. doi: 10.1177/1362361302006001007

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language, 30*(1), 213-236. doi: 10.1017/S0305000902005482

Chawarska, K., Macari, S., & Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *Journal of Child Psychology and Psychiatry, 53*(8), 903-913. doi: 10.1111/j.1469-7610.2012.02538.x

Coonrod, E. E., & Stone, W. L. (2004). Early concerns of parents of children with autistic and nonautistic disorders. *Infants & Young Children*, *17*(3), 258-268. doi: 10.1097/00001163-200407000-00007.

Courchesne, E., Townsend, J., Akshoomoff, N., Saitoh, O., Yeung-Courchesne, R., Lincoln, A., . . . Lau, L. (1994). Impairment in Shifting Attention in Autistic and Cerebellar Patients. *Behavioral Neuroscience, 108*(5), 848-865. doi: 10.1037/0735-7044.108.5.848

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers, 28,* 125-127.

Dautriche, I., Swingley, D., & Christophe, A. (2015). Learning novel phonological neighbors: Syntactic category matters. *Cognition*, *143*, 77-86. doi: 10.1016/j.cognition.2015.06.003

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society B, 364*, 3773-3800. doi: 10.1098/rstb.2009.0111

De Fossé, L., Hodge, S., Makris, N., Kennedy, D., Caviness, D., & McGrath, L. (2004). Language-association cortex asymmetry in autism and specific language impairment. *Annals of Neurology, 56,* 757-766. doi: 10.1002/ana.20275.

de Marchena, A., Eigsti, I.-M., Worek, A., Ono, K. E., & Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: Testing the pragmatic hypothesis. *Cognition, 119*(1), 96–113. doi: 10.1016/j.cognition.2010.12.011.

Devnani, P. A., & Hegde, A. U. (2015). Autism and sleep disorders. *Journal of Pediatric Neurosciences,10*(4),304-7. doi: 10.4103/1817-1745.174438.

Díaz-Román, A., Zhang, J., Delorme, R., Beggiato, A., & Cortese, S. (2018). Sleep in youth with autism spectrum disorders: systematic review and meta-analysis of subjective and objective studies. *Evidence-based Mental Health*, *21*(4), 146-154. doi: 10.1136/ebmental-2018-300037

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience, 11*(2), 114–126. doi: 10.1038/nrn2762

Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep Medicine Reviews 13*, 309–321. doi: 10.1016/j.smrv.2008.08.002

DiStefano, C., & Kasari, C. (2016). The window to language is still open: distinguishing between preverbal and minimally verbal children with ASD. *Perspectives of the ASHA Special Interest Groups*, *1*(1), 4-11. doi: 10.1044/persp1.SIG1.4.

Drosopoulos, S., Schulze, C., Fischer, S., & Born, J. (2007). Sleep's Function in the Spontaneous Recovery and Consolidation of Memories. *Journal of Experimental Psychology, 136*(2), 169–183. doi: 10.1037/0096-3445.136.2.169

Dumay, N., & Gaskell, M. G. (2007). Sleep associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35-39. doi: 10.1111/j.1467-9280.2007.01845.x.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British picture vocabulary scale* (2nd ed.), NFER-Nelson, Windsor.

Durkin, K., Conti-Ramsden, G., & Simkin Z. (2012). Functional outcomes of adolescents with a history of specific language impairment (SLI) with and without autistic symptomatology. *Journal of Autism and Developmental Disorders, 42*, 123–138. doi: 10.1007/s10803-011-1224-y.

Eigsti, I. M., de Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, *5*(2), 681-691. doi: 10.1016/j.rasd.2010.09.001.

El Zein, F., Gevarter, C., Bryant, B., Son, S-H., Bryant, D., Kim, M., Solis, M. (2016). A Comparison between iPad-Assisted and Teacher-Directed Reading Instruction for Students with Autism Spectrum Disorder (ASD). *Journal of Developmental and Physical Disabilities, (28)*2, 195-215. doi: 10.1007/s10882-015-9458-9

Elison, J. T., Sasson, N. J., Turner-Brown, L. M., Dichter, G. S., & Bodfish, J. W. (2012). Age trends in visual exploration of social and nonsocial information in children with autism. *Research in Autism Spectrum Disorders, 6*(2), 842–851. doi: 10.1016/j.rasd.2011.11.005

Ellis, E., Robledo Gonzalez, M. & Deák, G. (2014). Visual Prediction in Infancy: What is the Association with Later Vocabulary?. *Language Learning and Development*. 10. 36-50. doi: 10.1080/15475441.2013.799988.

Ellis Weismer, S., & Kover, S. T. (2015). Preschool language variation, growth, and predictors in children on the autism spectrum. *Journal of Child Psychology and Psychiatry*, *56*(12), 1327-1337. doi: 10.1111/jcpp.12406.

Elrod, M. G., & Hood, B. S. (2015). Sleep differences among children with autism spectrum disorders and typically developing peers: A meta-analysis. *Journal of Developmental & Behavioral Pediatrics, 36,* 166–177. doi: 10.1097/DBP.0000000000000140

Elsabbagh, M., Fernandes, J., Webb, S. J., Dawson, G., Charman, T., & Johnson, M. H. (2013). Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. *Biological Psychiatry, 74*,189–194. doi: 10.1016/j.biopsych.2012.11.030

Elsabbagh, M., Volein, A., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, S., & Johnson, M.H. (2009). Visual orienting in the early broader autism phenotype: Disengagement and facilitation. *Journal of Child Psychology and Psychiatry, 50*, 637–642. doi: 10.1111/j.1469-7610.2008.02051.x

Fantasia, V., Markant, D. B., Valeri, G., Perri, N., & Ruggeri, A. (2020). Memory enhancements from active control of learning in children with autism spectrum disorder. *Autism, 24*(8), 1995–2007. doi: 10.1177/1362361320931244

Fenson, L., Dale, P., Reznick, S.J., Bates, E., Thal, D.J. & Pethick, S.J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development, 59*, 1-173. doi: 10.2307/1166093

Ferman. (2021). Morpholexical Processes in Children with Autism Spectrum Disorder: Evidence from Artificial Word Learning. *Folia Phoniatrica et Logopaedica, 73*(3), 248–262. doi: 10.1159/000512849

Fernald, A. Perfors, A., & Marchman, V. A. (2006). Picking Up Speed in Understanding. *Developmental Psychology, 42*(1), 98–116. doi: 10.1037/0012-1649.42.1.98

Fitneva, S. A, & Christiansen, M. H. (2011). Looking in the Wrong Direction Correlates with More Accurate Word Learning. *Cognitive Science, 35*(2), 367-380. doi: 10.1111/j.1551-6709.2010.01156.x

Fletcher, F. E., Foster-Owens, M. D., Conduit, R., Rinehart, N. J., Riby, D. M., & Cornish, K. M. (2017). The developmental trajectory of parent-report and objective sleep profiles in autism spectrum disorder: Associations with anxiety and bedtime routines. *Autism, 21*(4), 493-503. doi: 10.1177/136236131665336

Fletcher, F., Knowland, V., Walker, S., Gaskell, M. G., Norbury, C., & Henderson, L. M. (2020). Atypicalities in sleep and semantic consolidation in autism. *Developmental Science, 23*(3), e12906. doi: 10.1111/desc.12906

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

Fletcher-Watson, S., Petrou, A., Scott-Barrett, J., Dicks, P., Graham, C., O'Hare, A., … McConachie, Helen. (2016). A Trial of an iPad™ Intervention Targeting Social Communication Skills in Children with Autism. *Autism: The International Journal of Research and Practice, 20*(7), 771-782. doi: 10.1177/1362361315605624

Foti, F., De Crescenzo, F., Vivanti, G., Menghini, D., & Vicari, S. (2015). Implicit learning in individuals with autism spectrum disorders: A meta-analysis. *Psychological Medicine, 45*(5), 897-910. doi: 10.1017/S0033291714001950

Frank, I., & Poulin-Dubois, D. (2002). Young monolingual and bilingual children’s responses to violation of the Mutual Exclusivity Principle. *International Journal of Bilingualism*, *6*(2), 125–146. doi: 10.1177/13670069020060020201

Friedrich, M., & Friederici, A.D. (2011). Word learning in 6-month-olds: Fast encoding – weak retention. *Journal of Cognitive Neuroscience, 23*(11), 3228–3240. doi: 10.1162/jocn_a_00002

Friedrich, M., Wilhelm, I., Born, J., & Friederici, A. D. (2015). Generalization of word meanings during infant sleep. *Nature Communications*, *6*, 6004. doi: 10.1038/ncomms7004

Frith, U. (1989). *Autism: Explaining the Enigma*. Oxford: Basil Blackwell.

Gais, S. & Born, J. (2004). Declarative memory consolidation: mechanisms acting during human sleep. *Learning & Memory, 11*(6), 679–685. doi: 10.1101/lm.80504

Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. *Learning & Memory. 13,* 259–262. doi: 10.1101/lm.132106

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychological Bulletin, 134*(1), 31. doi: 10.1037/0033-2909.134.1.31

Gavelek, J. R., & Raphael, T. E. (1985). Metacognition, instruction, and the role of questioning activities. *Metacognition, cognition, and human performance: Instructional practices*, *2*, 103-136. doi: 10.1016/B978-0-12-262302-8.50008-0

Gelman, S. (2003). *The Essential Child. Origins of Essentialism in Everyday Thought*. Oxford University Press; New York. doi: 10.1093/acprof:oso/9780195154061.001.0001

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*(1), 3–55. doi: 10.1207/s15327817la0101_2

Gliga, T., Elsabbagh, M., Hudry, K., Charman, T., Johnson, M. H., & BASIS team. (2012). Gaze following, gaze reading, and word learning in children at risk for autism. *Child Development*, *83*(3), 926-938. doi: 10.1111/j.1467-8624.2012.01750.x

Gliga, T., Skolnick, A., Liersch, U., Charman, T., Johnson, M. H., Bedford, R., & BASIS Team. (2022). Investigating the mechanisms driving referent selection and retention in toddlers at typical and elevated likelihood for autism spectrum disorder. *Journal of Child Language*, *49*(5), 1024-1036. doi: 10.1017/S0305000921000416

Goldman, S. (2013). Opinion: Sex, gender and the diagnosis of autism—A biosocial view of the male preponderance. *Research in Autism Spectrum Disorders,7*(6), 675–679. doi: 10.1016/j.rasd.2013.02.006

Goldstein, H. (2002). Communication intervention for children with autism: a review of treatment efficacy. *Journal of Autism and Developmental Disorders, 32*, 373–396. doi: 10.1023/a:1020589821992

Gómez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science, 17,* 670–674. doi: 10.1111/j.1467-9280.2006.01764.x

Goodlin-Jones, B., Tang, K., Liu, J., & Anders, T. F. (2008). Sleep Patterns in Preschool-Age Children With Autism, Developmental Delay, and Typical Development. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(8), 930–938. doi: 10.1097/CHI.0b013e3181799f7c

Gureckis, T. M., & Markant, D. B. (2012). Self-directed learning a cognitive and computational perspective. *Perspectives on Psychological Science, 7*(5), 464–481. doi: 10.1177/1745691612454304

Gurteen, P.M., Horne, P. J., & Erjavec, M. (2011). Rapid word learning in 13- and 17-month-olds in a naturalistic two-word procedure: Looking versus reaching measures. *Journal of Experimental Child Psychology, 109*(2), 201–217. doi: 10.1016/j.jecp.2010.12.001

Haebig, E., Saffran, J.R., & Ellis Weismer, S. (2017). Statistical word learning in children with autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry, 58*(11), 1251-1263. doi: 10.1111/jcpp.12734

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*(1), B23-B34. doi: 10.1016/S0010-0277(02)00186-5

Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology, 53*(4), 310–344. doi: 10.1016/j.cogpsych.2006.04.003

Hani, H. B., Gonzalez-Barrero, A. M., & Nadig, A. S. (2013). Children's referential understanding of novel words and parent labeling behaviors: Similarities across children with and without autism spectrum disorders. *Journal of Child Language, 40*(5), 971-1002. doi: 10.1017/S0305000912000426

Happé, F. (1995). The role of age and verbal ability in the theory of mind task: Performance of subjects with autism. *Child Development, 66*, 843-855. doi: 10.2307/1131954

Happé, F. (1999). Autism: cognitive deficit or cognitive style? *Trends in Cognitive Sciences, 3*, 216–222. doi: 10.1016/S1364-6613(99)01318-2

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. doi: 10.1007/s10803-005-0039-0

Hartley, C., & Allen, M. L. (2014). Brief Report: Generalisation of Word–Picture Relations in Children with Autism and Typically Developing Children. *Journal of Autism and Developmental Disorders*, *44,* 2064–2071. doi: 10.1007/s10803-014-2074-1

Hartley, C., Bird, L., & Monaghan, P. (2019). Investigating the relationship between fast mapping, retention, and generalisation of words in children with autism spectrum disorder and typical development. *Cognition, 187*, 126-138. doi: 10.1016/j.cognition.2019.03.001

Hartley, C., Bird, L., & Monaghan, P. (2020). Comparing cross-situational word learning, retention, and generalisation in children with autism and typical development. *Cognition, 200*, 104265. doi: 10.1016/j.cognition.2020.104265

Henderson, L., Powell, A., Gareth Gaskell, M., & Norbury, C. (2014). Learning and consolidation of new spoken words in autism spectrum disorder. *Developmental Science, 17*(6), 858-871. doi: 10.1111/desc.12169

Henderson, L., Van Rijn, E., James, E., Walker, S., Knowland, V., & Gaskell, M. G. (2021). Timing storytime to maximize children's ability to retain new vocabulary. *Journal of Experimental Child Psychology, 210,* 105207. doi: 10.1016/j.jecp.2021.105207

Henderson, L., Weighall, A., & Gaskell, M. G. (2013). Learning new vocabulary during childhood: effects of semantic training on lexical consolidation and integration. *Journal of Experimental Child Psychology, 116,* 572–592. doi: 10.1016/j.jecp.2013.07.004

Henderson, L.M., Weighall, A.R., Brown, H., & Gaskell, M.G. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science, 15*, 674–87. doi: 10.1111/j.1467-7687.2012.01172.x

Hill, C. M., Hogan, A. M., & Karmiloff-Smith, A. (2007). To sleep, perchance to enrich learning? *Archives of Disease in Childhood, 92,* 637–643. doi: 10.1136/adc.2006.096156

Hilton, M., & Westermann, G. (2017). The effect of shyness on children's formation and retention of novel word–object mappings. *Journal of Child Language*, *44*(6), 1394–1412. doi: 10.1017/S030500091600057X

Hilton, M., Twomey, K. E., & Westermann, G. (2019). Taking their eye off the ball: How shyness affects children's attention during word learning. *Journal of Experimental Child Psychology*, *183*, 134–145. doi: 10.1016/j.jecp.2019.01.023

Hilton, M., Twomey, K. E., & Westermann, G (2023). Face time: Effects of shyness and attention to faces on early word learning. *Language Development Research*. doi: 10.34842/2023.652

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., … Rocroi, C. (2000). Breaking the Language Barrier: An Emergentist Coalition Model for the Origins of Word Learning. *Monographs of the Society for Research in Child Development, 65*(3), 1-123. doi: 10.1111/1540-5834.00090

Honey, E., Rodgers, J., & McConachie, H. (2012). Measurement of restricted and repetitive behaviour in children with autism spectrum disorder: Selecting a questionnaire or interview. *Research in Autism Spectrum Disorders, 6*(2), 757–776. doi: 10.1016/j.rasd.2011.10.011

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behaviour Research Methods, 48*(4), 1393-1409. doi: 10.3758/s13428-015-0647-3

Horst, J. S., Parsons, K. L, & Bryan, N. M. (2011). Get the story straight: Contextual repetition promotes word learning from storybooks. *Frontiers in Psychology, 2*,17. doi: 10.3389/fpsyg.2011.00017

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy, 13*, 128–157. doi: 10.1080/15250000701795598.

Horst, J., Scott, E. J., & Pollard, J. A. (2010). The role of competition in word learning via referent selection. *Developmental Science, 13*(5), 706–713. doi: 10.1111/j.1467-7687.2009.00926.x

Horst, J. S., Twomey, K. E., Morse, A. F., Nurse, R., & Cangelosi, A. (2019). When object color is a red herring: extraneous perceptual information hinders word learning via referent selection. *IEEE Transactions on Cognitive and Developmental Systems*, *12*(2), 222-231. doi: 10.1109/TCDS.2019.2894507

Horváth, K., Liu, S., & Plunkett, K. (2016). A daytime nap facilitates generalization of word meanings in young toddlers. *Sleep*, *39*(1), 203-207. doi: 10.5665/sleep.5348

Horváth, K., Myers, K., Foster, R., & Plunkett, K. (2015). Napping facilitates word learning in early lexical development. *Journal of Sleep Research, 24,* 503–509. doi: 10.1111/jsr.12306

Horváth, K., & Plunkett, K. (2016). Frequent daytime naps predict vocabulary growth in early childhood. *Journal of Child Psychology and Psychiatry*, *57*(9), 1008-1017. doi: 10.1111/jcpp.12583.

Horváth, K., & Plunkett, K. (2018). Spotlight on daytime napping during early childhood. *Nature and science of sleep*, 97-104. doi: 10.2147/NSS.S126252

Howlin, P. (2003). Outcome in high-functioning adults with autism with and without early language delays: implications for the differentiation between autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *33*(1), 3-13. doi: 10.1111/1460-6984.12192.

Howlin, P., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *American Journal of Intellectual Developmental Disabilities, 114*, 23–41. doi: 10.1352/2009.114:23;nd41

Hughes, C. Russell, J. & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia*, *32*, 4, 477–492. doi: 10.1016/0028-3932(94)90092-2

Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M., & William, M. (2017). “Putting on my best normal”: Social camouflaging in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders, 47*(8), 2519–2534. doi: 10.1007/s10803-017-3166-5.

Hunter, M. A., Ames, E. W. (1988) A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research, 5*, 69–95.

Hupbach, A., Gómez, R. L., Bootzin, R. R., & Nadel, L. (2009). Nap dependent learning in infants. *Developmental Science, 12,* 1007–1012. doi: 10.1111/j.1467-7687.2009.00837.x

Hutman, T., Chela, M. K., Gillespie-Lynch, K., & Sigman, M. (2012). Selective visual attention at twelve months: Signs of autism in early social interactions. *Journal of Autism and Developmental Disorders, 42,* 487–498. doi: 10.1007/s10803-011-1262-5

Hyde, K. L., Samson, F., Evans, A. C., & Mottron, L. (2010). Neuroanatomical differences in brain areas implicated in perceptual and other core features of autism revealed by cortical thickness analysis and voxel-based morphometry. Human Brain Mapping, 31, 556-566. doi: 10.1002/hbm.20887

Ibanez, L. V., Messinger, D. S., Newell, L., Lambert, B., & Sheskin, M. (2008). Visual disengagement in the infant siblings of children with an autism spectrum disorder (ASD*). Autism: The International Journal of Research and Practice, 12*(5), 473–485. doi: 10.1177/1362361308094504

James, E., Gaskell, M. G., & Henderson, L. M. (2020). Sleep-dependent consolidation in children with comprehension and vocabulary weaknesses: it’ll be alright on the night?. *Journal of Child Psychology and Psychiatry, 61*(10), 1104-1115. doi: 10.1111/jcpp.13253

James, E., Gaskell, M., Weighall, A., & Henderson, L. (2017). Consolidation of vocabulary during sleep: The rich get richer? *Neuroscience and Biobehavioral Reviews, 77*, 1-13. doi: 10.1016/j.neubiorev.2017.01.054

Jones, S., Smith, L., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development*, *62*(3), 499–516. doi: 10.1111/j.1467-8624.1991.tb01547.x

Joseph, R. M., McGrath, L. M., & Tager-Flusberg, H. (2005). Executive dysfunction and its relation to language ability in verbal school-age children with autism. *Developmental Neuropsychology*, *27*(3), 361-78. doi: 10.1207/s15326942dn2703_4.

Kalashnikova, M., Mattock, K., & Monaghan, P. (2014). Disambiguation of novel labels and referential facts: A developmental perspective. *First Language, 34,* 125-135. doi: 10.1177/0142723714525946

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2,* 217–250.

Kapa, L. L., & Erikson, J. A. (2020). The Relationship Between Word Learning and Executive Function in Preschoolers With and Without Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research, 63*(7), 2293–2307. doi: 10.1044/2020_JSLHR-19-00342

Kasari, C., Paparella, T., Freeman, S., & Jahromi, L. B. (2008). Language outcome in autism: randomized comparison of joint attention and play interventions. *Journal of Consulting and Clinical Psychology, 76*(1), 125. doi: 10.1037/0022-006X.76.1.125

Kidd, E., & Donnelly, S. (2020). Individual differences in first language acquisition. *Annual Review of Linguistics*, *6*, 319-340. doi: 10.1146/annurev-linguistics-011619-030326

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS One*, *7*(5), e36399. doi: 10.1371/journal.pone.0036399

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, *85*(5), 1795-1804. doi: 10.1111/cdev.12263

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, *16(2-3)*, 287-308. doi: 10.1080/01690960042000058.

Klin, A., Danovitch, J., Merz, A., & Volkmar, F. (2007) Circumscribed interests in higher functioning individuals with autism spectrum disorders: an exploratory study. *Research & Practice for Persons with Severe Disabilities, 32*, 89–100. doi: 10.2511/rpsd.32.2.89

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, *459*(7244), 257-261. doi: 10.1038/nature07868

Kotagal, S., & Broomall, E. (2012). Sleep in children with autism spectrum disorder. *Pediatric Neurology, 47*(4), 242–51.2. doi: 10.1016/j.pediatrneurol.2012.05.007

Kover, S. T., McDuffie, A. S., Hagerman, R. J., & Abbeduto, L. (2013). Receptive vocabulary in boys with autism spectrum disorder: Cross-sectional developmental trajectories. *Journal of Autism and Developmental Disorders, 43,* 2696-2709. doi: 10.1007/s10803-013-1823-x

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2015). Slowing down fast mapping: Redefining the dynamics of word learning. *Child Development Perspectives, 9(2)*, 74–78. doi: 10.1111/cdep.12110

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2020). Sometimes it is better to know less: How known words influence referent selection and retention in 18- to 24-month-old children. *Journal of Experimental Child Psychology, 189,* 104705–104705. doi: 10.1016/j.jecp.2019.104705

Kucker, S. C., & Samuelson, L. K. (2012). The first slow step: Differential effects of object and word-form familiarization on retention of fast-mapped words. *Infancy, 17*(3), 295-323.

Kucker, S.C. & Seidler, E. (2023). The timescales of word learning in children with language delays: In-the-moment mapping, retention, and generalization. *Journal of Child Language, 50*(2), 245–29. doi: 10.1017/S0305000921000817

Kurdziel, L., Duclos, K., & Spencer, R. M. (2013). Sleep spindles in mid-day naps enhance learning in preschool children. *Proceedings of the National Academy of Sciences, 110*(43), 17267–17272. doi: 10.1073/pnas.1306418110

Lai, M. C., Lombardo, M. V., & Baron-Cohen, S. (2014). Autism. *The Lancet, 389*, 896-910. doi: 10.1016/S0140-6736(13)61539-1

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3,* 299 – 321. doi: 10.1016/0885-2014(88)90014-7

Landry, R., & Bryson, S.E. (2004). Impaired disengagement of attention in young children with autism. *Journal of Child Psychology and Psychiatry, 45*, 1115–1122. doi: 10.1111/j.1469-7610.2004.00304.x

Landry, S. H., & Loveland, K. A. (1988). Communication behaviors in autism and developmental language delay. *Journal of Child Psychology and Psychiatry, 29*(5), 621-634. doi: 10.1111/j.1469-7610.1988.tb01884.x

Larkin, F., Meins, E., Centifanti, L. C., Fernyhough, C., & Leekam, S. R. (2017). How does restricted and repetitive behavior relate to language and cognition in typical development?. *Development and Psychopathology*, *29*(3), 863-874. doi: 10.1017/S0954579416000535

Lau, H., Tucker, M. A., & Fishbein, W. (2010). Daytime napping: effects on human direct associative and relational memory. *Neurobiology of Learning and Memory. 93,* 554–560. doi: 10.1016/j.nlm.2010.02.003

Lee, A., Lang, R., Davenport, K., Moore, M., Rispoli, M., Van Der Meer, L., . . . Chung, C. (2015). Comparison of therapist implemented and iPad-assisted interventions for children with autism. *Developmental Neurorehabilitation, 18*(2), 97-103. doi: 10.3109/17518423.2013.830231

Leekam, S., Prior, M., & Uljarevic, M. (2011). Restricted and Repetitive Behaviors in Autism Spectrum Disorders: A Review of Research in the Last Decade. *Psychological Bulletin, 137*(4), 562-593. doi: 10.1037/a0023341

Leekam, S. R., & Ramsden, C. A. (2006). Dyadic orienting and joint attention in preschool children with autism. *Journal of Autism and Developmental Disorders, 36*, 185-197. doi: 10.1007/s10803-005-0054-1

Leekam, S., Tandos, J., McConachie, H., Meins, E., Parkinson, K., Wright, C., Turner, M., Arnott, B., Vittorini, L., & Le Couteur, A. (2007). Repetitive behaviours in typically developing 2-year-olds. *Journal of Child Psychology and Psychiatry, 48*(11), 1131-1138. doi: 10.1111/j.1469-7610.2007.01778.x

Leung, E. H., & Rheingold, H. L. (1981). Development of pointing as a social gesture. *Developmental Psychology, 17*, 215–220. doi: 10.1037/0012-1649.17.2.215

Lewis, M., Cristiano, V., Lake, B., Kwan, T., & Frank, M. (2020). The role of developmental change and linguistic experience in the mutual exclusivity effect. *Cognition*, *198,* 104191. doi: 10.1016/j.cognition.2020.104191

Lopez, B. R., Lincoln, L. A. J., Ozonoff, S., & Lai, Z. (2005). Examining the relationship between executive functions and restricted, repetitive symptoms of autistic disorder. *Journal of Autism and Developmental Disorders, 35*(4), 445–460. doi: 10.1007/s10803-005-5035-x

Lord, C. & Paul, R. (1997). Language and communication in autism. In D.J. Cohen & F.R. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders* (2nd ed.). (pp. 195–225). New York: John Wiley and Sons.

Lord, C., Risi, S., & Pickles, A. (2004). Trajectory of language development in autistic spectrum disorders. In *Developmental Language Disorders* (pp. 18-41). Psychology Press.

Lord, C., Rutter, M., & Le Couteur, A. (1994) Autism Diagnostic Interview – Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659-685. doi: 10.1007/BF02172145

Lord, C., Rutter, M., DiLavore, P.C., & Risi, S. (2002). *Autism diagnostic observation schedule.* (WPS ed.), Western Psychological Services, Los Angeles.

Louis, J., Cannard, C., Bastuji, H., & Challamel, M.-J. (1997). Sleep ontogenesis revisited: a longitudinal 24-hour home polygraphic study on 15 normal infants during the first two years of life. *Sleep, 20,* 323–333. doi: 10.1093/sleep/20.5.323

Lovaas, O. I., Koegel, R. L., & Schreibman, L. (1979). Stimulus overselectivity in autism: A review of research. *Psychological Bulletin, 86,* 1236–1254. doi: 10.1037/0033-2909.86.6.1236

Lovaas, O. I., Schreibman, L., Koegel, R., & Rehm, R. (1971). Selective responding by autistic children in multiple sensory input. *Journal of Abnormal Psychology, 77*, 211–222. doi: 10.1037/h0031015

Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language assessment and development in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38(8)*, 1426–1438. doi: 10.1007/s10803-007-0510-1.

Luyster, R., Lopez, K., & Lord, C. (2007) Characterizing communicative development in children referred for autism spectrum disorders using the MacArthur-Bates Communicative Development Inventory. *Journal of Child Language, 34*(3), 623-654. doi: 10.1017/s0305000907008094.

Luyster, R., & Lord, C. (2009). Word Learning in Children with Autism Spectrum Disorders. *Developmental Psychology, 45*(6), 1774–1786. doi: 10.1037/a0016223.

Ma, L., Twomey, K., & Westermann, G. (2022). The impact of perceived emotions on toddlers' word learning. *Child Development*, *93*(5), 1584–1600. doi: 10.1111/cdev.13799

Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word Learning in Infant- and Adult-Directed Speech. *Language Learning and Development, 7*(3), 185–201. doi: 10.1080/15475441.2011.579839

Macnamara J. (1972). Cognitive basis of language learning in infants. *Psychological Review, 79,* 1–13. doi: 10.1037/h0031901

MacWhinney, B. (1991). A reply to Woodward and Markman. *Developmental Review, 11,* 192-194. doi: 10.1016/0273-2297(91)90007-B

Markant, D., DuBrow, S., Davachi, L., & Gureckis, T. M. (2014). Deconstructing the effect of self-directed study on episodic memory. *Memory & Cognition, 42*(8), 1211– 1224. doi: 10.3758/s13421-014-0435-9

Markman, E. M. (1989). Mutual exclusivity. In E. M. Markman (Ed.), *Categorization and naming in children: Problems of induction* (pp. 187-215). Cambridge, MA: MIT Press

Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S. A. Gelman & J. P. Byrnes (Eds.), Perspectives on language and thought: Interrelations in development (pp. 72-106). New York, NY, US: Cambridge University Press. doi: 10.1017/CBO9780511983689.004

Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua*, *92*, 199-227. doi: 10.1016/0024-3841(94)90342-5

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16*(1), 1-27. doi: 10.1016/0010-0285(84)90002-1

Markman, E. M., & Wachtel, G. F. (1998). Children’s use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20,* 121-157. doi: 10.1016/0010-0285(88)90017-5

Martin, F., & Farnum, J. (2002). Animal-Assisted Therapy for Children with Pervasive Developmental Disorders. *Western Journal of Nursing Research, 24*(6), 657-670. doi: 10.1177/019394502320555403

Maski, K., Holbrook, H., Manoach, D., Hanson, E., Kapur, K., & Stickgold, R. (2015). Sleep dependent memory consolidation in children with autism spectrum disorder. *SLEEP, 38*(12), 1955–1963. doi: 10.5665/sleep.5248

Mathée-Scott, J., Larson, C., Venker, C., Pomper, R., Edwards, J., Saffran, J., & Ellis Weismer, S. (2021). Use of Mutual Exclusivity and its Relationship to Language Ability in Toddlers with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *52*(10), 4528–4539. doi: 10.1007/s10803-021-05321-0

Matson, J., & Neal, D. (2010). Differentiating communication disorders and autism in children. *Research in Autism Spectrum Disorders. 4,* 626-632. doi: 10.1016/j.rasd.2009.12.006.

McDuffie, A., Yoder, P., & Stone, W. (2006a). Fast-mapping in young children with autism spectrum disorders. *First Language, 26*(4), 421-438. doi: 10.1177/0142723706067438

McDuffie, A., Yoder, P., & Stone, W. (2006b). Labels increase attention to novel objects in children with autism and comprehension-matched children with typical development. *Autism*, *10*(3), 288–301. doi: 10.1177/1362361306063287

McGregor, K. K., Rost, G., Arenas, R., Farris-Trimble, A., & Stiles, D. (2013). Children with ASD can use gaze in support of word recognition and learning. *Journal of Child Psychology and Psychiatry, 54(7)*, 745–753. doi: 10.1111/jcpp.12073.

McGregor, K. K., Smolak, E., Jones, M., Oleson, J., Eden, N., Arbisi-Kelm, T., & Pomper, R. (2022). What children with developmental language disorder teach us about cross-situational word learning. *Cognitive Science, 46*(2), e13094. doi: 10.1111/cogs.13094

McMurray, B., Horst, J., & Samuelson, L. (2012). Word Learning Emerges From the Interaction of Online Referent Selection and Slow Associative Learning. *Psychological Review*, *119*(4), 831-877. doi: 10.1037/a0029872

Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development, 54*(3-4)*,* 1-132,

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name nameless category (N3c) principle. *Child Development, 65,* 16461 662. doi: 10.2307/1131285

Ming, X., Brimacombe, M., Chaaban, J., Zimmerman-Bier, B., & Wagner, G. C. (2008). Autism spectrum disorders: concurrent clinical disorders. *Journal of Child Neurology, 23*(1):6–13.4. doi: 10.1177/0883073807307102

Mody, M., & Belliveau, J. W. (2013). Speech and Language Impairments in Autism: Insights from Behavior and Neuroimaging. *North American Journal of Medical Sciences, 5*(3), 157-161. doi: 10.7156/v5i3p157.

Monaghan, P. (2017). Canalization of language structure from environmental constraints: A computational model of word learning from multiple cues. *Topics in Cognitive Science,* 9, 21-34. doi: 10.1111/tops.12239.

Monaghan, P., & Mattock, K. (2012). Integrating constraints for learning word referent mappings. *Cognition, 123*, 133-143. doi: 10.1016/j.cognition.2011.12.010

Montessori, M. (1912). *A critical consideration of the new pedagogy in its relation to modern science*. The Montessori method (A. E. George (Trans.) (pp. 1-27). New York: Frederick A. Stokes Company.

Montgomery-Downs, H. E., O'Brien, L. M., Gulliver, T. E., & Gozal, D. (2006). Polysomnographic characteristics in normal preschool and early school-aged children. *Pediatrics* 117, 741–753. doi: 10.1542/peds.2005-1067

Mullen, E. M. (1995). *Mullen scales of early learning* (AGS). Circle Pines, MN: American Guidance Service.

Mundy, P., Sigman, M., Ungerer, J., & Sherman, T. (1986). Defining the social deficits of autism: the contribution of non-verbal communication measures. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 27*(5), 657–669. doi: 10.1111/j.1469-7610.1986.tb00190.x

Nelson, C. A. (2007). A Neurobiological Perspective on Early Human Deprivation. *Child Development Perspectives, 1*, 13–18. doi: 10.1111/j.1750-8606.2007.00004.x

Neveu, A., & Kaushanskaya, M. (2023). Paired-associate versus cross-situational: How do verbal working memory and word familiarity affect word learning?. *Memory & Cognition,* 1-13. doi: 10.3758/s13421-023-01421-7

Norbury, C. F., Griffiths, H. & Nation, K. (2010). Sound before meaning: Word learning in autistic disorders. *Neuropsychologia*, *48*, 14, 4012–4019. doi: 10.1016/j.neuropsychologia.2010.10.015.

Norrelgen, F., Fernell, E., Eriksson, M., Hedvall, Å., Persson, C., Sjölin, M., ... & Kjellmer, L. (2015). Children with autism spectrum disorders who do not develop phrase speech in the preschool years. *Autism*, *19*(8), 934-943. doi: 10.1177/1362361314556782.

Noterdaeme, M., Mildenberger, K., Minow, F., & Amorosa, H. (2002). Evaluation of neuromotor deficits in children with autism and children with a specific speech and language disorder. *European Child & Adolescent Psychiatry, 11*(5), 219–225. doi: 10.1007/s00787-002-0285-z

Nowell, K. P., Bernardin, C. J., Brown, C., & Kanne, S. (2021). Characterization of special interests in autism spectrum disorder: A brief review and pilot study using the special interests survey. *Journal of Autism and Developmental Disorders*, *51*, 2711-2724. doi: 10.1007/s10803-020-04743-6

Oakes, L. (2011). *Infant perception and cognition recent advances, emerging theories, and future directions*. New York ; Oxford: Oxford University Press.

Omaki, A. & Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition, 22*, 158–192. doi: 10.1080/10489223.2014.943903

Oudeyer, P., & Smith, L. (2016). How Evolution May Work Through Curiosity-Driven Developmental Process. *Topics in Cognitive Science, 8*(2), 492-502. doi: 10.1111/tops.12196

Ozonoff, S., Cook, I, Coon, H, Dawson, G., Joseph, R. M., Klin, A., . . . Wrathall, D. (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. doi: 10.1023/B:JADD.0000022605.81989.cc

Ozonoff, S., Pennington, B. F., & Rogers. S. J. (1991). Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *32*, 7, 1081–1105. doi: 10.1111/j.1469-7610.1991.tb00351.x

Ozonoff, S., Strayer, D., McMahon, W., & Filloux, F. (1994). Executive Function Abilities in Autism and Tourette Syndrome: An Information Processing Approach. *Journal of Child Psychology and Psychiatry, 35*(6), 1015-1032. doi: 10.1177/1087054714550336

Palincsar, A. S., & Brown, D. A. (1987). Enhancing instructional time through attention to metacognition. *Journal of learning disabilities*, *20*(2), 66-75. doi: 10.1177/002221948702000201

Parish-Morris, J., Hennon, E. A., Hirsh-Pasek, K., Golinkoff, R. M., & Tager-Flusberg, H. (2007). Children with autism illuminate the role of social intention in word learning. *Child Development, 78*, 1265–1287. doi: 10.1111/j.1467-8624.2007.01065.x.

Parsons, J. P., Bedford, R., Jones, E. J., Charman, T., Johnson, M. H., & Gliga, T. (2019). Gaze following and attention to objects in infants at familial risk for ASD. *Frontiers in psychology, 10,* 1799. doi: 10.3389/fpsyg.2019.01799

Partridge, E., McGovern, M. G., Yung, A., & Kidd, C. (2015). *Young children's self-directed information gathering on touchscreens.* In D. C. Noelle, R. Dale, A. A. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), Proceedings of the37th annual meeting of the Cognitive Science Society (pp. 1835– 1840). Austin, TX: Cognitive Science Society.

Patael, S., & Diesendruck, G. (2008). Intentions help children learn meaningful rules. *Journal of Child Language, 35,* 221–234. doi: 10.1017/ S0305000907008331.

Pereira, A.F., Smith, L.B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review, 21*, 178–185. doi: 10.3758/s13423-013-0466-4.

Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, *85*(1), 53-78. doi: 10.1016/s0010-0277(02)00073-2

Piaget, J. (1970). *Structuralism*. New York: Harper & Row.

Pickles, A., Anderson, D. K., & Lord, C. (2014). Heterogeneity and plasticity in the development of language: A 17 year follow up of children referred early for possible autism. *Journal of Child Psychology and Psychiatry, 55(12)*, 1354–1362. doi: 10.1111/jcpp.12269.

Pierce, K., Conant, D., Hazin, R., Stoner, R., & Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. *Archives of General Psychiatry, 68*(1), 101–109. doi: 10.1001/ archgenpsychiatry.2010.113.

Pierce, K., Marinero, S., Hazin, R., McKenna, B., Barnes, C. C., & Malige, A. (2016). Eye-tracking reveals abnormal visual preference for geometric images as an early biomarker of an ASD subtype associated with increased symptom severity. *Biological Psychiatry, 79*, 657–666. doi: 10.1016/ j.biopsych.2015.03.032.

Pinker, S. (1989). *Learnability and Cognition*, The MIT Press, Cambridge, MA.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition, 95,* 201–236 doi: 10.1016/j.cognition.2004.08.004

Plunkett, K. (1997). Theories of early language acquisition. *Trends in Cognitive Sciences*, *1*(4), 146-153.

Polimeni, M.A., Richdale, A., & Francis, A. (2005). A survey of sleep problems in autism, Asperger's disorder and typically developing children. *Journal of Intellectual Disability Research, 49*(4):260–8. doi: 10.1111/j.1365-2788.2005.00642.x

Pomper, R., & Saffran, J. R. (2018). Familiar object salience affects novel word learning. *Child Development*, *90*(2), e246-e262. doi: 10.1111/cdev.13053

Posner, M.I. (1988). Structures and functions of selective attention. In T. Boll & B. Bryant (Eds.), *Clinical neuropsychology and brain function: Research, measurement and practice* (pp. 169–202). Washington, DC: American Psychological Association.

Posner, M. I. & Dehaene, R. S. (1994). Attentional networks. *Trends in Neuroscience, 17*, 75–79. doi: 10.1016/0166-2236(94)90078-7

Posner, M. I. & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, *12*, 3, 427–441. doi: 10.1017/s0954579400003096

Potrzeba, E. R., Fein, D., & Naigles, L. (2015). Investigating the shape bias in typically developing children and children with autism spectrum disorders. *Frontiers in Psychology*, *6*, 446. doi: 10.3389/fpsyg.2015.00446

Preissler, M., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition, 97*(1), B13-B23. doi: 10.1016/j.cognition.2005.01.008

Prizant, B. M. (1983). Language acquisition and communicative behavior in autism: Toward an understanding of the ‘‘whole’’ of it. *Journal of Speech and Hearing Disorder, 48,* 296–307. doi: 10.1044/jshd.4803.296

Prothmann, A., Ettrich, C., & Prothmann, S. (2009). Preference for, and Responsiveness to, People, Dogs and Objects in Children with Autism. *Anthrozoös, 22*(2), 161-171. doi: 10.2752/175303709X434185

Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Development, 77*(2), 266-280. doi: 10.1111/j.1467-8624.2006.00869.x

Pupura, D. J. & Lonigan, C. J. (2009). Conners' Teacher Rating Scale for Preschool Children: A Revised, Brief, Age-Specific Measure. *Journal of Clinical Child and Adolescent Psychology, 38*(2), 263–272. doi: 10.1080/15374410802698446.

Quine, W. V. O. (1960). *Word and object: An inquiry into the linguistic mechanisms of objective reference.* Cambridge, MA: MIT Press. doi: 10.1068/p6639.

Ray-Subramanian, C. E., & Weismer, S. E. (2012). Receptive and expressive language as predictors of restricted and repetitive behaviors in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42,* 2113–2120. doi: 10.1007/s10803-012-1463-6

Reed, S. R., Stahmer, A. C., Suhrheinrich, J., & Schreibman, L. (2013). Stimulus overselectivity in typical development: Implications for teaching children with autism. *Journal of Autism and Developmental Disorders*, *43*, 1249-1257. doi: 10.1007/s10803-012-1658-x

Richler, J., Bishop, S.L., & Kleinke, J. R., & Lord, C. (2007). Restricted and repetitive behaviors in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37,* 73-85. doi: 10.1007/s10803-006-0332-6

Ricketts, J., Dockrell, J. E., Patel, N., Charman, T., & Lindsay, G. (2015). Do children with specific language impairment and autism spectrum disorders benefit from the presence of orthography when learning new spoken words? *Journal of Experimental Child Psychology, 134*, 43–61. doi: 10.1016/j.jecp.2015.01.015

Rincover, A., & Ducharme, J. M. (1987). Variables influencing stimulus overselectivity and 'tunnel vision' in developmentally delayed children. *American Journal of Mental Deficiency, 91*, 422–430.

Robertson, E. M. (2009). From creation to consolidation: a novel framework for memory processing. *PLoS Biology, 7*(1), e1000019. doi: 10.1371/journal.pbio.1000019

Robinson, E. B., Lichtenstein, P., Anckarsäter, H., Happé, F., & Ronald, A. (2013). Examining and interpreting the female protective effect against autistic behavior. *Proceedings of the National Academy of Sciences of the United States of America, 110*(13), 5258–5262. doi: 10.1073/pnas.1211070110

Robinson, S., Goddard, L., Dritschel, B., Wisley, M., & Howlin, P. (2009). Executive functions in children with autism spectrum disorders. *Brain and Cognition*, *71*, 362-368. doi: 10.1016/j.bandc.2009.06.007

Roder, B. J., Bushnell, E. W., Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy 1*(4), 491–507. doi: 10.1207/S15327078IN0104_9

Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology, 37,* 8–38. doi: 10.1080/15374410701817808

Roid, G. H., Miller, L. J., Pomplun, M., & Koch, C. (2013). *Leiter International Performance Scale* (3rd ed.). Los Angeles, CA: Western Psychological Services.

Rose, V. Trembath, D., Keen, D. & Paynter. J. (2016). The proportion of minimally verbal children with autism spectrum disorder in a community-based early intervention programme. *Journal of Intellectual Disability Research, 60*(5), 464-477. doi: 10.1111/jir.12284

Roser, M., Aslin, R., McKenzie, R., Zahra, D., & Fiser, J. (2015). Enhanced Visual Statistical Learning in Adults with Autism. *Neuropsychology, 29*(2), 163-172. doi: 10.1037/neu0000137

Rowland, C. (2014). *Understanding child language acquisition.* Abingdon, UK: Routledge.

Sacrey L. A. R., Armstrong V. L., Bryson S. E., & Zwaigenbaum, L. (2014). Impairments to visual disengagement in autism spectrum disorder: A review of experimental studies from infancy to adulthood. *Neuroscience & Biobehavioral Reviews, 47*, 559–577. doi: 10.1016/j.neubiorev. 2014.10.011.

Sacrey, L. A. R., Bryson, S. E., & Zwaigenbaum, L. (2013). Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder: A longitudinal study from 6 to 36 months of age. *Behavioural brain research*, *256*, 441-450. doi: 10.1016/j.bbr.2013.08.028

Sáez, L., Folsom, J. S., Al Otaiba, S., & Schatschneider, C. (2012). Relations among student attention behaviors, teacher practices, and beginning word reading skill. *Journal of Learning Disabilities, 45*(5), 418-432. doi: 10.1177/0022219411431243

Samuelson, L. K., & McMurray, B. (2017). What does it take to learn a word? *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(1-2), e1421. doi: 10.1002/wcs.1421

Samuelson, L., Kucker, S., & Spencer, J. (2017). Moving Word Learning to a Novel Space: A Dynamic Systems View of Referent Selection and Retention. *Cognitive Science, 41*(S1), 52-72. doi: 10.1111/cogs.12369

Samuelson, L., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category organization and syntax correspond? *Cognition, 73*(1), 1–33. doi: 10.1016/S0010-0277(99)00034-7

Samuelson, L. K., & Smith, L. B. (2000). Grounding development in cognitive processes. *Child Development, 71*(1), 98–106. doi: 10.1111/1467-8624.00123

Sasson, N. J., Elison, J. T., Turner-Brown, L. M., Dichter, G. S., & Bodfish, J. W. (2011). Brief report: Circumscribed attention in young children with autism. *Journal of Autism and Developmental Disorders, 41*, 242–247. doi: 10.1007/s10803-010-1038-3

Sasson, N. J., Turner-Brown, L. M., Holtzclaw, T. N., Lam, K. S. L., & Bodfish, J. W. (2008). Children with autism demonstrate circumscribed attention during passive viewing of complex social and nonsocial picture arrays. *Autism Research, 1*, 31–42, doi: 10.1002/aur.4

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, *69*(2), 309-320. doi: 10.1111/j.1467-8624.1998.tb06190.x

Schmidt, C., Collette, F., Cajochen, C., & Peigneux, P. (2007). A time to think: circadian rhythms in human cognition. *Cognitive Neuropsychology, 24*(7), 755-789. doi: 10.1080/02643290701754158

Schopler, E., Van Bourgondien, M.E., Wellman, G.J., & Love, S.R. (2010). *Childhood autism rating scale CARS-2* (2nd ed.), Western Psychological Services, Los Angeles.

Schover, L. R., & Newsom, C. D. (1976). Overselectivity, developmental level, and overtraining in autistic and normal children. *Journal of Abnormal Child Psychology, 4*(3), 289–298. doi: 10.1007/BF00917765

Schwichtenberg, A. J., Iosif, A-M., Goodlin-Jones, B., Tang, K., & Anders, T. (2011). Daytime Sleep Patterns in Preschool Children with Autism, Developmental Delay, and Typical Development. *American Journal on Intellectual and Developmental Disabilities, 116* (2), 142–152. doi: 10.1352/1944-7558-116.2.142

Seegers, V., Touchette, E., Dionne, G., Petit, D., Seguin, J. R., Montplaisir, J., Vitaro, F., Falissard, B., Boivin, M., & Tremblay, R. E. (2016). Short persistent sleep duration is associated with poor receptive vocabulary performance in middle childhood. *Journal of Sleep Research, 25*(3), 325–332. doi: 10.1111/jsr.12375

Sigman, M., & McGovern, C. W. (2005). Improvement in cognitive and language skills from preschool to adolescence in autism. *Journal of Autism and Developmental Disorders, 35*(1), 15–23. doi: 10.1007/s10803-004-1027-5

Skuse, D. H. (2000). Imprinting, the X-chromosome, and the male brain: Explaining sex differences in the liability to autism. *Pediatric Research, 47*, 9. doi: 10.1203/00006450-200001000-00006

Smith, L.B. (2000). *Becoming a word learner: A debate on lexical development.* Oxford University Press; New York: Learning how to learn words: An associative crane; p. 51-80.

Smith, T. & Eikeseth, S. (2011). O. Ivar Lovaas: pioneer of applied behavior analysis and intervention for children with autism. *Journal of Autism and Developmental Disorders, 41*, 375–378. doi: 10.1007/s10803-010-1162-0

Smith, S. M., & Handy, J. D. (2014). Effects of varied and constant environmental contexts on acquisition and retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(6), 1582–1593. doi: 10.1037/xlm0000019

Smith, L., Jones, S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition, 60*(2), 143-171. doi: 10.1016/0010-0277(96)00709-3

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13*, 13–19. doi: 10.1111/1467-9280.00403.

Smith, K., Smith, A. D. M., & Blythe, R. A. (2011). Cross-situational learning: An experimental study of word-learning mechanisms. *Cognitive Science, 35,* 480–498. doi: 10.1111/j.1551-6709.2010.01158.x

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. doi: 10.1016/j.cognition.2007.06.010

Smith, L., & Yu, C. (2013). Visual Attention Is Not Enough: Individual Differences in Statistical Word-Referent Learning in Infants*, Language Learning and Development, 9*(1), 25-49. doi: 10.1080/15475441.2012.707104

So Yeon, K., Rispoli, M., Lory, C., Gregori, E., & Brodhead, M. T. (2018). The Effects of a Shared Reading Intervention on Narrative Story Comprehension and Task Engagement of Students with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 48*(10), 3608–3622. doi: 10.1007/s10803-018-3633-7

Solomon, M., Olsen, E., Niendam, T., Ragland, J., Yoon, J., & Minzenberg, M. (2011). From lumping to splitting and back again: atypical social and language development in individuals with clinical-high-risk for psychosis, first episode schizophrenia, and autism spectrum disorders. *Schizophrenia Research, 131,* 146-151. doi: 10.1016/j.schres.2011.03.005

Souders, M. C., Mason, T. B., Valladares, O., Bucan, M., Levy, S. E., Mandell, D. S., Weaver, T. E., & Pinto-Martin, J. (2009). Sleep behaviors and sleep quality in children with autism spectrum disorders. *Sleep, 32*(12), 1566-1578. doi: 10.1093/sleep/32.12.1566

Southall, C. (2013). Use of technology to accommodate differences associated with autism spectrum disorder in the general curriculum and environment. *Journal of Special Education Technology*, *28*, 23–34. doi: 10.1177/016264341302800103

Spiegel, C., & Halberda, J. (2011). Rapid Fast-Mapping Abilities in 2-Year-Olds. *Journal of Experimental Child Psychology, 109*(1), 132-140. doi: 10.1016/j.jecp.2010.10.013

Stone, W. L., & Yoder, P. J. (2001). Predicting spoken language level in children with autism spectrum disorders. *Autism, 5*(4), 341-361. doi: 10.1177/1362361301005004002

Studdert-Kennedy, M. (1986). *Some developments in research on language behavior.* N.J. Smelser, D.R. Gerstein (Eds.), Behavioral and social science: Fifty years of discovery: In commemoration of the fiftieth anniversary of the "Ogburn report: Recent social trends in the United States", National Academy Press, Washington, DC, pp. 208-248.

Swensen, L. D., Kelley, E., Fein, D., & Naigles, L. R. (2007). Processes of Language Acquisition in Children with Autism: Evidence from Preferential Looking. *Child Development, 78*(2), 542–557. doi: 10.1111/j.1467-8624.2007.01022.x

Swettenham, J., Baron-Cohen, S., Charman, T., Cox, A., Baird, G., Drew, A., Rees, L., & Wheelwright, S. (1998). The Frequency and Distribution of Spontaneous Attention Shifts between Social and Nonsocial Stimuli in Autistic, Typically Developing, and Nonautistic Developmentally Delayed Infants. *Journal of Child Psychology and Psychiatry, 39*(5), 747–753. doi: 10.1017/S0021963098002595

Swingley, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of memory and language*, *60*(2), 252-269. doi: 10.1016/j.jml.2008.11.003

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, *71*(2), 73-108. doi: 10.1016/S0010-0277(99)00021-9

Tager-Flusberg, H., Joseph, R., & Folstein, S. (2001). Current directions in research on autism. *Mental retardation and developmental disabilities research reviews, 7*(1), 21-29. doi: 10.1002/1098-2779(200102)7:1<21::AID-MRDD1004>3.0.CO;2-3

Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Research, 6*, 468–478. doi: 10.1002/aur.1329.

Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and Communication in Autism. *In Handbook of Autism and Pervasive Developmental Disorders: Diagnosis, Development, Neurobiology, and Behavior*; Volkmar, F. R., Paul, R., Klin, A., Cohen, D., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA; pp. 335–364.

Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., ... & Yoder, P. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research, 52*(3), 643 – 652. doi: 10.1044/1092-4388(2009/08-0136)

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *Journal of Neuroscience, 30*(43), 14356–14360. doi: 10.1523/JNEUROSCI.3028-10.2010

Tek, S., Jaffery, G., Fein, D., & Naigles, L. R. (2008). Do children with autism spectrum disorders show a shape bias in word learning?. *Autism Research*, *1*(4), 208-222. doi: 10.1002/aur.38

Tek, S., Mesite, L., Fein, D., & Naigles, L. (2014). Longitudinal analyses of expressive language development reveal two distinct language profiles among young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44,* 75–89. doi: 10.1007/s10803-013-1853-4

Tenenbaum, E. J., Amso, D., Abar, B., & Sheinkopf, S. J. (2014). Attention and word learning in autistic, language delayed and typically developing children. *Frontiers in Psychology, 5,* 490–490. doi: 10.3389/fpsyg.2014.00490

Tenenbaum, E. J., Amso, D., Righi, G., & Sheinkopf, S. J. (2017). Attempting to "Increase Intake from the Input": Attention and Word Learning in Children with Autism. *Journal of Autism and Developmental Disorders, 47*, 1791–1805. doi: 10.1007/s10803-017-3098-0

Thurm, A., Lord, C., Lee, L.-C., & Newschaffer, C. (2007). Predictors of language acquisition in preschool children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37(9)*, 1721–1734. doi: 10.1007/s10803-006-0300-1.

Tilley, A., & Warren, P. (1983). Retrieval from semantic memory at different times of day. *Journal of Experimental Child Psychology. Learning, Memory, and Cognition, 9,* 718–724. doi: 10.1037//0278-7393.9.4.718

Tomasello, M. (2000). *Perceiving intentions and learning words in the second year of life*, in: M. Bowerman, S. Levinson (Eds.), Language Acquisition and Conceptual Development, Cambridge University Press, Cambridge, pp. 111–128.

Tomasello, M. (2003). The key is social cognition. *Language in mind: Advances in the study of language and thought*, 47-57.

Touchette, É., Petit, D., Séguin, J. R., Boivin, M., Tremblay, R. E., & Montplaisir, J. Y. (2007). Associations between sleep duration patterns and behavioral/cognitive functioning at school entry. *Sleep, 30*(9), 1213-1219. doi: 10.1093/sleep/30.9.1213

Tse, A. C. Y., Yu, C. C. W., & Lee, P. H. (2020). Comparing sleep patterns between children with autism spectrum disorder and children with typical development: A matched case–control study. *Autism: The International Journal of Research and Practice, 24*(8), 2298–2303. doi: 10.1177/1362361320936827

Turner, L., Stone, W., Pozdol, S.L., & Coonrod, E. (2006). Follow-up of children with autism spectrum disorders fromage 2 to age 9. *Autism, 10*, 243–265. doi: 10.1177/1362361306063296

Twomey, K. E., Ma, L., & Westermann, G. (2018). All the right noises: Background variability helps early word learning. *Cognitive Science*, *42*, 413-438. doi: 10.1111/cogs.12539

Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's more like it: Multiple exemplars facilitate word learning. *Infant and Child Development*, *23*(2), 105-122. doi: 10.1002/icd.1824

Twomey, K. E., Morse, A. F., Cangelosi, A., & Horst, J. S. (2016) Children's referent selection and word learning: insights from a developmental robotic system. *Interaction Studies, 17*(1), 101-127. doi: 10.1075/is.17.1.05two

Uljarević, M., Alvares, G. A., Steele, M., Edwards, J., Frazier, T. W., Hardan, A. Y., & Whitehouse, A. J. (2022). Toward better characterization of restricted and unusual interests in youth with autism. *Autism: The International Journal of Research and Practice, 26*(5), 1296–1304. doi: 10.1177/13623613211056720

Van Der Geest, J. N., Kemner, C., Verbaten, M. N., & Van Engeland, H. (2002). Gaze behavior of children with pervasive developmental disorder toward human faces: a fixation time study. *Journal of Child Psychology and Psychiatry, 43*(5), 669–678. doi: 10.1111/1469-7610.00055

Venker, C. E. (2019). Cross-situational and ostensive word learning in children with and without autism spectrum disorder. *Cognition, 183*, 181–191. doi: 10.1016/j.cognition.2018.10.025.

Venker, C. E, Bean, A., & Kover, S. T. (2018). Auditory-visual misalignment: A theoretical perspective on vocabulary delays in children with ASD. *Autism Research*, 11(12), 1621-1628. doi: 10.1002/aur.2038

Venker, C., Kover, S., & Weismer, S. (2015). Brief Report: Fast Mapping Predicts Differences in Concurrent and Later Language Abilities Among Children with ASD. *Journal of Autism and Developmental Disorders, 46*(3), 1118-1123. doi: 10.1007/s10803-015-2644-x

Venker, C. E., Mathée, J., Neumann, D., Edwards, J., Saffran, J., & Ellis Weismer, S. (2021). Competing Perceptual Salience in a Visual Word Recognition Task Differentially Affects Children With and Without Autism Spectrum Disorder. *Autism Research. 14*(6), 1147-1162. doi: 10.1002/aur.2457

Venker, C., Neumann, D., & Aladé, F. (2022). Visual perceptual salience and novel referent selection in children with and without autism spectrum disorder. *Autism & Developmental Language Impairments, 7,* 239694152210854–23969415221085476. doi: 10.1177/23969415221085476

Vivanti, G., Hocking, D. R., Fanning, P., & Dissanayake, C. (2016). Verbal labels increase the salience of novel objects for preschoolers with typical development and Williams syndrome, but not in autism. *Journal of Neurodevelopmental Disorders, 8*(1), 46–46. doi: 10.1186/s11689-016-9180-7

Vlach, H. A., & DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 45(4), 700–711. doi: 10.1037/xlm0000611.

Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time?: Memory processes support children's retention of learned words. *Frontiers in Psychology, 3*, 1–8. doi: 10.3389/fpsyg.2012.00046

Vlach, H. A., & Sandhofer, C. M. (2014). Retrieval Dynamics and Retention in Cross-Situational Statistical Word Learning. *Cognitive Science, 38*(4), 757–774. doi: 10.1111/cogs.12092

Volden, J., Smith, I. M., Szatmari, P., Bryson, S., Fombonne, E., Mirenda, P., ... & Thompson, A. (2011). Using the preschool language scale, fourth edition to characterize language in preschoolers with autism spectrum disorders. *American Journal of Speech-Language Pathology, 20*(3), 200–208. doi: 10.1044/1058-0360(2011/10-0035)

Volkmar, F.R. (2014). Editorial: the importance of early intervention. *Journal of Autism and Developmental Disorders, 44*(12), 2979–2980. doi: 10.1007/s10803-014-2265-9

Volkmar, F., Chawarska, K., & Klin, A. (2005). Autism in infancy and early childhood. *Annual Review of Psychology, 56*, 315-336. doi: 10.1146/annurev.psych.56.091103.070159

Wainwright, B., Allen, M. L., & Cain, K. (2020). Symbolic Understanding and Word–Picture–Referent Mapping from iPads in Autism Spectrum Condition: The Roles of Iconicity and Engagement. *Journal of Autism and Developmental Disorders*, *50*(8), 2941–2956. doi: 10.1007/s10803-020-04404-8

Walker, S., Gaskell, M. G., Knowland, V. C. P., Fletcher, F. E., Cairney, S. A., & Henderson, L. M. (2020). Growing up with interfering neighbours: The influence of time of learning and vocabulary knowledge on written word learning in children. *Royal Society Open Science, 7*(3), 191597. doi: 10.1098/rsos.191597

Walton, & Ingersoll, B. R. (2013). Expressive and receptive fast-mapping in children with autism spectrum disorders and typical development: The influence of orienting cues. *Research in Autism Spectrum Disorders, 7*(6), 687–698. doi: 10.1016/j.rasd.2013.02.012

Watt, N., Wetherby, A., & Shumway, S. (2006). Prelinguistic Predictors of Language Outcome at 3 Years of Age. *Journal of Speech, Language, and Hearing Research, 49*(6), 1224–1237. doi: 10.1044/1092-4388(2006/088)

Waxman, S.R., & Kosowski, T.D. (1990) Nouns mark category relations: Toddlers' and preschoolers' word learning biases. *Child Development, 61*(5):1461–73. doi: 10.1111/j.1467-8624.1990.tb02875.x

Weismer, S. E., Lord, C., & Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *Journal of Autism and Developmental Disorders, 40,* 1259-1273. doi: 10.1007/s10803-010-0983-1

Wilhelm, H., & Lovaas, O. I. (1976). Stimulus overselectivity: a common feature in autism and mental retardation. *American Journal of Mental Deficiency*, *81*(1), 26-31.

Wilkinson, K. M., Ross, E., & Diamond, A. (2003). Fast mapping of multiple words: Insights into when “the information provided” does and does not equal “the information perceived.” *Applied Developmental Psychology, 24,* 739-762. doi: 10.1016/j.appdev.2003.09.006

Williams, K.T. (2007). *Expressive Vocabulary Test.* 2nd Edn. Minneapolis, MN: Pearson Assessments.

Williams, S. E., & Horst, J. S. (2014). Goodnight book: Sleep consolidation improves word-learning via storybooks. *Frontiers in Psychology, 5*, 184. doi: 10.3389/fpsyg.2014.00184

Woodward, A. (2003). *Infants' use of action knowledge to get a grasp on words*. MIT Press.

Yang, D.C., & Li, M. N. (2013). Assessment of animated self-directed learning activities modules for children's number sense development. *Journal of Educational Technology & Society*, *16*(3), 44-58.

Yu, C., & Ballard., D. H. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing, 70*, 2149-2165. doi: 10.1016/j.neucom.2006.01.034

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414-420. doi: 10.1111/j.1467-9280.2007.01915.x

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition, 125*, 244–262. doi: 10.1016/j.cognition.2012.06.016

Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition, 145*, 53-62, doi: 10.1016/j.cognition.2015.07.013

Yurovsky, D., Smith, L. B. & Yu, C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental Science, 16*(6), 959-966. doi: 10.1111/desc.12036

Yurovsky, D., Yu, C., & Smith, L. B. (2013). Competitive processes in cross-situational word learning. *Cognitive Science*, *37*(5), 891-921. doi: 10.1111/cogs.12035

Zosh, J. M., Brinster, M., & Halberda, J. (2013). Optimal contrast: Competition between two referents improves word learning. *Applied Developmental Science, 17*(1), 20–28. doi: 10.1080/10888691.2013.748420

Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research, 50*(6), 1562-1592. doi: 10.1044/1092-4388(2007/106)

Zwaigenbaum L., Bauman M. L., Choueiri R., Kasari C., Carter A., Granpeesheh D., … Natowicz, M. R. (2015). Early intervention for children with autism spectrum disorder under 3 years of age: Recommendations for practice and research. *Pediatrics*, *136*(1), S60–S81. doi: 10.1542/peds.2014-3667E

Appendix A: Animal Interests Questionnaire

Interests Questionnaire

Dear Parent/Caregiver,

As mentioned in the information sheet, the objective of this research is to explore how children's interests in certain things influence their ability to learn new words. Animals are a common fascination of many young children and we expect that they may pay more attention to them than other things.

Please answer the following questions, as accurately as possible, in relation to your child's current interests.

In the questions, "realistic animals" refer to animals that are similar to those seen in real life, for example TV shows like Meerkat Manor, Planet Earth, Waffle the Wonder Dog.

"Human-like animals" are animals that are made to look or behave like humans, such as wearing clothes, performing jobs, talking, and exhibiting human emotions. Examples include programmes like Tom and Jerry, The Lion King, Paw Patrol.

Please circle the most appropriate answer.

1. **Does your child like animals?**

Yes No

If yes, please answer the following questions by circling the most appropriate answer:

2. **How much does your child like animals?**

1	2	3	4
They don't mind animals	They like animals a little	They like animals a lot	They really, really like animals

3. How much does your child enjoy listening to stories about realistic animals? (e.g. books about farm animals)

1	2	3	4
They don't particularly enjoy it	They enjoy it a little	They enjoy it a lot	They really, really enjoy it

4. How much does your child enjoy listening to stories about human-like animals? (e.g. The Lion King)

1	2	3	4
They don't particularly enjoy it	They enjoy it a little	They enjoy it a lot	They really, really enjoy it

5. How much does your child enjoy playing with realistic animal toys? (e.g. toy zoo animals)

1	2	3	4
They don't particularly enjoy it	They enjoy it a little	They enjoy it a lot	They really, really enjoy it

6. How much does your child enjoy playing with human-like animal toys? (e.g. Peppa Pig)

1	2	3	4
They don't particularly enjoy it	They enjoy it a little	They enjoy it a lot	They really, really enjoy it

7. **How much does your child enjoy watching television programmes, videos, and films involving realistic animals? (e.g. Roar)**

1	2	3	4
They don't particularly enjoy it	They enjoy it a little	They enjoy it a lot	They really, really enjoy it

8. **How much does your child enjoy watching television programmes, videos and films involving human-like animals? (e.g. Secret Life of Pets)**

1	2	3	4
They don't particularly enjoy it	They enjoy it a little	They enjoy it a lot	They really, really enjoy it

9. **How often does your child enjoy interacting with real animals? (e.g. pets or farmyard animals?**

1	2	3	4
They don't particularly enjoy it	They enjoy it a little	They enjoy it a lot	They really, really enjoy it

10. **Does your child usually enjoy activities that involve animals more than activities that do not?**

Yes No

Appendix B: Supplementary materials for analyses in Chapter 2

Is autistic children's word learning facilitated or hindered by high interest distractors?

Model building sequences for results reported in the main text

All models were conducted using the glmer and lmer functions from the lme4 package in R (Bates et al., 2015). Population was contrast coded as -0.5 (neurotypical) and 0.5 (autistic). Condition was contrast coded as -0.5 (novel object) and 0.5 (novel animal). Trial type was coded as -0.5 (known) and 0.5 (novel). Receptive vocabulary was coded as the participant's age equivalent based on their raw score on the British Picture Vocabulary Scale 2 (BPVS; Dunn et al., 1997). Expressive Vocabulary was coded as the participant's age equivalent based on their raw score on the Expressive Vocabulary Test 2 (EVT; Williams, 2007) or the expressive language module of the Mullen's Scales of Early Learning (Mullen, 1995). Attention was coded as the participant's raw score on the Conner's Teacher Rating Scale 15 (Pupura & Lonigan, 2009). Autism severity was coded as the participant's raw score on the Childhood Autism Rating Scale 2 (CARS; Schopler et al., 2010). Non-verbal intellectual abilities were coded as children's raw score on the Leiter-3 test of non-verbal intelligence (NVIQ; Roid et al., 2013). Repetitive behaviour was coded as children's raw score on the Repetitive Behaviour Questionnaire (RRB; Leekam et al., 2007). Animal Interest was coded as children's raw score on our Animal Interest questionnaire. Chronological Age was measured in months. Referent selection accuracy was coded as -0.5 (incorrect) and 0.5 (correct). Total accuracy at referent selection for novel trials was coded as 0-4. Number of labels at referent selection per novel word was coded as 1-6. 5-minute retention accuracy was coded as 0-8. Trial-level accuracy for analyses at each experimental phase was coded as 1(correct) or 0 (incorrect).

All analyses were undertaken using the same procedure. Models were built up sequentially, adding in one fixed effect at a time and comparing each model with the previous best-fitting model using log-likelihood tests. Each model was built up from a null model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant (referent selection analyses) or Condition slopes per participant (5 minute and 24-hour retention analyses). If some models in a sequence were singular fitting or failed to converge, the random effects were simplified until all models in the sequence successfully converged (the final random effects structure for each variable are provided below).

Referent Selection

Accuracy

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. Fixed effects of Population (Model 2), Condition (Model 3), and Trial Type (Model 4) were entered individually. The addition of Condition ($\chi^2 = 3.41$, $p = .065$) yielded a borderline improvement in fit than the baseline model. The addition of Trial Type ($\chi^2 = 16.39$, $p < .001$) yielded a significant improvement in fit when compared with the baseline model. The addition of Population ($p = .40$) did not significantly improve fit. Model 5 included fixed effects of Population + Trial Type, which was significantly better fitting than Model 3 ($\chi^2 = 15.21$, $p <.001$) but not Model 4 ($p = .14$). Model 6 contained fixed effects of Condition + Trial Type, yielding a significantly better fit than Model 3 ($\chi^2 = 13.08$, $p <.001$), but not Model 4 ($p = .76$). Model 7 included fixed effects of Trial Type + Population + Condition, yielding a significantly better fit than Model 3 ($\chi^2 = 16.66$, $p <.001$), but not Model 4 ($p = .16$). Model 8 included the Trial Type x Population interaction, yielding a significantly better fit than Model 3 ($\chi^2 = 16.03$, $p <.001$), but not Model 4 ($p = .22$). Model 9 included the Trial Type x Condition interaction, yielding a significantly better fit than Model 3 ($\chi^2 = 19.20$, $p <.001$), and Model 4 ($\chi^2 = 6.21$, $p = .045$), and containing a significant interaction effect ($z = -4.89$, $p <.001$). Model 10 included the Population x Condition interaction but did not significantly improve fit compared to Model 3 ($p = .66$) or Model 4 ($p = 1.00$). Finally, the three-way interaction was entered (Model 11), and was borderline significantly better fitting than Model 9 ($\chi^2 = 9.01$, $p = .061$), but did not contain a significant three-way interaction effect ($z = 1.42$, $p = .15$). Thus, Model 9, containing the Trial Type x Condition interaction, was the best fit to the observed data.

Accuracy – individual differences

Informed by the best fitting model described above, we began with a baseline model including the Condition x Trial Type interaction, and by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. These models analyse data from the autistic and neurotypical groups separately.

Autistic

Fixed effects of Age (Model 2), RRB (Model 3), Animal Interest (Model 4), Receptive Vocabulary (Model 5), and Expressive Vocabulary (Model 6), were entered individually. NVIQ, Attention, and CARS were not added as fixed effects for the autistic population as not all participants successfully completed these measures. The addition of Age ($p = .59$), RRB ($p = 1.00$), and Animal Interest ($p = .16$) did not improve fit compared to the baseline model. However, the individual addition of Receptive Vocabulary ($\chi^2 = 11.46$ $p <.001$), and Expressive Vocabulary ($\chi^2 = 12.07$, $p <.001$), yielded significant improvements in fit when compared to the baseline model. Therefore, Model 7 included fixed effects of Expressive Vocabulary + Receptive Vocabulary but did not differ significantly from Model 5 ($p= .18$) or Model 6 ($p = .28$). Model 6, with a fixed effect of Expressive Vocabulary, was concluded to be the overall best fitting model to the observed data as its AIC value of 171.3 and BIC value of 229.6 were lower than those of Model 5 (AIC = 171.9; BIC = 230.2). While the best fitting model was determined by the model with the lowest AIC and BIC values, individual effects of Receptive Vocabulary may have marginally contributed to children's referent selection accuracy and could overlap with Expressive Vocabulary.

Neurotypical

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Animal Interest (Model 6), Receptive Vocabulary (Model 7), Expressive Vocabulary (Model 8), and NVIQ (Model 9) were entered individually. The addition of Age ($p = 1.00$), Attention ($p = .87$), CARS ($p = .18$), RRB ($p = .11$), and Animal Interest ($p = .77$), did not improve fit compared to the baseline model. The individual addition of Receptive Vocabulary ($\chi^2 = 7.48$,

$p = .006$), Expressive Vocabulary ($\chi^2 = 6.10$, $p = .014$), and NVIQ ($\chi^2 = 10.08$, $p = .001$) yielded a significantly better fit than the baseline model. Therefore, Model 10 added fixed effects of Receptive Vocabulary + Expressive Vocabulary + NVIQ and was borderline significantly better fitting than Model 7 ($\chi^2 = 5.28$, $p = .07$), significantly better fitting than Model 8 ($\chi^2 = 6.66$, $p = .036$), but was not significantly better fitting than Model 9 ($p = .26$). Model 11 included fixed effects of Receptive Vocabulary + Expressive Vocabulary but did not differ significantly from Model 7 ($p = .73$), Model 8 ($p = .22$), and Model 9 ($p = 1.00$). Model 12 added fixed effects of Receptive Vocabulary + NVIQ and did not differ significantly from Model 7 ($p = .10$) or Model 9 ($p = .74$) but did differ significantly from Model 8 ($\chi^2 = 4.10$, $p = .043$). Model 13 added fixed effects of Expressive Vocabulary + NVIQ and did not differ significantly from Model 9 ($p = .14$) but did differ significantly from Model 7 ($\chi^2 = 4.77$, $p = .029$), and Model 8 ($\chi^2 = 6.15$, $p = .013$). Therefore, Model 9 – containing a fixed effect of NVIQ – was selected as the final model because it did not differ in fit when compared with more complex models.

Correct response times

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. Fixed effects of Population (Model 2), Condition (Model 3), and Trial Type (Model 4) were entered individually. The addition of Trial Type ($\chi^2 = 15.14$, $p <.001$) yielded a significant improvement in fit when compared with the baseline model. The addition of Population ($p = .74$) and Condition ($p = .97$) did not improve fit compared to the baseline model. Model 5 added fixed effects of Population alongside Trial Type, which did not improve fit compared to Model 4 ($p = .64$). Model 6 added fixed effects of Trial Type and Condition ($p = .95$), which did not improve fit compared to Model 4. Model 7 included all three fixed effects but did not differ significantly from Model 4 ($p = .90$). Model 8 included the Trial Type x Population interaction ($\chi^2 = 8.23$, $p = .016$), yielding a significant improvement in fit in comparison to Model 4. Model 9 included the Trial Type x Condition interaction ($p = .93$), and Model 10 included the Population x Condition interaction ($p = 1.00$), but neither model differed significantly from Model 4. Finally, the three-way interaction was entered (Model 11); doing so did not improve fit in comparison to Model 8 ($p = .61$). Thus, Model 8, containing the Trial Type x Population interaction, provided the best fit to the observed data.

5-Minute Retention

Accuracy

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Fixed effects of Population (Model 2), Condition (Model 3), Referent Selection Accuracy (Model 4), Novel Trial Referent Selection Accuracy (Model 5), and Number of Labels at Referent Selection (Model 6) were entered individually. The addition of Population ($p = .08$), Condition ($p = .68$), Referent Selection Accuracy ($p = .33$), Total Accuracy at Referent Selection for Novel Trials ($p = .94$), and Number of Labels at Referent Selection ($p = .19$), did not yield a significantly better fitting model compared to the baseline. Models 7-12 added the individual two-way interactions, however, none of these provided a better fit compared to the baseline model (Population x

Condition: $p = .23$; Population x Novel Trial Referent Selection Accuracy: $p = .19$; Population x Number of Labels at Referent Selection: $p = .16$; Condition x Referent Selection Accuracy: $p = .61$; Condition x Novel Trial Referent Selection Accuracy: $p = .78$; Condition x Number of Labels at Referent Selection: $p = .12$). Model 13 contained the Population x Referent Selection Accuracy interaction and approached a significant improvement in fit compared to the baseline model ($\chi^2 = 7.50$, $p =.057$). Model 14 added the three-way interaction of Population x Condition x Referent Selection Accuracy, which did not provide a better fit compared to Model 13 ($p = .52$). Model 15 added the three-way interaction of Population x Condition x Novel Trial Referent Selection Accuracy, which did not provide a better fit compared to Model 13 ($p = .72$). Model 16 added the three-way interaction of Population x Condition x Number of Labels at Referent Selection Referent Selection for Novel Trials, which did not provide a better fit compared to Model 13 ($p = .50$). Therefore, the best fitting model was Model 13, containing the Population x Referent Selection Accuracy interaction. However, the model comparison only approached significance, so we must interpret this result with caution.

Accuracy – individual differences

The following models were performed for autistic and neurotypical populations separately. For the autistic population, we began with a baseline model containing a fixed effect of Referent Selection Accuracy, and by-participant and by-word random intercepts, with a random slope of Condition per participant.

Autistic

Fixed effects of Age (Model 2), RRB (Model 3), Animal Interest (Model 4), Receptive Vocabulary (Model 5), and Expressive Vocabulary (Model 6) were entered individually. NVIQ, Attention, and CARS were not added as fixed effects for the autistic population as not all participants successfully completed this measure. The addition of Age ($p = .71$), RRB ($p = .08$), Receptive Vocabulary ($p = .48$) and Expressive Vocabulary ($p = .79$) did not improve fit compared to the baseline model. The inclusion of Animal Interest ($\chi^2 = 3.61$, $p = .058$) provided a borderline significantly better fit to the observed data than the baseline model. Therefore Model 4, containing a fixed effect of Animal Interest, was determined to be the best fit to the observed data. This result should be taken with caution however due to the approaching significant model comparison.

For the neurotypical model we began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant.

Neurotypical

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Animal Interest (Model 6), Receptive Vocabulary (Model 7), Expressive Vocabulary (Model 8), and NVIQ (Model 9), were entered individually. The addition of Attention ($p = .96$), CARS ($p = .31$), RRB ($p = .81$), and Animal Interest ($p = .92$) did not improve fit compared to the baseline model. The addition of Age ($\chi^2 = 10.52$, $p =.001$), Receptive Vocabulary ($\chi^2 = 8.90$, $p =.003$), Expressive Vocabulary ($\chi^2 = 14.44$, $p <.001$) and NVIQ ($\chi^2 = 7.34$, $p = .007$) significantly improved fit compared to the baseline model. Model 10 added fixed effects of Age + Receptive Vocabulary + Expressive Vocabulary + NVIQ. Model 10 did not differ significantly from Model 8 ($p = .09$) but was significantly better fitting than Model 2 ($\chi^2 =$

10.51, $p = .015$), Model 7 ($\chi^2 = 12.13$, $p = .007$) and Model 9 ($\chi^2 = 13.69$, $p = .003$). Model 11 included fixed effects of Expressive Vocabulary + Age ($p = .79$) but was not significantly better fitting than Model 8. Model 12 included fixed effects of Expressive Vocabulary + Receptive Vocabulary, which was significantly better fitting than Model 8 ($\chi^2 = 4.50$, $p = .034$). Model 13 included fixed effects of Expressive Vocabulary + NVIQ, which approached a significantly better fit than Model 8 ($\chi^2 = 4.02$, $p = .045$). Model 14 included fixed effects of Expressive Vocabulary + Receptive Vocabulary + NVIQ but was not significantly better fitting than Model 12 ($p = .18$) or Model 13 ($p = .13$). As such, Models 12 and 13 were retained as the current best fitting models. Model 12, with fixed effects of Expressive Vocabulary and Receptive Vocabulary, was concluded to be the overall best fitting model to the observed data as its AIC value of 390.4 and BIC value of 416.3 were lower than those of Model 13 (AIC = 390.9; BIC = 416.7). While the best fitting model was determined by the model with the lowest AIC and BIC values, individual effects of NVIQ may have marginally contributed to children's 5-minute retention accuracy and could overlap with Receptive Vocabulary.

Correct response times

We began with a baseline model containing a by-participant random intercept. Fixed effects of Population (Model 2) and Condition (Model 3) were entered individually. The addition of Population ($p = .86$) and Condition ($p = .34$) did not improve fit compared to the baseline model. Model 4 included fixed effects of Population + Condition but did not differ significantly from the baseline model ($p = .62$). Model 5 included the Population x Condition interaction but did not differ significantly from the baseline model ($p = .23$). Therefore, the inclusion of fixed effects did not improve predictive power.

24-Hour Retention

Accuracy

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Fixed effects of Population (Model 2), Condition (Model 3), Total Accuracy at 5-minute Retention (Model 4), Referent Selection Accuracy (Model 5), Total Accuracy at Referent Selection for Novel Trials (Model 6), and Number of Labels at Referent Selection (Model 7) were entered individually. The addition of Population ($p = .83$), Condition ($p = .33$), Referent Selection Accuracy ($p = .13$), Total Accuracy at Referent Selection for Novel Trials ($p = .27$), and Number of Labels at Referent Selection ($p = .41$) did not improve model fit. The addition of Total Accuracy at 5-minute Retention ($\chi^2 = 13.45$, $p <.001$) yielded a significant improvement in fit when compared with the baseline model. The addition of Population (Model 8, $p = .77$), Condition (Model 9, $p = .42$), Referent Selection Accuracy (Model 10, $p = .11$), Total Accuracy at Referent Selection for Novel Trials (Model 11, $p = .27$), and Number of Labels at Referent Selection (Model 12, $p = .39$) alongside Total Accuracy at 5-minute Retention did not yield any significant improvements in fit compared to the baseline model. Including a fixed effect of Total Accuracy at 5-minute Retention plus two-way interactions of Population x Referent Selection Accuracy interaction (Model 13; $p = .45$), Population x Total Accuracy at Referent Selection for Novel Trials (Model 14; $p = .71$), Population x Number of Labels at Referent Selection

(Model 15; p = .85), Condition x Referent Selection Accuracy (Model 16; p = .36), Condition x Total Accuracy at Referent Selection for Novel Trials (Model 17; p = .46), Condition x Number of Labels at Referent Selection (Model 18; p = .70) did not yield any significant improvements in fit compared to the baseline model. Including Total Accuracy at 5-minute Retention plus the Population x Condition interaction yielded an approaching significantly better fit than Model 4 (Model 19; χ^2 = 7.17, p = .067), and contained a significant Population x Condition interaction effect (z = -2.68, p = .008), and a significant fixed effect of Total Accuracy at 5-minute Retention (z = 3.59, p <.001). Model 20 contained the Population x Total Accuracy at 5-minute Retention interaction (p = .96), and Model 21 contained the Condition x Total Accuracy at 5-minute Retention interaction (p = .46), but neither were significantly better fitting than Model 4. Including a fixed effect of Total Accuracy at 5-minute Retention plus three-way interactions of Condition x Population x Referent Selection Accuracy in Model 22 did not yield a significantly better fit than Model 4 (p = .14) or Model 19 (p = .43). Inclusion of the Condition x Population x Total Accuracy at Referent Selection for Novel Trials interaction plus fixed effect of Total Accuracy at 5-minute Retention in Model 23 did not yield a significantly better fit than Model 4 (p = .19) or Model 19 (p = .59). For Model 24, inclusion of the Condition x Population x Number of Labels at Referent Selection interaction alongside the fixed effect of Total Accuracy at 5-minute Retention did not yield a significantly better fit than Model 4 (p = .21) or Model 19 (p = .66). Model 25 contained the Condition x Population x Total Accuracy at 5-minute Retention interaction but was not significantly better fitting than Model 4 (p = .27) or Model 19 (p = .93). Overall, Model 19 (Total Accuracy at 5-minute Retention + Population x Condition) was determined to be the final model as it best accounted for the data, although given that the model comparison only approached significance, we must interpret this result with caution.

Accuracy – individual differences

The following models analysed data from autistic and neurotypical populations separately. For the autistic population, informed by the best fitting model described above, we began with a baseline model including fixed effects of Condition and Total Accuracy at 5-minute Retention, plus by-participant and by-word random intercepts, with a random slope of Condition per participant.

Autistic

Fixed effects of Age (Model 2), RRB (Model 3), Animal Interest (Model 4), Receptive Vocabulary (Model 5), and Expressive Vocabulary (Model 6), were entered individually. NVIQ, Attention, and CARS were not added as a fixed effect for the autistic population as not all participants successfully completed these measures. The addition of Age (p = .15), RRB (p = .40), and Receptive Vocabulary (p = .10) did not significantly improve fit compared to the baseline model. The individual addition of Animal Interest (χ^2 = 6.69, p = .010) and Expressive Vocabulary (χ^2 = 4.35, p = .037) yielded a significant improvement in fit when compared with the baseline model. Model 7 included fixed effects of Expressive Vocabulary + Animal Interest and was not significantly better fitting than Model 4 (p = .44) or Model 6 (p = .09). Therefore, Models 4 and 6 were retained for comparison. Model 4, containing a fixed effect of Animal Interests, was determined to be the best fitting model, as it's AIC value of 356.2 and BIC value of 384.9 was lower than that of Model 6 (AIC = 358.5, BIC = 387.2).

For the neurotypical population, we began with a baseline model including a fixed effect of Total Accuracy at 5-minute Retention, and by-participant and by-word random intercepts, with a random slope of Condition per participant.

Neurotypical

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Animal Interest (Model 6), Receptive Vocabulary (Model 7), Expressive Vocabulary (Model 8), and NVIQ (Model 9) were entered individually. The addition of Age (p = .26), CARS (p = .47), RRB (p = .60), Animal Interest (p = .63), Receptive Vocabulary (p = .13), and Expressive Vocabulary (p = .08) did not significantly improve fit compared to the baseline model. The individual addition of Attention (χ^2 = 8.25, p = .004) and NVIQ (χ^2 = 4.16, p = .041) yielded significant improvements in fit compared to the baseline model. Model 10 added fixed effects of Attention + NVIQ and was borderline significantly better fitting than Model 3 (χ^2 = 3.65, p = .056), and significantly better fitting than Model 9 (χ^2 = 7.74, p = .005). Therefore, Model 10, containing fixed effects of Attention and NVIQ, was determined to be the best fitting model to the observed data. However, one of the model comparisons only approached significance, so we must interpret this result with caution.

Correct response times

We began with a baseline model containing a by-participant random intercept. Fixed effects of Population (Model 2) and Condition (Model 3) were entered individually. The addition of Population (p = .87) and Condition (p = .34) did not improve fit compared to the baseline model. Model 4 included fixed effects of Population + Condition but did not differ significantly from the baseline model (p = .33). Model 5 included the Population x Condition interaction but did not differ significantly from Model 2 (p = .38). Therefore, the inclusion of fixed effects did not improve predictive power.

Appendix C: Supplementary materials for analyses in Chapter 3

How do autistic and neurotypical children's interests influence their visual attention and accuracy during novel word learning?

Model building sequences for results reported in the main text

All models were conducted using the glmer and lmer functions from the lme4 package in R (Bates et al., 2015). Population was contrast coded as -0.5 (neurotypical) and 0.5 (autistic). Condition was contrast coded as -0.5 (novel object) and 0.5 (novel animal). Trial type was coded as -0.5 (known) and 0.5 (novel).

Receptive vocabulary was coded as the participant's age equivalent based on their raw score on the British Picture Vocabulary Scale 2 (BPVS; Dunn et al., 1997). Expressive Vocabulary was coded as the participant's age equivalent based on their raw score on the Expressive Vocabulary Test 2 (EVT; Williams, 2007) or the expressive language module of the Mullen's Scales of Early Learning (Mullen, 1995). Attention was coded as the participant's raw score on the Conner's Teacher Rating Scale 15 (CTRS-15; Pupura & Lonigan, 2009). Autism severity was coded as the participant's raw score on the Childhood Autism Rating Scale 2 (CARS; Schopler et al., 2010). Non-verbal intellectual abilities were coded as children's raw score on the Leiter-3 test of non-verbal intelligence (NVIQ; Roid et al., 2013). Repetitive behaviour was coded as children's raw score on the Repetitive Behaviour Questionnaire (RBQ; Leekam et al., 2007). Animal Interest was coded as children's raw score on our Animal Interest questionnaire. Chronological Age was measured in months. Referent selection accuracy was coded as -0.5 (incorrect) and 0.5 (correct). Total accuracy at referent selection for novel trials was coded as 0-4. Number of labels at referent selection per novel word was coded as 1-7. Total accuracy at 5-minute retention was coded as 0-8. Trial-level accuracy for analyses at each experimental phase was coded as 1 (correct) or 0 (incorrect).

Proportion of time spent looking at the target object on each trial was scored between 0 and 1 for all participants. Number of looks to the target object on each trial ranged from 0 to 14, with every new look towards the target stimuli being counted. The longest look to novel or known stimuli was coded as -0.5 (longest look to known object) or 0.5 (longest look to novel object) for the predicting accuracy analyses and known (0) novel (1) for the between group differences analyses. This variable was only included in referent selection analyses, as the retention phases involved only novel objects. The longest look to target or foil stimuli was coded as -0.5 (longest look to foil stimuli) and 0.5 (longest look to target stimuli) for the predicting accuracy analyses, and foil (0) target (1) for the between group differences analyses.

All analyses were undertaken using the same procedure. Models were built up sequentially, adding in one fixed effect at a time and comparing each model with the previous best-fitting model using log-likelihood tests. Each model was built up from a null model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant (referent selection analyses) or Condition slopes per participant (5 minute and 24-hour retention analyses). If some models in a sequence were singular fitting or failed to converge, the random effects were simplified until all models in the sequence successfully converged (the final random effects structure for each variables are provided below).

Do autistic and neurotypical children differ in accuracy and response times when learning names for high interest and neutral interest stimuli?

Referent Selection

Accuracy

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. Fixed effects of Population (Model 2), Condition (Model 3), and Trial Type (Model 4) were entered individually. The addition of Trial Type (χ^2 = 12.83, p <.001) yielded a significant improvement in fit when compared with the baseline model. The addition of Population (p = .10) and Condition (p = .45) did not significantly improve fit. Model 5 included fixed effects of Population + Trial Type but did not differ significantly from Model 4 (p = .18). Model 6 included fixed effects of Condition + Trial Type but did not differ significantly from Model 4 (p = .42). Model 7 included fixed effects of Trial Type + Population + Condition but did not differ significantly from Model 4 (p = .18). Models 8, 9 and 10 included the individual two-way interactions. None of these models differed significantly from Model 4 (Trial Type x Population: p = .30; Trial Type x Condition: p = .35; Population x Condition: p = 1.00). Finally, the three-way interaction was entered (Model 11); doing so did not significantly improve fit in comparison to Model 4 (p = .09). Thus, Model 4, containing a fixed effect of Trial Type, provided the best fit to the observed data.

Accuracy – individual differences

Informed by the best fitting model described above, we began with a baseline model including a fixed effect of Trial Type and by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. These models analyse data from the autistic and neurotypical groups separately.

Autistic

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Animal Interest (Model 6), Receptive Vocabulary (Model 7), and Expressive Vocabulary (Model 8), were entered individually. NVIQ was not added as a fixed effect for the autistic population as not all participants successfully completed this measure. The addition of Age (p = .53) or RRB (p = .31), did not improve fit compared to the baseline model. However, the individual addition of Attention (χ^2 = 6.34, p = .012), CARS (χ^2 = 10.64, p = .001), Animal Interest (χ^2 = 5.24, p = .022), Receptive Vocabulary (χ^2 = 10.17, p = .001), and Expressive Vocabulary (χ^2 = 12.71, p <.001), yielded significant improvements in fit when compared to the baseline model. Therefore, Model 9 included fixed effects of Expressive Vocabulary + Attention + CARS + Animal Interest + Receptive Vocabulary but did not differ significantly from Model 8 (p = .18) containing only Expressive Vocabulary, or Model 4 (p = .08) containing only CARS. Model 9 yielded a marginally significant improvement in fit compared to Model 7 (χ^2 = 8.81, p = .066), and differed significantly from Model 3 (χ^2 = 12.63, p = .013), and Model 6 (χ^2 = 13.73, p = .008). Model 10 included fixed effects of Expressive Vocabulary + CARS, differing significantly from Model 4 (χ^2 = 6.69, p = .010) and Model 8 (χ^2 = 4.62, p = .032). The individual addition of Attention (Model 11; p = .32), Receptive Vocabulary (Model 12; p = 1.00) and Animal Interests (Model 13; p = .67) alongside Expressive Vocabulary and CARS did not yield significant improvements in fit compared to Model 10. Model 14 included fixed effects of Expressive Vocabulary + CARS + Receptive Vocabulary + Animal Interests but was not significantly better fitting than Model 10 (p = .90). Model 15 included fixed effects of Expressive Vocabulary + CARS + Receptive Vocabulary + Attention but was not significantly better fitting than Model 10 (p = .73). Model 16 included fixed effects of Expressive Vocabulary + CARS + Animal Interest + Attention but was not significantly better fitting than Model 10 (p = .54). As such, Model 10, containing fixed effects of Trial Type + Expressive Vocabulary + CARS, provided the best fit to the observed data.

Neurotypical

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Animal Interest (Model 6), Receptive Vocabulary (Model 7), Expressive Vocabulary (Model 8), and NVIQ (Model 9) were entered individually. The addition of Attention (p = .44), CARS (p = .57), RRB (p = .92), and Animal Interest (p = .20) did not improve fit compared to the baseline model. However, the addition of Age (χ^2 = 4.51, p = .034), Expressive Vocabulary (χ^2 = 3.88, p = .049), and NVIQ (χ^2 = 5.03, p = .025) differed significantly from the baseline model. The addition of Receptive Vocabulary (χ^2 = 3.53, p = .060) approached a significant improvement in fit compared to the baseline model. Therefore, Model 10 added fixed effects of Age + Receptive Vocabulary + Expressive Vocabulary + NVIQ but did not yield a significant improvement in fit compared to Model 2 (p = .78), Model 7 (p = .55), Model 8 (p = .63) or Model 9 (p = .90). Model 11 included fixed effects of Age + Receptive Vocabulary + Expressive Vocabulary but did not differ significantly from Model 2 (p = .93), Model 7 (p = .57), Model 8 (p = .67) or Model 9 (p = 1.00). Model 12 included fixed effects of Age + Receptive Vocabulary + NVIQ but did not yield a significant improvement in fit than Model 2 (p = .68), Model 7 (p = .42), Model 8 (p = .50) or Model 9 (p = .88). Model 13 included fixed effects of Age + Expressive Vocabulary + NVIQ but did not differ significantly from Model 2 (p = .62), Model 7 (p = .38), Model 8 (p = .45) or Model 9 (p = .80). Model 14 included fixed effects of Receptive Vocabulary + Expressive Vocabulary + NVIQ but did not provide a significantly better fit than Model 2 (p = .68), Model 7 (p = .41), Model 8 (p = .49) or Model 9 (p = .88). Model 15 included fixed effects of Receptive Vocabulary + Age but did not differ significantly from Model 2 (p = .76), Model 7 (p = .30), Model 8 (p = .39) or Model 9 (p = 1.00). Model 16 included fixed effects of Expressive Vocabulary + Age but did not provide a significantly better fit than Model 2 (p = .83), Model 7 (p = .31), Model 8 (p = .41) or Model 9 (p = 1.00). Model 17 included fixed effects of NVIQ + Age but did not differ significantly from Model 2 (p = .43), Model 7 (p = .20), Model 8 (p = .26) or Model 9 (p = .74). Model 18 included fixed effects of Receptive Vocabulary + Expressive Vocabulary but did not yield a significant improvement in fit compared to Model 2 (p = 1.00), Model 7 (p = .55), Model 8 (p = .99) or Model 9 (p = 1.00). Model 19 included fixed effects of Receptive Vocabulary + NVIQ but did not differ significantly from Model 2 (p = .40), Model 7 (p = .19), Model 8 (p = .25) or Model 9 (p = .67). Model 20 included fixed effects of Expressive Vocabulary + NVIQ but did not differ significantly from Model 2 (p = .38), Model 7 (p = .18), Model 8 (p = .24) or Model 9 (p = .61). Thus, the four best fitting models were Model 2, Model 7, Model 8, and Model 9. We compared AIC and BIC values, which were as follows: Model 2 - AIC = 148.2, BIC = 197.9;

Model 7 - AIC = 149.2, BIC = 198.9; Model 8 - AIC = 148.9, BIC = 198.5; Model 9 - AIC = 147.7, BIC = 197.4). The final best fitting model with the lowest AIC and BIC values was Model 9 (Trial Type + NVIQ).

Correct response times

We began with a baseline model containing a by-participant random intercept. Fixed effects of Population (Model 2), Condition (Model 3), and Trial Type (Model 4) were entered individually. The individual addition of Population ($\chi^2 = 4.36, p = .037$) and Trial Type ($\chi^2 = 18.25, p <.001$) yielded a significant improvement in fit when compared with the baseline model. The addition of Condition ($p = .91$) did not significantly improve fit. Model 5 added fixed effects Population alongside Trial Type, with a significant improvement in fit compared to Model 3 ($\chi^2 = 18.39, p <.001$) and Model 4 ($\chi^2 = 4.50, p = .034$). Model 6 added fixed effects of Trial Type and Condition, yielding a significant improvement in fit compared to Model 3 ($\chi^2 = 13.90, p <.001$) but not Model 4 ($p = .92$). Model 7 included all three fixed effects but did not differ significantly from Model 5 ($p = .94$). Models 8, 9 and 10 included the individual two-way interactions. None of these models differed significantly from Model 5 (Trial Type x Population: $p = .66$; Trial Type x Condition: $p = 1.00$; Population x Condition: $p = 1.00$). Finally, the three-way interaction was entered (Model 11); doing so did not improve fit in comparison to Model 5 ($p = .10$). However, the individual effects of Trial Type ($t = 4.34, p <.001$) and Population ($t = 2.17, p = .038$), and the Population x Condition interaction ($t = 2.62, p = .009$) within Model 11 were significant. Model 12 included the Population x Condition interaction plus fixed effect of Trial Type and was significantly better fitting than Model 5 ($\chi^2 = 6.94, p = .031$), with a significant Population x Condition interaction ($t = 2.65, p = .008$) and significant fixed effects of Trial Type ($t = 4.40, p <.001$) and Population ($t = 2.19, p = .037$). Therefore, Model 12 containing a fixed effect of Trial Type, plus the Population x Condition interaction, provided the best fit to the observed data.

5-Minute Retention

We began with a baseline model a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Fixed effects of Population (Model 2), Condition (Model 3), Referent Selection Accuracy (Model 4), Total Accuracy at Referent Selection for Novel Trials (Model 5), and Number of Labels at Referent Selection (Model 6) were entered individually. The addition of Population ($p = .95$), Condition ($p = .76$), Referent Selection Accuracy ($p = .13$), Total Accuracy at Referent Selection for Novel Trials ($p = .33$), and Number of Labels at Referent Selection ($p = .90$), did not yield significantly a better fit compared to the baseline model. Models 7-13 added the individual two-way interactions. The inclusion of the following interactions did not yield a significantly better fit compared to the baseline model: Population x Referent Selection Accuracy: $p = .49$; Population x Total Accuracy at Referent Selection for Novel Trials: $p = .44$; Population x Number of Labels at Referent Selection: $p = 1.00$; Condition x Total Accuracy at Referent Selection for Novel Trials: $p = .76$; Condition x Number of Labels at Referent Selection: $p = .99$). The inclusion of the Condition x Referent Selection Accuracy interaction in Model 12 ($\chi^2 = 7.59, p =.055$) yielded a close to significantly better fit than the baseline model. Model 13 included the Population x Condition interaction ($\chi^2 = 7.84, p =.050$) and yielded a significant improvement in fit over the baseline model. Model 14 added the Population x Condition x Referent Selection Accuracy interaction, which did not provide a significantly better fit compared to Model 12 (*p* = .14) or Model 13 (*p* = .15). Therefore, Models 12 and 13 were compared as the best fitting models. Model 13, containing the Population x Condition interaction, was determined to be the best fitting model as it's AIC value of 656.7 and BIC value of 690.2 were lower than those of Model 12 (AIC = 657.0, BIC = 690.4).

Accuracy – individual differences

For the autistic population, we began with a baseline model including a fixed effect of Condition, and by-participant and by-word random intercepts, with a random slope of Condition per participant.

Autistic

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Animal Interest (Model 6), Receptive Vocabulary (Model 7), and Expressive Vocabulary (Model 8), were entered individually. NVIQ was not added as a fixed effect for the autistic population as not all participants successfully completed this measure. The addition of Age (*p* = .59), Attention (*p* = .12), CARS (*p* = .12), RRB (*p* = .35), Animal Interest (*p* = .89), and Receptive Vocabulary (*p* = .29) did not improve fit compared to the baseline model. Model 8, which included Expressive Vocabulary (χ^2 = 3.32, *p* =.069), yielded a close to significant improvement in fit compared to the baseline model, with a significant effect of Expressive Vocabulary (*z* = 2.17, *p* = .030). Thus, the best fitting model to the observed data was Model 8, containing a fixed effect of Expressive Vocabulary.

For the neurotypical population, we began with a baseline model including by-participant and by-word random intercepts, with a random slope of Condition per participant.

Neurotypical

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Animal Interest (Model 6), Receptive Vocabulary (Model 7), Expressive Vocabulary (Model 8), and NVIQ (Model 9), were entered individually. The addition of Age (*p* = .08), Attention (*p* = .28), CARS (*p* = .99), RRB (*p* = .83), and Animal Interest (*p* = .62), did not improve fit compared to the baseline model. The inclusion of Expressive Vocabulary (χ^2 = 3.70, *p* = .054) yielded an approaching significant improvement in fit compared to the baseline model. The inclusion of Receptive Vocabulary (χ^2 = 4.00, *p* = .046), and NVIQ (χ^2 = 4.44, *p* = .035) yielded significant improvements in fit compared to the baseline model. Model 10 therefore added fixed effects of Expressive Vocabulary + Receptive Vocabulary + NVIQ but was not significantly better fitting than Model 7 (*p* = .70), Model 8 (*p* = .60) or Model 9 (*p* = .87). Model 11 added fixed effects of Receptive Vocabulary + NVIQ but was not significantly better fitting than Model 7 (*p* = .50), Model 8 (*p* = .39) or Model 9 (*p* = .97). Model 12 added fixed effects of Expressive Vocabulary + NVIQ but was not significantly better fitting than Model 7 (*p* = .43), Model 8 (*p* = .34) or Model 9 (*p* = .67). Model 13 added fixed effects of Expressive Vocabulary + Receptive Vocabulary but was not significantly better fitting than Model 7 (*p* = .83), Model 8 (*p* = .56) or Model 9 (*p* = 1.00). Model 9, containing a fixed effect of NVIQ, was determined to be the best fitting model, as it's AIC value of 346.9 and BIC value of 368.1 were lower than those of Model 7 (AIC = 347.4, BIC = 368.6) and Model 8 (AIC = 347.7, BIC = 368.9).

Correct response times

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Models 2 and 3, fixed effects of Population (p = .51) and Condition (p = .66) did not improve fit compared to the baseline model. Model 4 included fixed effects of Population + Condition but did not differ significantly from the baseline model (p = .72). Model 5 included the Population x Condition interaction but did not yield a significantly better fit than the baseline model (p = .12). Inclusion of fixed effects did not improve predictive power.

24-Hour Retention

Accuracy

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Fixed effects of Population (Model 2), Condition (Model 3), Total Accuracy at 5-minute Retention (Model 4), Referent Selection Accuracy (Model 5), Total Accuracy at Referent Selection for Novel Trials (Model 6), and Number of Labels at Referent Selection (Model 7) were entered individually. The addition of Condition (p = .74) and Total Accuracy at Referent Selection for Novel Trials (p = .20) did not significantly improve fit compared to the baseline model. The individual addition of Population (χ^2 = 4.18, p = .041), Total Accuracy at 5-minute Retention (χ^2 = 12.79, p <.001), Referent Selection Accuracy (χ^2 = 6.83, p = .009), and Number of Labels at Referent Selection (χ^2 = 6.30, p = .012) all yielded a significant improvement in fit when compared with the baseline model. Model 8 included fixed effects of Population + Total Accuracy at 5-minute Retention + Referent Selection Accuracy + Number of Labels at Referent Selection, which yielded a significant improvement in fit when compared with Model 2 (χ^2 = 32.26, p <.001), Model 4 (χ^2 = 23.66, p <.001), Model 5 (χ^2 = 29.61, p <.001), and Model 7 (χ^2 = 30.14, p <.001). Model 9 included fixed effects of Population + Total Accuracy at 5-minute Retention + Referent Selection Accuracy but was significantly worse fitting than Model 8 (χ^2 = 10.31, p = .001). Model 10 contained fixed effects of Population + Total Accuracy at 5-minute Retention + Number of Labels at Referent Selection but was significantly worse fitting than Model 8 (χ^2 = 7.97, p = .005). Model 11 contained fixed effects of Population + Referent Selection Accuracy + Number of Labels at Referent Selection but was significantly worse fitting than Model 8 (χ^2 = 17.85, p <.001). Model 12 contained fixed effects of Total Accuracy at 5-minute Retention + Referent Selection Accuracy + Number of Labels at Referent Selection but approached a significantly worse fit than Model 8 (χ^2 = 3.44, p = .063). As such, Model 8 was retained as the current best fitting model. Model 13 contained the Population x Condition interaction, plus fixed effects of Total Accuracy at 5-minute Retention + Referent Selection Accuracy + Number of Labels at Referent Selection but did not provide a significant improvement in fit compared to Model 8 (p = .99). Model 14 contained the Population x Total Accuracy at 5-minute Retention interaction, plus fixed effects of Referent Selection Accuracy + Number of Labels at Referent Selection but did not significantly improve fit when compared to Model 8 (p = .31). Model 15 contained the Population x Referent Selection Accuracy interaction, plus fixed effects of Total Accuracy at

5-minute Retention + Number of Labels at Referent Selection but did not provide a significant improvement in fit compared to Model 8 (*p* = .09). Model 16 contained the Population x Number of Labels at Referent Selection interaction, plus fixed effects of Total Accuracy at 5-minute Retention + Referent Selection Accuracy but did not significantly improve the fit when compared to Model 8 (*p* = .45). Model 17 contained the Condition x Total Accuracy at 5-minute Retention interaction, plus fixed effects of Number of Labels at Referent Selection + Referent Selection Accuracy + Population but did not significantly improve model fit when compared to Model 8 (*p* = .51). Model 18 contained the Condition x Referent Selection Accuracy interaction, plus fixed effects of Total Accuracy at 5-minute Retention + Number of Labels at Referent Selection + Population but did not yield a better fit when compared to Model 8 (*p* = .53). Model 19 contained the Condition x Number of Labels at Referent Selection interaction, plus fixed effects of Total Accuracy at 5-minute Retention + Referent Selection Accuracy + Population but did not yield a better fit when compared to Model 8 (*p* = .41). Model 20 contained the Population x Condition x Total Accuracy at 5-minute Retention interaction, plus fixed effects of Referent Selection Accuracy + Number of Labels at Referent Selection but did not yield a better fit when compared to Model 8 (*p* = .71). Model 21 contained the Population x Condition x Referent Selection Accuracy interaction, plus fixed effects of Total Accuracy at 5-minute Retention + Number of Labels at Referent Selection but did not yield a significantly better fit compared to Model 8 (*p* = .52). Model 22 contained the Population x Condition x Number of Labels at Referent Selection interaction, plus fixed effects of Referent Selection Accuracy + Total Accuracy at 5-minute Retention but did not yield a significantly better fit when compared with Model 8 (*p* = .29). Therefore, Model 8, containing fixed effects of Population + Total Accuracy at 5-minute Retention + Referent Selection Accuracy + Number of Labels at Referent Selection, was determined to be the best fitting model for the observed data. However, since one of the model comparisons for the final model approached significance, we must be cautious when interpreting these results.

Accuracy – individual differences

Informed by the best fitting model described above, for the autistic group, we began with a baseline model including fixed effects of Total Accuracy at 5-minute Retention + Referent Selection Accuracy + Number of Labels at Referent Selection, plus by-participant and by-word random intercepts, with a random slope of Condition per participant. As these models analysed data from autistic and neurotypical groups separately, the fixed effect of Population was removed.

Autistic

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Animal Interest (Model 6), Receptive Vocabulary (Model 7), and Expressive Vocabulary (Model 8), were entered individually. NVIQ was not added as a fixed effect for the autistic population as not all participants successfully completed this measure. The addition of Age (*p* = .08), Attention (*p* = .74), Animal Interest (*p* = .48), Receptive Vocabulary (*p* = .80), and Expressive Vocabulary (*p* = .81) did not significantly improve fit compared to the baseline model. The addition of RRB ($\chi^2 = 4.14$, *p* = .042) significantly improved fit, and the addition of CARS ($\chi^2 = 3.67$, *p* = .055) resulted in approaching significant improvement in fit compared to the baseline model. Model 9 included fixed effects of RRB + CARS, which yielded a significantly better fit than Model 4 ($\chi^2 = 6.10$, $p = .014$) and Model 5 ($\chi^2 = 5.64$, $p = .018$). Therefore Model 9, containing fixed effects of RRB + CARS, was concluded to be the final best fitting model to the observed data.

Neurotypical

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Animal Interest (Model 6), Receptive Vocabulary (Model 7), Expressive Vocabulary (Model 8), and NVIQ (Model 9) were entered individually. The addition of Age ($p = .87$), Attention ($p = .22$), CARS ($p = .74$), RRB ($p = .14$), Animal Interest ($p = .68$), Receptive Vocabulary ($p = .19$), Expressive Vocabulary ($p = .28$), and NVIQ ($p = .45$) did not improve fit compared to the baseline model. The inclusion of fixed effects did not improve predictive power.

Correct response times

We began with a baseline model containing a by-participant random intercept. Addition of individual fixed effects of Condition in Model 2 ($p = .40$) and Population in Model 3 ($p = .07$) did not significantly improve fit compared to the baseline model. Model 4 included fixed effects of Population + Condition but did not differ significantly from the baseline model ($p = .14$). Model 5 included the Population x Condition interaction, which yielded a significantly better fit than the baseline model ($\chi^2 = 11.67$, $p = .009$), and contained a significant interaction effect ($t = 2.82$, $p = .005$). As such, Model 5 containing the Population x Condition interaction was determined to be the best fitting model for the observed data.

Does autistic and neurotypical children's visual attention differ while learning names for high and neutral interest stimuli, and does variability in visual attention predict learning accuracy?

Between-population comparisons

Referent selection

Proportion of time spent looking at the target stimuli

We began with a baseline model containing a by-participant random intercept. In Model 2, Population was included as a fixed effect ($\chi^2 = 5.91$, $p = .015$), which was significantly better fitting the baseline model. In Model 3, Condition was included as a fixed effect ($p = 1.00$), which was not significantly better fitting than the baseline model. In Model 4, Trial Type was included as a fixed effect ($\chi^2 = 18.42$, $p <.001$), which was significantly better fitting than the baseline model. Model 5 included fixed effects of Trial Type + Population and was significantly better fitting compared to Model 2 ($\chi^2 = 18.42$, $p <.001$) and Model 4 ($\chi^2 = 5.91$, $p = .015$). Model 6 (Population + Condition) did not significantly improve fit compared to Model 2 ($p = 1.00$) or Model 4 ($p = 1.00$). Model 7 (Trial Type + Condition) was a significantly better fitting model than Model 2 ($\chi^2 = 12.51$, $p <.001$) but not Model 4 ($p = 1.00$). Model 8 included fixed effects of Trial Type + Population + Condition but did not significantly improve fit compared to Model 5 ($p = 1.00$). Inclusion of individual interactions

(Model 9, Trial Type x Population: *p* = .24; Model 10, Condition x Population: *p* = 1.00; Model 11, Trial Type x Condition: *p* = .22), did not provide a significantly better fit compared to Model 5. Model 12 included the Trial Type x Population x Condition interaction (χ^2 = 10.28, *p* = .068) which yielded an approaching significantly better fit compared to Model 5, but the three-way interaction effect within the model was not significant (*t* = - 1.19, *p* = .23). However, the fixed effect of Population (*t* = - 2.55, *p* = .016) and the Trial Type x Condition interaction (*t* = 2.70, *p* = .007) within the model were significant. As such, Model 13 included the Trial Type x Condition interaction, plus a fixed effect of Population, which afforded a significantly better fit than Model 5 (χ^2 = 7.42, *p* = .025). Therefore, the best fitting model was Model 12, containing the Trial Type x Condition interaction plus fixed effect of Population.

Number of looks to the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type random slopes per participant. In Model 2, Population was included as a fixed effect (χ^2 = 8.75, *p* =.003), which was significantly better fitting than the baseline model. In Model 3, Condition was included as a fixed effect (*p* = .12), which did not provide a significantly better fit than the baseline model. Model 4 included a fixed effect of Trial Type (χ^2 = 4.99, *p* = .025), which was significantly better fitting than the baseline model. Model 5 included fixed effects of Trial Type + Population, which provided a significant improvement in over Model 2 (χ^2 = 5.01, *p* = .025), and Model 4 (χ^2 = 8.77, *p* = .003). Model 6 included fixed effects of Population + Condition, and significantly improved fit compared to Model 4 (χ^2 = 6.24, *p* = .012), but not Model 2 (*p* = .12). Model 7 included fixed effects of Trial Type + Condition but did not significantly improve fit compared to Model 2 (*p* = 1.00) or Model 4 (*p* = .11). Model 8 included fixed effects of Trial Type + Population + Condition (*p* = .11) but did not significantly improve fit compared to Model 5. Inclusion of individual interactions in Models 9-11 (Trial Type x Population: *p* = .20; Condition x Population: *p* = 1.00; Trial Type x Condition: *p* = 1.00), did not provide a significantly better fit compared to Model 5. Model 12 included the Trial Type x Population x Condition interaction (*p* = .33) but did not improve fit compared to Model 5. Therefore, the best fitting model was Model 5, containing fixed effects of Trial Type + Population.

Longest look novel or known

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. In Models 2 and 3, adding Population (*p* = .34) and Condition (*p* = .14) as fixed effects did not provide a significantly better fit for the observed data than the baseline model. Model 4 included the fixed effect of Trial Type (χ^2 = 37.45, *p* <.001), which yielded a significantly better fit for the observed data than the baseline model. Inclusion of multiple fixed effects in Model 5 (Condition + Population, *p* = 1.00), Model 6 (Trial Type + Population, *p* = .34), Model 7 (Trial Type + Condition, *p* = .18) and Model 8 (Trial Type + Population + Condition, *p* = .24) did not significantly improve fit compared to Model 4. Model 9 contained the Condition x Population interaction (*p* = 1.00), which did not provide a significantly better fit than Model 4. Model 10 contained an interaction effect of Trial Type x Population (χ^2 = 5.80, *p* = .055), which approached a significant improvement in over Model 4, with a significant interaction effect (*z* = -2.36, *p* = .018). Model 11 included the Trial Type x Condition interaction (*p* = .38) but was not significantly better fitting than Model 4. Model 12 included the Trial Type x Population x Condition interaction (*p* = .53) but was not significantly better fitting than Model 10. Therefore, the best fitting model was Model 10, including the Trial Type x Population interaction. However, since the model comparison for the final model approached significance, we must be cautious when interpreting these results.

5-minute retention

Proportion of time spent looking at the target stimuli

We began with a baseline model containing a by-participant random intercept, although simplification of the random effects still yielded a singular fit. Inclusion of fixed effects in Model 2 (Population, *p* = .67), and Model 3 (Condition, *p* = .37), did not provide a significantly better fit than the baseline model. Model 4 included fixed effects of Condition + Population (*p* = .61) but was not significantly better fitting compared to the baseline model. Model 5 included the Condition x Population interaction (*p* = .34) but was not significantly better fitting than the baseline model. The inclusion of fixed effects did not improve predictive power.

Number of looks to the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Including fixed effects of Population in Model 2 (*p* = .12) and Condition in Model 3 (*p* = .64) did not yield a significantly better fit than the baseline model. Model 4 included fixed effects of Condition + Population (*p* = .27) but was not significantly better fitting than the baseline model. Model 5 included the Condition x Population interaction ($\chi^2 = 7.26$, *p* = .064), which approached a significant improvement in over the baseline model and contained a significant interaction effect (*t* = 2.24, *p* = .033). Therefore, the final model was Model 5, including the Condition x Population interaction. However, since the model comparison for the final model approached significance, we must be cautious when interpreting these results.

Longest look target or foil

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. The addition of fixed effects of Population (Model 2; *p* = .95), and Condition (Model 3; *p* = .43), did not provide a significantly better fit than the baseline model. Model 4 contained a fixed effect of Condition + Population (*p* = .73) but was not significantly better fitting than the baseline model. Model 5 included the Condition x Population interaction (*p* = .11), which was not significantly better fitting than the baseline model. The inclusion of fixed effects did not improve predictive power.

24-hour retention

Proportion of time spent looking at the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts. Model 2 included the fixed effect of Population (*p* = .28), and Model 3 included the fixed effect of Condition (*p* = .52), but neither provided a significantly better fit than the baseline model. Model 4 included a fixed effect of Condition + Population (*p* = .46) but did not significantly improve fit compared to the baseline model. Model 5 included the Condition x Population interaction (*p* = .33), which was not significantly better fitting than the baseline model. The baseline model therefore provided the best fit to the observed data.

Number of looks to the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Including Population as a fixed effect in Model 2 yielded a significant improvement in fit over the baseline model (χ^2 = 4.85, *p* = .028), but including a fixed effect of Condition in Model 3 did not (*p* = .81). Model 4 included fixed effects of Condition + Population (*p* = .74) but was not significantly better fitting than Model 2. Model 5 included the Condition x Population interaction (*p* = .38), which was not significantly better fitting than Model 2. Therefore, the final model was Model 2, including a fixed effect of Population.

Longest look target or foil

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Model 2 contained a fixed effect of Population (*p* = .50), and Model 3 included a fixed effect of Condition (*p* = .42), but neither models were significantly better fitting than the baseline model. Model 4 included a fixed effect of Condition + Population (*p* = .60), which was not significantly better fitting than the baseline model. Model 5 included the Condition x Population interaction (*p* = .57), which was not significantly better fitting than the baseline model. The inclusion of fixed effects did not improve predictive power.

Looking time predicting accuracy

The following analyses examined whether children's referent selection accuracy was influenced by a different looking behaviour, and whether the effects of those behaviours differed across populations and conditions (by testing whether the inclusion of x population and x condition interaction terms significantly improved model fit). All models were conducted using the glmer function from the lme4 package in R (Bates et al., 2015).

Referent Selection

Proportion of time spent looking at the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. In Model 2, Proportion to Target was included as a fixed effect (χ^2 = 153.30, *p* <.001), which was significantly better fitting than the baseline model. Model 3 included the Proportion to Target x Trial Type interaction (χ^2 = 7.05, *p* = .029), which was significantly better fitting than Model 2, but the interaction effect within the model was not significant (*z* = -0.64, *p* = .52). Neither Model 4 (Proportion to

Target x Condition: $p = .71$, or Model 5 (Proportion to Target x Population: $p = .81$), provided a significantly better fit than Model 2. None of the three-way interactions in Model 6 (Proportion to Target x Population x Trial Type: $p = .16$), Model 7 (Proportion to Target x Population x Condition: $p = .39$), and Model 8 (Proportion to Target x Condition x Trial Type: $p = .34$), afforded a significantly better fit than Model 2. Therefore, the final best fitting model for the observed data was determined to be Model 2, including a fixed effect of Proportion to Target.

Number of looks to the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. In Model 2, Number Looks to Target was included as a fixed effect ($\chi^2 = 36.68$, $p <.001$), which was significantly better fitting than the baseline model. Model 3 included the Number Looks to Target x Trial Type interaction ($\chi^2 = 17.70$, $p <.001$), which was significantly better fitting than Model 2, but did not have a significant interaction effect ($z = 1.33$, $p = .18$). Model 4 contained the Number Looks to Target x Condition interaction ($p = .61$) but was not significantly better fitting than Model 2. Model 5 included the Number Looks to Target x Population interaction and was significantly better fitting than Model 2 ($\chi^2 = 8.19$, $p = .017$), with a significant interaction effect ($z = -2.00$, $p = .046$). Model 6 included the Number Looks to Target x Population x Trial Type interaction, but this model would not converge. Model 7 included the Number Looks to Target x Population x Condition interaction ($p = .10$) but was not significantly better fitting than Model 5. Model 8 included the Number Looks to Target x Condition x Trial Type interaction ($\chi^2 = 10.92$, $p = .028$), which was significantly better fitting than Model 5, but did not have a significant three-way interaction effect ($z = 0.37$, $p = .71$). Therefore, Model 5, including the Number Looks to Target x Population interaction was established as the final model.

Longest look novel or known

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. In Model 2, Longest Look Novel was included as a fixed effect ($\chi^2 = 4.64$, $p = .031$), which yielded a significantly better fit than the baseline model. Model 3 included the Longest Look Novel x Trial Type interaction ($\chi^2 = 62.97$, $p <.001$), which was significantly better fitting than Model 2, and had a significant interaction effect ($z = 6.20$, $p <.001$). Inclusion of the Longest Look Novel x Condition interaction ($p = 1.00$) in Model 4 did not provide a significantly better fit than Model 2. Model 5 included the Longest Look Novel x Population interaction ($p = .22$) which was not significantly better fitting than Model 2. Models 6-8 included three-way interactions (Longest Look Novel x Condition x Trial Type: $p = .82$; Longest Look Novel x Trial Type x Population: $p = .44$; Longest Look Novel x Population x Condition: $p = 1.00$), also did not provide a significantly better fit than Model 3. Overall, the final best fitting model for the observed data was Model 3, including the Longest Look Novel x Trial Type interaction.

5-minute retention

Proportion of time spent looking at the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Proportion to Target was included as a fixed effect ($\chi^2 = 297.24$, $p < .001$), which was significantly better fitting than the baseline model. Model 3 included the Proportion to Target x Condition interaction ($p = .54$), and Model 4 included the Proportion to Target x Population interaction ($p = .77$), but neither were significantly better fitting than Model 2. Model 5 included the Proportion to Target x Population x Condition interaction ($p = .31$) but was not significantly better fitting than Model 2. Therefore, the final best fitting model for the observed data was determined to be Model 2, including a fixed effect of Proportion to Target.

Number of looks to the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Number Looks to Target was included as a fixed effect ($\chi^2 = 52.13$, $p < .001$), which was significantly better fitting than the baseline model. Inclusion of the Number Looks to Target x Condition interaction in Model 3 approached a significant improvement in fit over Model 2 ($\chi^2 = 5.70$, $p = .058$), and contained a significant interaction effect ($z = 2.60$, $p = .009$). Model 4 contained the Number Looks to Target x Population interaction ($p = .26$) but did not yield a significantly better fit than Model 2. Model 5 included the Number Looks to Target x Population x Condition interaction but was not significantly better fitting than Model 3 ($p = .28$). Therefore, the final best fitting model for the observed data was determined to be Model 3, including the Number Looks to Target x Condition interaction. However, since the model comparison for the final model approached significance, we must be cautious when interpreting these results.

Longest look target or foil

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Longest Look Target was included as a fixed effect ($\chi^2 = 179.44$, $p < .001$), which was significantly better fitting than the baseline model. Model 3 included the Longest Look Target x Condition interaction ($p = .87$), and Model 4 included the Longest Look Target x Population interaction ($p = .20$), but neither were significantly better fitting than Model 2. Model 5 included the Longest Look Target x Population x Condition interaction ($p = .19$) but was not significantly better fitting than Model 2. Therefore Model 2, containing a fixed effect of Longest Look Target, was determined to be the final model.

Proportion of time spent looking at the target stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Proportion to Target was included as a fixed effect ($p = .67$), which was not significantly better fitting than the baseline model. Model 3 included the Proportion to Target x Condition interaction ($p = .65$), and Model 4 included the Proportion to Target x Population interaction ($p = .87$), but neither were significantly better fitting than the baseline model. Model 5 included the Proportion to Target x Population x Condition interaction ($p = .32$), which not significantly better fitting than the baseline model. The inclusion of fixed effects did not improve predictive power.

Number of looks to the target stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Number Looks to Target was included as a fixed effect (χ^2 = 5.35, *p* = .021), which was significantly better fitting than the baseline model. Model 3 included the Number Looks to Target x Condition interaction (*p* = .29), and Model 4 included the Number Looks to Target x Population interaction (*p* = .58), but neither were significantly better fitting than Model 2. Model 5 included the Number Looks to Target x Population x Condition interaction (χ^2 = 11.80, *p* = .067), which was borderline significantly better fitting than Model 2, but the interaction effect was not significant (*z* = 0.04, *p* = .97). Therefore, the final best fitting model for the observed data was determined to be Model 2 (Number Looks to Target).

Longest look to target or foil stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Longest Look Target was included as a fixed effect (*p* = .24), which was not significantly better fitting than the baseline model. Model 3 included the Longest Look Target x Condition interaction (*p* = .69), and Model 4 included the Longest Look Target x Population interaction (*p* = .68), but neither were significantly better fitting than the baseline model. Model 5 included the Longest Look Target x Population x Condition interaction (*p* = .26) but was not significantly better fitting than the baseline model. The inclusion of fixed effects did not improve predictive power.

24-hour retention

Proportion of time spent looking at the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Proportion to Target was included as a fixed effect (χ^2 = 263.96, *p* <.001), which was significantly better fitting than the baseline model. Model 3 included the Proportion to Target x Condition interaction (*p* = .54), which was not significantly better fitting than Model 2. Model 4 included the Proportion to Target x Population interaction (χ^2 = 8.93, *p* = .012), which was significantly better fitting than Model 2, but did not contain a significant interaction effect (*z* = -1.50, *p* = .13). Model 5 included the Proportion to Target x Population x Condition interaction (χ^2 = 12.87, *p* = .045), which was significantly better fitting than Model 2, but did not contain a significant three-way interaction effect (*z* = -1.46, *p* = .14). Therefore, Model 2 was retained as the final best fitting model, containing a fixed effect of Proportion to Target.

Number of looks to the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Number Looks to Target was included as a fixed effect (χ^2 = 65.37, *p* <.001), which was significantly better fitting than the baseline model. Model 3 included the Number Looks to Target x Condition interaction (*p* = .22) but was not significantly better fitting than Model 2. Model 4 included the Number Looks to Target x Population interaction (χ^2 = 11.07, *p* = .004), which was significantly better fitting than Model 2 and contained a significant interaction effect (*z* = -3.02, *p* = .002). Model 5 included the Number Looks to Target x Population x Condition interaction (*p* = .27) but was not significantly better fitting than Model 4. Therefore, the final best fitting model for the observed data was determined to be Model 4, including the Number Looks to Target x Population interaction.

Longest look target or foil

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Longest Look Target was included as a fixed effect (χ^2 = 143.27, *p* <.001), which was significantly better fitting than the baseline model. Model 3 included the Longest Look Target x Condition interaction (*p* = .89), which was not significantly better fitting than Model 2. Model 4 included the Longest Look Target x Population interaction (χ^2 = 14.95, *p* <.001), which was significantly better fitting than Model 2 and contained a significant interaction effect (*z* = -2.78, *p* = .005). Model 5 included the Longest Look Target x Population x Condition interaction (*p* = .82), which was not significantly better fitting than Model 4. Overall, Model 4 containing the Longest Look Target x Population interaction was the best fitting model.

Proportion of time spent looking at the target stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Proportion to Target was included as a fixed effect (χ^2 = 3.65, *p* = .056), which approached a significant improvement in over the baseline model. Model 3 included the Proportion to Target x Condition interaction (*p* = .56), which was not significantly better fitting than Model 2. Model 4 included the Proportion to Target x Population interaction (χ^2 = 5.49, *p* = .064), which approached a significant improvement in fit over Model 2 but did not contain a significant interaction effect (*z* = 0.81, *p* = .42). Model 5 included the Proportion to Target x Population x Condition interaction (χ^2 = 12.33, *p* = .055), which approached a significant improvement in fit over Model 2 and contained a significant interaction effect (*z* = -2.03, *p* = .042). Model 5 (Proportion to Target x Population x Condition) was therefore determined to be the best fitting model. However, since the model comparison for the final model approached significance, we must be cautious when interpreting these results.

Number of looks to the target stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Number Looks to Target was included as a fixed effect (χ^2 = 11.01, *p* <.001), which was significantly better fitting than the baseline model. Model 3 included the Number Looks to Target x Condition interaction (*p* = .64) but was not significantly better fitting than Model 2. Model 4 included the Number Looks to Target x Population interaction (*p* = .28), which was not significantly better fitting than Model 2. Model 5 included the Number Looks to Target x Population x Condition interaction (*p* = .59) but was not significantly better fitting than Model 2. Therefore, the final best fitting model for the observed data was determined to be Model 2 (Number Looks to Target).

Longest look to target or foil stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Longest Look Target was included as a fixed effect (χ^2 = 6.22, *p* = .013), which was significantly better fitting than the baseline model. Model 3 included the Longest Look Target x Condition interaction (*p* = .58), which was not significantly better fitting than Model 2. Model 4 included the Longest Look Target x Population interaction (χ^2 = 5.50, *p* = .064), which approached a significant improvement in fit over Model 2 but did not contain a significant interaction effect (*z* = 0.70, *p* = .49). Model 5 included the Longest Look Target x Population x Condition interaction (*p* = .11), which was not significantly better fitting than Model 2. Overall, Model 2 (Longest Look Target) was the best fitting model.

Appendix D: Supplementary materials for analyses in Chapter 4

Do autistic and neurotypical children's stimulus preferences influence their accuracy and visual attention when learning novel words?

Model building sequences for results reported in the main text

All models were conducted using the glmer and lmer functions from the lme4 package in R (Bates et al., 2015). Population was contrast coded as -0.5 (neurotypical) and 0.5 (autistic). Condition was contrast coded as -0.5 (novel object) and 0.5 (novel animal). Trial type was coded as -0.5 (known) and 0.5 (novel).

Receptive vocabulary was coded as the participant's age equivalent based on their raw score on the British Picture Vocabulary Scale 2 (BPVS; Dunn et al., 1997). Expressive Vocabulary was coded as the participant's age equivalent based on their raw score on the Expressive Vocabulary Test 2 (EVT; Williams, 2007) or the expressive language module of the Mullen's Scales of Early Learning (Mullen, 1995). Attention was coded as the participant's raw score on the Conner's Teacher Rating Scale 15 (CTRS-15; Pupura & Lonigan, 2009). Autism severity was coded as the participant's raw score on the Childhood Autism Rating Scale 2 (CARS; Schopler et al., 2010). Non-verbal intellectual abilities were coded as children's raw score on the Leiter-3 test of non-verbal intelligence (NVIQ; Roid et al., 2013). Repetitive behaviour was coded as children's raw score on the Repetitive Behaviour Questionnaire (RBQ; Leekam et al., 2007). Animal Interest was coded as children's raw score on our Animal Interest questionnaire. Chronological Age was measured in months. Referent selection accuracy was coded as -0.5 (incorrect) and 0.5 (correct). Number of labels at referent selection for each novel word was coded as 1-5. By-word 5-minute retention accuracy was coded as 0-2. Trial-level accuracy for analyses at each experimental phase was coded as 1 (correct) or 0 (incorrect).

Proportion of time spent looking at the target object on each trial was scored between 0 and 1 for all participants. Number of looks to the target object on each trial ranged from 0 to 12, with every new look towards the target stimuli being counted. The longest look to novel or known stimuli was coded as -0.5 (longest look to known object) or 0.5 (longest look to novel object) for the predicting accuracy analyses and known (0) novel (1) for the between group differences analyses. This variable was only included in referent selection analyses, as the retention phases involved only novel objects. The longest look to target or foil stimuli was coded as -0.5 (longest look to foil stimuli) and 0.5 (longest look to target stimuli) for the predicting accuracy analyses, and foil (0) target (1) for the between group differences analyses.

All analyses were undertaken using the same procedure. Models were built up sequentially, adding in one fixed effect at a time and comparing each model with the previous best-fitting model using log-likelihood tests. Each model was built up from a null model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant (referent selection analyses) or Condition slopes per participant (5 minute and 24-hour retention analyses). If some models in a sequence were singular fitting or failed to converge, the random effects were simplified until all models in the sequence successfully converged (the final random effects structure for each variable are provided below).

Do autistic and neurotypical children differ in accuracy and response times when learning names for liked and disliked stimuli?

Referent Selection

Accuracy

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. Fixed effects of Population (Model 2), Condition (Model 3), and Trial Type (Model 4) were entered individually. The addition of Population ($\chi^2 = 4.05$, $p = .044$) and Trial Type ($\chi^2 = 4.51$, $p = .034$) approached a significant improvement in fit when compared with the baseline model. The addition of Condition ($p = .46$) did not significantly improve fit. Model 5 included fixed effects of Population + Trial Type, yielding a significantly better fit than Model 2 ($\chi^2 = 5.32$, $p = .021$), and Model 3 ($\chi^2 = 4.87$, $p = .027$). Model 6 included fixed effects of Condition + Trial Type but was not significantly better fitting than Model 2 ($p = .34$), or Model 3 ($p = .51$). Model 7 included fixed effects of Trial Type + Population + Condition but did not yield a significantly better fit than Model 5 ($p = .57$). Model 8 included the Trial Type x Population interaction but did not afford a significantly better fit than Model 5 ($p = .73$). Model 9 included the Trial Type x Condition interaction but did not differ significantly from Model 5 ($p = 1.00$). Model 10 included the Population x Condition interaction but did not significantly improve fit compared to Model 5 ($p = 1.00$). Finally, the three-way interaction was entered (Model 11), but did not significantly improve fit compared to Model 5 ($p = .90$). Thus, Model 5, containing fixed effects of Population and Trial Type, was the best fit to the observed data.

Accuracy – individual differences

Informed by the best fitting model described above, we began with a baseline model including a fixed effect of Trial Type, and by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. These models analyse data from the autistic and neurotypical groups separately, so the fixed effect of Population was removed.

Autistic

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5) and Receptive Vocabulary (Model 6) were entered individually. NVIQ and Expressive Vocabulary were not added as fixed effects for the ASD population as not all participants successfully completed this measure. The addition of Age ($p = .18$), Attention ($p = .73$), and CARS ($p = .12$) did not improve fit compared to the baseline model. The individual addition of RRB yielded a borderline significantly better fit compared to the baseline model ($\chi^2 = 3.80$, $p = .051$) and the addition of Receptive Vocabulary yielded a significantly better fit compared to the baseline model ($\chi^2 = 12.71$, $p < .001$). Therefore, Model 7 included fixed effects of RRB + Receptive Vocabulary, yielding a significantly better fit than Model 5 ($\chi^2 = 14.92$, $p < .001$) and Model 6 ($\chi^2 = 6.00$, $p = .014$). As such, Model 7, with fixed effects of RRB and Receptive Vocabulary, was concluded to be the overall best fitting model to the observed data.

Neurotypical

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Receptive Vocabulary (Model 6), Expressive Vocabulary (Model 7), and NVIQ (Model 8) were entered individually. The addition of Attention (*p* = 1.00), CARS (*p* = 1.00), and RRB (*p* = .09) did not improve fit compared to the baseline model. However, the addition of Age (χ^2 = 6.56, *p* = .010), Receptive Vocabulary (χ^2 = 6.20, *p* = .013), Expressive Vocabulary (χ^2 = 7.90, *p* = .005), and NVIQ (χ^2 = 5.01, *p* = .025) differed significantly from the baseline model. Therefore, Model 9 added fixed effects of Age + Receptive Vocabulary + Expressive Vocabulary + NVIQ. Model 9 did not differ significantly from Model 2 (*p* = .40), Model 6 (*p* = .35), Model 7 (*p* = .66), or Model 8 (*p* = .21). Model 10 included fixed effects of Age + Receptive Vocabulary + Expressive Vocabulary but did not differ significantly from Model 2 (*p* = .25), Model 6 (*p* = .21), Model 7 (*p* = .49), or Model 8 (*p* = .11). Model 11 added fixed effects of Age + Receptive Vocabulary + NVIQ and did not differ significantly from Model 2 (*p* = .74), Model 6 (*p* = .62), Model 7 (*p* = 1.00), or Model 8 (*p* = .34). Model 12 added fixed effects of Age + Expressive Vocabulary + NVIQ and did not differ significantly from Model 2 (*p* = .39), Model 6 (*p* = .33), Model 7 (*p* = .77), or Model 8 (*p* = .18). Model 13 included fixed effects of Receptive Vocabulary + Expressive Vocabulary + NVIQ but did not yield a significantly better fit compared to Model 2 (*p* = .25), Model 6 (*p* = .21), Model 7 (*p* = .49), or Model 8 (*p* = .12). Model 14 added fixed effects of Receptive Vocabulary + Age but was not significantly better fitting than Model 2 (*p* = .50), Model 6 (*p* = .37), Model 7 (*p* = 1.00), or Model 8 (*p* = .16). Model 15 added fixed effects of Expressive Vocabulary + Age but was not significantly better fitting than Model 2 (*p* = .25), Model 6 (*p* = .19), Model 7 (*p* = .93), or Model 8 (*p* = .09). Model 16 added fixed effects of NVIQ + Age but did not yield a significantly better fit compared to Model 2 (*p* = .90), Model 6 (*p* = .54), Model 7 (*p* = 1.00), or Model 8 (*p* = .21). Model 17 added fixed effects of Receptive Vocabulary + Expressive Vocabulary, which did not yield a significantly better fit compared to Model 2 (*p* = .12), Model 6 (*p* = .09), or Model 7 (*p* = .29), but was significantly better fitting than Model 8 (χ^2 = 4.03, *p* = .045). Model 18 added fixed effects of Receptive Vocabulary + NVIQ, which did not yield a significantly better fit compared to Model 2 (*p* = 1.00), Model 6 (*p* = .97), Model 7 (*p* = 1.00), or Model 8 (*p* = .27). Model 19 added fixed effects of Expressive Vocabulary + NVIQ, which did not yield a significantly better fit compared to Model 2 (*p* = .18), Model 6 (*p* = .14), or Model 7 (*p* = .48), but was borderline significantly better fitting than Model 8 (χ^2 = 3.39, *p* = .066). Therefore, Models 2, 6, and 7 were retained for comparison. Overall, Model 7, containing a fixed effect of Expressive Vocabulary, was determined to be the best fitting model as its AIC value of 178.4 and BIC value of 228.8 were lower than those of Model 2 (AIC = 179.7; BIC = 230.1) and Model 6 (AIC = 180.1; BIC = 230.5).

Correct response times

We began with a baseline model containing a by-participant random intercept. Fixed effects of Population (Model 2), Condition (Model 3), and Trial Type (Model 4) were entered individually. The addition of Population (*p* = .43) did not improve fit compared to the baseline model. The addition of Condition (χ^2 = 3.83, *p* = .050) and Trial Type (χ^2 = 33.56, *p* <.001) yielded a significant improvement in fit compared to the baseline model. Model 5 added fixed effects of Population alongside Trial Type, which was significantly better fitting than Model 3 (χ^2 = 30.44, *p* <.001) but did not significantly improve fit compared to Model 4 (*p* = .40). Model 6 added fixed effects of Trial Type and Condition, which was significantly better fitting than Model 3 (χ^2 = 33.98, *p* <.001) and Model 4 (χ^2 = 4.24, *p* = .040). Model 7 included all three fixed effects but did not significantly improve fit in comparison to Model 6 (*p* = .40). Model 8 included the Trial Type x Population interaction (*p* = .79) but did not yield a significant improvement in fit compared to Model 6. Model 9 included the Trial Type x Condition interaction (*p* = .86), and Model 10 included the Population x Condition interaction (*p* = 1.00), but neither model differed significantly from Model 6. Finally, the three-way interaction was entered (Model 11); doing so did not improve fit in comparison to Model 6 (*p* = .48). As such Model 6 containing fixed effects of Trial Type and Condition provided the best fit to the observed data.

5-Minute Retention

Accuracy

We began with a baseline model a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Fixed effects of Population (Model 2), Condition (Model 3), Referent Selection Accuracy (Model 4), and Number of Labels at Referent Selection (Model 5) were entered individually. The addition of Population (*p* = .62), Condition (*p* = .83), and Number of Labels at Referent Selection (*p* = .45) did not yield a significantly better fitting model compared to the baseline model. The addition of Referent Selection Accuracy ($\chi^2 = 6.01$, *p* = .014) yielded a significant improvement in fit compared to the baseline model. Models 6-10 contained the two-way interactions plus significant fixed effect of Referent Selection Accuracy. Model 6 (Population x Condition + Referent Selection Accuracy, *p* = .34), Model 7 (Population x Referent Selection Accuracy, *p* = .67), Model 8 (Population x Number of Labels at Referent Selection + Referent Selection Accuracy (*p* = .75), Model 9 (Condition x Referent Selection Accuracy, *p* = .88), and Model 10 (Condition x Number of Labels at Referent Selection + Referent Selection Accuracy, *p* = .96) did not yield a significantly better fit than Model 4. Model 11 contained the Population x Condition x Referent Selection Accuracy interaction but did not yield a significantly better fit than Model 4 (*p* = .61). Model 12 contained the Population x Condition x Number of Labels at Referent Selection interaction plus fixed effect of Referent Selection Accuracy but did not yield a significantly better fit than Model 4 (*p* = .66). Therefore, Model 4 containing a fixed effect of Referent Selection Accuracy was determined to be the best fitting model.

Accuracy – individual differences

We began with a baseline model containing a fixed effect of Referent Selection Accuracy, and by-participant and by-word random intercepts, with a random slope of Condition per participant. These models were performed for autistic and neurotypical populations separately.

Autistic

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5) and Receptive Vocabulary (Model 6) were entered individually. NVIQ and Expressive Vocabulary were not added as fixed effects for the ASD population as not all participants successfully completed this measure. The addition of Age (*p* = .63), Attention (*p* = .93), CARS (*p* = .29), and RRB (*p* = .29) did not improve fit compared to the baseline model. The inclusion of

Receptive Vocabulary ($\chi^2 = 4.72$, $p = .030$) did provide a significantly better fit compared to the baseline model. Therefore, Model 6, containing a fixed effect of Receptive Vocabulary, was the best fit to the observed data.

Neurotypical

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Receptive Vocabulary (Model 6), Expressive Vocabulary (Model 7), and NVIQ (Model 8) were entered individually. The addition of Age ($p = .64$), Attention ($p = .56$), CARS ($p = .93$), RRB ($p = .16$), Receptive Vocabulary ($p = .10$), Expressive Vocabulary ($p = .23$) and NVIQ ($p =.10$) did not significantly improve fit compared to the baseline model. Inclusion of fixed effects did not improve predictive power.

Correct response times

We began with a baseline model containing a by-participant random intercept. Fixed effects of Population (Model 2) and Condition (Model 3) were entered individually. The addition of Population ($p = .82$) and Condition ($p = .87$) did not improve fit compared to the baseline model. Model 4 included fixed effects of Population + Condition but did not differ significantly from the baseline model ($p = .96$). Model 5 included the Population x Condition interaction but did not significantly improve fit compared to the baseline model ($p = .71$). Therefore, the inclusion of fixed effects did not improve predictive power.

24-Hour Retention

Accuracy

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Fixed effects of Population (Model 2), Condition (Model 3), 5-minute Retention Accuracy (Model 4), Referent Selection Accuracy (Model 5), and Number of Labels at Referent Selection (Model 6) were entered individually. The addition of Population ($p = .62$), Condition ($p = .39$), and Number of Labels at Referent Selection ($p = .11$) did not improve fit compared to the baseline model. The individual addition of by-word 5-minute Retention Accuracy ($\chi^2 = 7.93$, $p = .005$) and Referent Selection Accuracy ($\chi^2 = 4.96$, $p = .026$) yielded a significant improvement in fit when compared with the baseline model. Model 7 added fixed effects of 5-minute Retention Accuracy + Referent Selection Accuracy, which was borderline significantly better fitting than Model 4 ($\chi^2 = 3.34$, $p = .068$) and significantly better fitting than Model 5 ($\chi^2 = 6.31$, $p = .012$). Model 8 contained a two-way interaction of Population x Condition, plus fixed effects of by-word 5-minute Retention Accuracy + Referent Selection Accuracy, which was not significantly better fitting than Model 7 ($p = .44$). Model 9 contained a two-way interaction of Population x 5-minute Retention Accuracy, plus a fixed effect of Referent Selection Accuracy, which was not significantly better fitting than Model 7 ($p = .87$). Model 10 contained a two-way interaction of Population x Referent Selection Accuracy, plus a fixed effect of 5-minute Retention Accuracy, which was not significantly better fitting than Model 7 ($p = .80$). Model 11 contained a two-way interaction of Population x Number of Labels at Referent Selection, plus fixed effects of Referent Selection Accuracy and 5-minute Retention Accuracy but was not significantly better fitting than Model 7 ($p = .58$). Model 12 included the Condition x 5-minute Retention Accuracy, plus a fixed effect of Referent Selection Accuracy, which was not significantly better fitting compared to Model 7 (p = .12). Model 13 contained a two-way interaction of Condition x Referent Selection Accuracy, plus a fixed effect of 5-minute Retention Accuracy which was not significantly better fitting than Model 7 (p = .40). Model 14 contained a two-way interaction of Condition x Number of Labels at Referent Selection, plus fixed effects of Referent Selection Accuracy and 5-minute Retention Accuracy but was not significantly better fitting than Model 7 (p = .27). Model 15 contained a three-way interaction of Population x Condition x Referent Selection Accuracy, plus a fixed effect of by-word 5-minute retention accuracy but was not significantly better fitting than Model 7 (p = .61). Model 16 contained a three-way interaction of Population x Condition x 5-minute Retention Accuracy, plus a fixed effect of Referent Selection Accuracy but was not significantly better fitting than Model 7 (p = .20). Model 17 contained a three-way interaction of Population x Condition x Number of Labels at Referent Selection, plus fixed effects of 5-minute Retention Accuracy and Referent Selection Accuracy but was not significantly better fitting than Model 7 (p = .62). As such, Model 7 containing the fixed effects of 5-minute Retention Accuracy and Referent Selection Accuracy was retained as the final model. However, one of the model comparisons only approached significance, so we must interpret this result with caution.

Accuracy – individual differences

Informed by the best fitting model described above, we began with a baseline model including fixed effects of Referent Selection Accuracy, 5-minute Retention Accuracy, and Condition, plus by-participant and by-word random intercepts, with a random slope of Condition per participant. These models analysed data from autistic and neurotypical groups separately.

Autistic

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5) and Receptive Vocabulary (Model 6) were entered individually. NVIQ and Expressive Vocabulary were not added as fixed effects for the ASD population as not all participants successfully completed this measure. The addition of Age (p = .77), Attention (p = 1.00), CARS (p = .45), RRB (p = .36) and Receptive Vocabulary (p = .18) did not significantly improve fit compared to the baseline model. The inclusion of fixed effects did not improve predictive power.

Neurotypical

Fixed effects of Age (Model 2), Attention (Model 3), CARS (Model 4), RRB (Model 5), Receptive Vocabulary (Model 6), Expressive Vocabulary (Model 7), and NVIQ (Model 8) were entered individually. The addition of Age (p = .84), CARS (p = .23), RRB (p = .44), Receptive Vocabulary (p = .45), Expressive Vocabulary (p = .58), and NVIQ (p = .78) did not improve fit compared to the baseline model. The addition of Attention (χ^2 = 5.95, p = .015) did yield a significantly better fit compared to the baseline model, and thus was determined to be the best fitting model to the observed data.

Correct response times

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. The individual addition of Population (Model 2; $p = .19$) and Condition (Model 3; $p = .28$) did not significantly improve fit compared to the baseline model. Model 4 included fixed effects of Population + Condition but did not differ significantly from the baseline model ($p = .23$). Model 5 included the two-way interaction between Population x Condition ($p = .23$) but did not differ significantly from the baseline model. The inclusion of fixed effects did not improve predictive power.

Do autistic and neurotypical children's visual attention differ while learning names for liked and disliked stimuli, and does variability in visual attention predict their learning accuracy?

Between-population comparisons

Referent selection

Proportion of time spent looking at the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts. In Models 2 and 3, the addition of individual fixed effects of Population ($p = .17$), Condition ($p = .75$), did not yield a significantly better fit than the baseline model. In Model 4, Trial Type was included as a fixed effect ($\chi^2 = 21.15, p <.001$), which was significantly better fitting than the baseline model. Models 5, 6 and 7 included pairs of fixed effects (Trial Type + Population, $p = .16$; Condition + Population, $p = 1.00$; Condition + Trial Type, $p = .66$), but none of the models were significantly better fitting than Model 4. Model 8 included fixed effects of Trial Type + Population + Condition ($p = .34$) but did not significantly improve fit compared to Model 4. Inclusion of individual two-way interactions in Models 9-11 (Trial Type x Population: $p = .17$; Condition x Population: $p = 1.00$; Trial Type x Condition: $p = .49$), did not provide a significantly better fit compared to Model 4. Model 12 included the included the Trial Type x Population x Condition interaction ($p = .51$) but did not improve fit compared to Model 4. Therefore, the best fitting model was Model 4, containing a fixed effect of Trial Type.

Number of looks to the target stimuli

We began with a baseline model containing by-participant random intercepts. The individual addition of fixed effects of Population (Model 2; $\chi^2 = 4.89, p =.027$), Condition (Model 3; $\chi^2 = 11.10, p <.001$), and Trial Type (Model 4; $\chi^2 = 9.77, p = .002$), all yielded a significant improvement in fit over the baseline model. Model 5 included fixed effects of Trial Type + Population, which significantly improved fit compared to Model 2 ($\chi^2 = 9.75, p = .002$) and Model 4 ($\chi^2 = 4.86, p = .027$), and yielded a borderline significantly better fit than Model 3 ($\chi^2 = 3.54, p = .060$). Model 6 included fixed effects of Population + Condition, and significantly improved fit compared to Model 2 ($\chi^2 = 11.09, p <.001$), Model 3 ($\chi^2 = 4.89, p = .027$) and Model 4 ($\chi^2 = 6.21, p = .013$). Model 7 included fixed effects of Trial Type + Condition, and significantly improved fit compared to Model 2 ($\chi^2 = 16.20, p <.001$), Model

3 (χ^2 = 9.99, p = .002), and Model 4 (χ^2 = 11.31, p <.001). Model 8 included fixed effects of Trial Type + Population + Condition, and significantly improved fit compared to Model 5 (χ^2 = 11.31, p <.001), Model 6 (χ^2 = 9.96, p = .002), and Model 7 (χ^2 = 4.86, p =.027). Model 9 included the interaction between Trial Type x Population plus a fixed effect of Condition (p = .85), Model 10 included the Condition x Population interaction plus a fixed effect of Trial Type (p = .26), and Model 11 included the Trial Type x Population interaction plus a fixed effect of Population (p = .11), but none of these models yielded a significantly better fit in comparison to Model 8. Model 12 included the Trial Type x Population x Condition interaction (p = .40) but did not significantly improve fit compared to Model 8. Therefore, the best fitting model was Model 8, containing fixed effects of Trial Type + Population + Condition.

Longest look novel or known (Longest Look Novel)

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. In Models 2 and 3, adding Population (p = .53) and Condition (p = .19) as fixed effects did not provide a significantly better fit than the baseline model. Model 4 included the fixed effect of Trial Type (χ^2 = 33.55, p <.001), which yielded a significantly better fit for the observed data than the baseline model. Inclusion of multiple fixed effects in Model 5 (Condition + Population, p = 1.00), Model 6 (Trial Type + Population, p = .53), Model 7 (Trial Type + Condition, p = .25) and Model 8 (Trial Type + Population + Condition, p = .42) did not improve fit compared to Model 4. Inclusion of individual two-way interactions in Models 9-11 (Trial Type x Population: p = .71; Condition x Population: p = 1.00; Trial Type x Condition: p = .43), did not provide a significantly better fit compared to Model 4. Model 12 included the Trial Type x Population x Condition interaction (p = .76) but was not significantly better fitting than Model 4. Therefore, the best fitting model was Model 4, including Trial Type as a fixed effect.

5-minute retention

Proportion of time spent looking at the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts. In Models 2 and 3, adding individual fixed effects of Population (p = .31) and Condition (p = .33) did not yield a significantly better fit than the baseline model. Model 4 included fixed effects of Population + Condition (p = .37) but did not significantly improve fit compared to the baseline model. Model 5 included the Condition x Population (p = .27) interaction but was not significantly better fitting than baseline model. The inclusion of fixed effects did not improve predictive power.

Number of looks to the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Adding a fixed effect of Population in Model 2 (p = .21) or Condition in Model 3 (p = .10) did not significantly improve fit compared to the baseline model. Model 4 included fixed effects of Condition + Population (p = .12) but was not significantly better fitting than the baseline model. Model 5 included the Condition x Population interaction and was borderline significantly better fitting than the baseline model ($\chi^2 = 7.53$, $p = .057$). As such, Model 5 containing the Condition x Population interaction was determined to be the best fitting model, although should be interpreted with caution as the model comparison only approached significance.

Longest look target or foil

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Adding fixed effects of Population (Model 2; $p = .57$) and Condition (Model 3; $p = 1.00$) did not yield a significantly better fit compared to the baseline model. Model 4 contained a fixed effect of Condition + Population but was not significantly better fitting than the baseline model ($p = .81$). Model 5 included the Condition x Population interaction but was not significantly better fitting than the baseline model ($p = .55$). Therefore, inclusion of fixed effects did not improve predictive power.

24-hour retention

Proportion of time spent looking at the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Model 2 included the fixed effect of Population ($p = .66$), and Model 3 included the fixed effect of Condition ($p = .14$), but neither provided a significantly better fit than the baseline model. Model 4 included a fixed effect of Condition + Population ($p = .29$) but did not significantly improve fit compared to the baseline model. Model 5 included the Condition x Population interaction ($p = .27$), which was not significantly better fitting than the baseline model. Therefore, inclusion of fixed effects did not improve predictive power.

Number of looks to the target stimuli

We began with a baseline model containing a by-participant random intercept. Including Population as included as a fixed effect in Model 2 yielded a significant improvement in fit over the baseline model ($\chi^2 = 4.88$, $p = .027$), but including a fixed effect of Condition in Model 3 did not ($p = .37$). Model 4 included fixed effects of Condition + Population ($p = .36$) but was not significantly better fitting than Model 2. Model 5 included the Condition x Population interaction ($p = .09$), which was not significantly better fitting than Model 2. Therefore, the final model was Model 2, including a fixed effect of Population.

Longest look target or foil

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. Model 2 contained a fixed effect of Population ($p = .47$), and Model 3 included a fixed effect of Condition ($p = .28$), but neither models were significantly better fitting than the baseline model. Model 4 included a fixed effect of Condition + Population ($p = .41$), which was not significantly better fitting than the baseline model. Model 5 included the Condition x Population interaction ($p = .26$), which was not significantly better fitting than the baseline model. The inclusion of fixed effects did not improve predictive power.

Looking time predicting accuracy

The following analyses examined whether children's referent selection accuracy was influenced by a different looking behaviour, and whether the effects of those behaviours differed across populations and conditions (by testing whether the inclusion of x population and x condition interaction terms significantly improved model fit). All models were conducted using the glmer function from the lme4 package in R (Bates et al., 2015).

Referent Selection

Proportion of time spent looking at the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. In Model 2, Proportion to Target was included as a fixed effect ($\chi^2 = 199.32$, $p < .001$), which was significantly better fitting than the baseline model. Model 3 included the Proportion to Target x Trial Type interaction, which was not significantly better fitting than Model 2 ($p = .95$). Model 4 containing the Proportion to Target x Condition did not yield a significantly better fit compared to Model 2 ($p = .89$). Model 5, including the Proportion to Target x Population interaction, yielded a significantly better fit than Model 2 ($\chi^2 = 9.76$, $p = .008$), but did not contain a significant interaction effect ($z = 0.98$, $p = .33$). Model 6 included the Proportion to Target x Population x Condition interaction, which afforded an approaching significantly better fit than Model 2 ($\chi^2 = 12.09$, $p = .060$), but did not contain a significant interaction effect ($z = -1.24$, $p = .21$). Neither Model 7 (Proportion to Target x Population x Trial Type: $p = .11$), or Model 8 (Proportion to Target x Condition x Trial Type: $p = 1.00$), afforded a significantly better fit than Model 2. Therefore, the final best fitting model for the observed data was determined to be Model 2, including a fixed effect of Proportion to Target.

Number of looks to the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. In Model 2, Number Looks to Target was included as a fixed effect, which was significantly better fitting than the baseline model ($\chi^2 = 34.54$, $p < .001$). Model 3 included the Number Looks to Target x Trial Type interaction, and yielded a significantly better fit than Model 2 ($\chi^2 = 7.65$, $p = .022$), but the two-way interaction effect was not significant ($z = 0.93$, $p = .35$). Model 4, containing the Number Looks to Target x Condition interaction ($p = .12$) did not yield a significantly better fit compared to Model 2. Model 5, including the Number Looks to Target x Population interaction, which was significantly better fitting than Model 2 ($\chi^2 = 21.57$, $p < .001$) and contained a significant two-way interaction effect ($z = -3.30$, $p < .001$). Model 6 included the Number Looks to Target x Population x Trial Type interaction, which was borderline significantly better fitting than Model 5 ($\chi^2 = 8.67$, $p = .070$) but did not contain a significant three-way interaction effect ($z = 0.60$, $p = .55$). Model 7 included the Number Looks to Target x Population x Condition interaction, which was not significantly better fitting than Model 5 ($p = .54$). Model 8 included the Number Looks to Target x Condition x Trial Type interaction but was not significantly better fitting than Model 5 ($p = 1.00$). Therefore, Model 5, including the Number Looks to Target x Population interaction was established as the final model as it remained consistently better fitting than more complex models.

Longest look novel or known

We began with a baseline model containing by-participant and by-word random intercepts, with Condition x Trial Type slopes per participant. In Model 2, Longest Look Novel was included as a fixed effect, which was not significantly better fitting than the baseline model (*p* = .11). Model 3 included the Longest Look Novel x Trial Type interaction (χ^2 = 76.21, *p* <.001), which was significantly better fitting than the baseline model, and had a significant two-way interaction effect (*z* = 7.80, *p* <.001). Inclusion of the two-way Longest Look Novel x Condition interaction in Model 4 did not provide a significantly better fit than Model 3 (*p* = .33). Model 5 included the Longest Look Novel x Population interaction but was not significantly better fitting than Model 3 (*p* = .08). Models 6-8 included three-way interactions (Longest Look Novel x Condition x Trial Type: *p* = .24; Longest Look Novel x Trial Type x Population: *p* = .19; Longest Look Novel x Population x Condition: *p* = 1.00), but also did not provide a significantly better fit than Model 3. Overall, the final best fitting model for the observed data was Model 3, including the Longest Look Novel x Trial Type interaction.

5-minute retention

Proportion of time spent looking at the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Proportion to Target was included as a fixed effect (χ^2 = 332.98, *p* <.001), which was significantly better fitting than the baseline model. Model 3 included the Proportion to Target x Condition interaction (*p* = .83) but was not significantly better fitting than Model 2. Model 4 included the Proportion to Target x Population interaction (*p* = .12) but was not significantly better fitting than Model 2. Model 5 included the Proportion to Target x Population x Condition interaction (*p* = .21) but was not significantly better fitting than Model 2. Therefore, the final best fitting model for the observed data was determined to be Model 2, including the fixed effect of Proportion to Target.

Number of looks to the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Number Looks to Target was included as a fixed effect (χ^2 = 69.61, *p* <.001), which was significantly better fitting than the baseline model. Inclusion of the Number Looks to Target x Population interaction in Model 3 approached a significant improvement in fit over Model 2 (χ^2 = 5.56, *p* = .062), and contained a borderline significant interaction effect (*z* = -1.84, *p* = .065). Model 4 contained the Number Looks to Target x Condition interaction (*p* = .15) but did not yield a significantly better fit than Model 2. Model 5 included the Number Looks to Target x Population x Condition interaction but was not significantly better fitting than Model 2 (*p* = .10) or Model 3 (*p* = .28). Therefore, the final best fitting model for the observed data was determined to be Model 3, including the Number Looks to Target x Population interaction. However, since this model approached significance, we must be cautious when interpreting these results.

Longest look target or foil

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Longest Look Target was included as a fixed effect (χ^2 = 192.71, *p* <.001), which was significantly better fitting than the baseline model. Model 3 included the Longest Look Target x Condition interaction (*p* = 1.00), and Model 4 included the Longest Look Target x Population interaction (*p* = .99), but neither were significantly better fitting than Model 2. Model 5 included the Longest Look Target x Population x Condition interaction (*p* = .89) but was not significantly better fitting than Model 2. Overall, Model 2, including Longest Look Target as a fixed effect, was the best fit to the observed data.

Proportion of time spent looking at the target stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Proportion to Target was included as a fixed effect (χ^2 = 2.94, *p* = .086), which approached a significantly better fit than the baseline model, with an approaching significant fixed effect (*z* = 1.79, *p* = .074). Model 3 included the Proportion to Target x Condition interaction but was not significantly better fitting than the baseline model (*p* = .37) or Model 2 (*p* = .89). Model 4 included the Proportion to Target x Population interaction but was not significantly better fitting than the baseline model (*p* = .33) or Model 2 (*p* = .79). Model 5 included the Proportion to Target x Population x Condition interaction but was not significantly better fitting than the baseline model (*p* = .42) or Model 2 (*p* = .65). Model 2 (Proportion to Target) was taken to be the final model. However, the model comparison and interaction effect only approached significance, so we must interpret this result with caution.

Number of looks to the target stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts and Condition slopes per participant. In Model 2, Number Looks to Target was included as a fixed effect, which was borderline significantly better fitting than the baseline model (χ^2 = 3.14, *p* = .077) and an approaching significant interaction effect (*z* = 1.85, *p* = .064). Model 3 included the Number Looks to Target x Condition interaction but was not significantly better fitting than the baseline model (*p* = .24) or Model 2 (*p* = .57). Model 4 included the Number Looks to Target x Population interaction but was not significantly better fitting than the baseline model (*p* = .24) or Model 2 (*p* = .58). Model 5 included the Number Looks to Target x Population x Condition interaction which was not significantly better fitting than the baseline model (*p* = .25) or Model 2 (*p* = .43). Therefore, the final best fitting model for the observed data was determined to be Model 2 (Number Looks to Target). However, the model comparison and interaction effect only approached significance, so we must interpret this result with caution.

Longest look to target or foil stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Longest Look Target was included as a fixed effect (*p* = .73), which was not significantly better fitting than the baseline model. Model 3 included the Longest Look Target x Condition interaction (*p* = .96), and Model 4 included the Longest Look Target x Population interaction (*p* = .92), but neither were significantly better fitting than the baseline model. Model 5 included the Longest Look

Target x Population x Condition interaction (*p* = .72) but was not significantly better fitting than the baseline model. The inclusion of fixed effects did not improve predictive power.

24-hour retention

Proportion of time spent looking at the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Proportion to Target was included as a fixed effect (χ^2 = 337.08, *p* <.001), which was significantly better fitting than the baseline model. Model 3 included the Proportion to Target x Condition interaction (*p* = .99), which was not significantly better fitting than Model 2. Model 4 included the Proportion to Target x Population interaction (*p* = .81), which was not significantly better fitting than Model 2. Model 5 included the Proportion to Target x Population x Condition interaction (*p* = .88), which was not significantly better fitting than Model 2. Therefore, Model 2 was retained as the final best fitting model, containing a fixed effect of Proportion to Target.

Number of looks to the target stimuli

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Number Looks to Target was included as a fixed effect (χ^2 = 74.47, *p* <.001), which was significantly better fitting than the baseline model. Model 3 included the Number Looks to Target x Condition interaction (*p* = .55) but was not significantly better fitting than Model 2. Model 4 included the Number Looks to Target x Population interaction (*p* = .22), which was not significantly better fitting than Model 2. Model 5 included the Number Looks to Target x Population x Condition interaction (*p* = .33) but was not significantly better fitting than Model 2. Therefore, the final best fitting model for the observed data was determined to be Model 2, including Number Looks to Target as a fixed effect.

Longest look target or foil

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Longest Look Target was included as a fixed effect (χ^2 = 203.90, *p* <.001), which was significantly better fitting than the baseline model. Model 3 included the Longest Look Target x Condition interaction (*p* = .53), which was not significantly better fitting than Model 2. Model 4 included the Longest Look Target x Population interaction (*p* = .34), which was not significantly better fitting than Model 2. Model 5 included the Longest Look Target x Population x Condition interaction (*p* = .45), which was not significantly better fitting than Model 2. Overall, Model 2 containing a fixed effect of Longest Look Target

Proportion of time spent looking at the target stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts, with a random slope of Condition per participant. In Model 2, Proportion to Target was included as a fixed effect, which approached a significantly better fit than the baseline model (χ^2 = 3.02, *p* = .082) and contained a borderline significant fixed effect (z = 1.79, *p* = .073). Model 3 included the Proportion to Target x Condition interaction which was not significantly better fitting than the baseline model (*p* = .23) or Model 2 (*p* = .52). Model 4 included the Proportion to Target x Population interaction, which was not significantly better fitting than the baseline model (*p* = .37) or Model 2 (*p* = .93). Model 5 included the Proportion to Target x Population x Condition interaction, which was not significantly better fitting than the baseline model (*p* = .40) or Model 2 (*p* = .64). Overall, Model 2, including a fixed effect of Proportion to Target was the best fitting model, although caution must be taken given the approaching significant *p* values.

Number of looks to the target stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts and Condition slopes per participant. In Model 2, Number Looks to Target was included as a fixed effect (*p* = .28) but was not significantly better fitting than the baseline model. Model 3 included the Number Looks to Target x Condition interaction (*p* = .58) but was not significantly better fitting than the baseline model. Model 4 included the Number Looks to Target x Population interaction (*p* = .69), which was not significantly better fitting than the baseline model. Model 5 included the Number Looks to Target x Population x Condition interaction (*p* = .58) but was not significantly better fitting than the baseline model. The inclusion of fixed effects did not improve predictive power.

Longest look to target or foil stimuli during referent selection

We began with a baseline model containing by-participant and by-word random intercepts and Condition slopes per participant. In Model 2, Longest Look Target was included as a fixed effect (*p* = .13), which was not significantly better fitting than the baseline model. Model 3 included the Longest Look Target x Condition interaction (*p* = .36), which was not significantly better fitting than the baseline model. Model 4 included the Longest Look Target x Population interaction (*p* = .35), which was not significantly better fitting than the baseline model. Model 5 included the Longest Look Target x Population x Condition interaction (*p* = .44), which was not significantly better fitting than the baseline model. The inclusion of fixed effects did not improve predictive power.

Appendix E: Participation of Children Across Multiple Experiments

Note: the order in which children experienced the studies was counterbalanced. Different children experienced different studies in their first, second and third exposures (where relevant). Auditory and visual stimuli used in the experimental stages of each study did not overlap.

Participation of autistic children across multiple experiments

Autistic participant	Study 1	Study 2	Study 3
1	✓	✓	✓
2	✓	✓	✓
3	✓	✓	✓
4	✗	✓	✓
5	✓	✓	✓
6	✗	✗	✓
7	✓	✓	✓
8	✓	✓	✓
9	✓	✓	✗
10	✓	✓	✓
11	✓	✓	✗
12	✓	✓	✓
13	✓	✓	✓
14	✓	✓	✗
15	✓	✗	✓
16	✓	✗	✓
17	✓	✓	✓
18	✗	✗	✓
19	✗	✗	✓
20	✗	✗	✓
21	✓	✗	✗
22	✓	✗	✗
23	✗	✓	✗
24	✓	✗	✗

Participation of neurotypical children across multiple experiments

Neurotypical participant	**Study 1**	**Study 2**	**Study 3**
1	✗	✓	✓
2	✓	✗	✓
3	✗	✓	✗
4	✓	✓	✓
5	✓	✗	✗
6	✗	✓	✗
7	✓	✓	✓
8	✓	✓	✓
9	✓	✗	✓
10	✗	✓	✓
11	✓	✓	✓
12	✓	✓	✓
13	✓	✗	✓
14	✓	✓	✓
15	✓	✗	✓
16	✓	✓	✓
17	✓	✓	✓
18	✓	✓	✓
19	✓	✗	✗
20	✓	✗	✗
21	✓	✓	✗
22	✓	✗	✗
23	✗	✓	✗
24	✗	✓	✓
25	✓	✗	✗
26	✗	✗	✓