Introduction

Successful word learning requires children to pay attention to corresponding auditory and visual input during naming events (Pereira et al., 2014; Zosh et al., 2013). However, autism is characterized by Restricted and Repetitive Patterns of Behaviour, Interests, or Activities (RRBIs) which can impact attentional mechanisms and cause perseverative fixation on specific stimuli (American Psychiatric Association, 2022). RRBIs may prevent autistic children from paying attention to the right audio-visual stimuli at the right times, reducing their likelihood of successfully disambiguating and retaining novel word meanings (Venker et al., 2018, 2022). Yet, when acquiring vocabulary associated with stimuli that align with their interests, autistic children's heightened attentional focus may facilitate their word learning (Akechi et al., 2013; Parish-Morris et al., 2007; Rothwell et al., 2024). Existing evidence indicates that word learning is influenced by visual attention (Akechi et al., 2011; Axelsson et al., 2012; Bion et al., 2013; Hilton et al., 2019; Hollich et al., 2000; Venker et al., 2022; Yu & Smith, 2011), but research has yet to determine how neurotypical and autistic children's attention to semantic categories of differing interest levels affects novel word identification and retention. Here, we examine how autistic and neurotypical children's visual attention differs during naming events associated with stimuli that do, and do not, correspond with pre-existing interests, and how these differences influence their referent selection and retention accuracy.

Word learning is a crucial developmental milestone that underpins children's language acquisition (Patael & Diesendruck, 2008). To learn a new word, children must disambiguate its intended meaning (*referent selection*) and store a word-referent representation in memory for later retrieval (*retention*; Gleitman, 1990). In natural learning environments, each novel word can have many potential referents (Quine, 1960). Children must therefore overcome the challenge of 'referential ambiguity' by narrowing their attention to a single target and its corresponding label (Markman & Wachtel, 1988). The process of referent selection has been demonstrated very early on in neurotypical development, with research suggesting that children as young as six months can utilise statistical learning mechanisms to correctly identify and encode novel word-referent mappings (Bergelson & Swingley, 2012). By 24 months, accurate referent selection is facilitated by children applying heuristics that help them to rapidly decipher word-referent mappings in the context of naming events ('fast mapping'; Carey & Bartlett, 1978). Such heuristics include the mutual exclusivity principle (the assumption that each referent has only one label; Monaghan & Mattock, 2012), the novel name name-less category principle (the assumption that a novel label refers to novel stimuli, rather than familiar stimuli; Mervis & Bertrand, 1994), and attentional preferences for unfamiliar stimuli (Mather & Plunkett, 2012). Following this, retention involves integration of novel words into the lexicon, which is constrained by external inputs and cues during learning events (Kucker et al., 2020).

Crucially, neurotypical children's word learning is intrinsically related to visual attention. Statistical learning is initially underpinned by broadly distributed shifts of attention between distractor and target objects, with Friedrich and Friederici (2011) suggesting that infants can exclude non-target competitors and map correct word-referent associations from 6 months. Yu and Smith (2011) showed that as neurotypical infants begin to learn correct word-referent associations in the first year of life, looking patterns become more focused and systematic. Neurotypical children's visual attention can also be influenced by specific properties of stimuli. Hollich et al. (2000) demonstrated that 12-month-olds failed a referent selection task when relatively dull objects were labelled but succeeded when perceptually salient stimuli were labelled (see Taxitari et al., 2020 for similar results). Additionally, studies examining attentional processes during word learning indicate how neurotypical children's looking behaviour during referent selection affects subsequent retention (Hilton & Westermann, 2017). Bion et al. (2013) found that novel word retention accuracy increased when neurotypical 18–30-month-olds looked proportionally longer at a novel target after labelling. Likewise, Hilton et al. (2019) found that 20- and 26-month-olds who showed increased attention to novel objects during labelling achieved more accurate retention.

Axelsson et al. (2012) influenced neurotypical 2-year-olds' attention towards targets during referent selection by increasing their salience (e.g. via illumination) relative to competitors. This manipulation increased retention accuracy for word–object mappings after 5 minutes, demonstrating that heightened attentional focus on targets during labelling facilitates subsequent retention. Together, these findings indicate that increasing visual attention towards targets through a variety of means enhances neurotypical children's referent selection and retention accuracy.

Increasing evidence suggests that autistic and neurotypical children may not differ in terms of how they disambiguate and retain word meanings. While early studies suggested that autistic children struggle to interpret social pragmatic cues to inform word-referent mappings (Baron-Cohen et al., 1997; Preissler & Carey, 2005), more recent research shows that autistic children can utilise social and non-social word learning cues in a similar manner to neurotypically developing children (de Marchena et al., 2011; Luyster & Lord, 2009). A growing number of studies also demonstrate that autistic children can retain novel word meanings as accurately as neurotypical peers matched on receptive vocabulary (Carter & Hartley, 2021; Hartley et al., 2019, 2020; Rothwell et al., 2024). Yet, the factors that determine whether autistic children successfully disambiguate and retain labels following naming events are largely unknown. It is possible that differences in accuracy across children or trials may be attributable to their patterns of attention when receiving audio-visual input.

During fast mapping, differences in visual attention displayed by autistic individuals may affect the accuracy and/or strength of newly encoded word meanings (e.g., Akechi et al., 2011). Prior studies suggest that autistic children may allocate their attention to stimuli less flexibly when acquiring new vocabulary, resulting in the loss of important informants from the environment (Tenenbaum et al., 2017). Autistic children can also experience "sticky" attentional fixations, defined by a prolonged latency to engage or disengage focus (Sacrey et al., 2014). As such, autistic children's attention may be captured by salient perceptual features to an atypical degree (Hartley & Allen, 2014). Venker et al. (2022) reported that autistic children's novel word referent selection was disrupted more than neurotypical children's when stimuli in a referent selection task were of high perceptual salience. Similarly, previous looking time studies have demonstrated that autistic children are slower, or unable, to disengage their attention from irrelevant stimuli (Elsabbagh et al., 2013). These findings indicate that autistic children's word learning may be impacted by attentional allocation; failure to attend to informative referential cues, or focusing on incorrect target referents during naming events, may result in spurious word-referent associations.

It is important to note that studies investigating autistic children's word learning via looking measures often interpret differences in fixation duration towards targets as evidence for differences in accuracy. However, differences in visual attention and differences in learning accuracy may not necessarily be congruent. Venker et al. (2021) investigated how visual allocation differs for autistic and neurotypical children when perceptual salience competes with linguistic information. They found that both groups of children recognised words, although competing perceptual salience significantly decreased autistic children's attention to targets. Conversely, in Akechi et al. (2013), autistic children often attended to a speaker's face and followed gaze as frequently as neurotypical children, but their referent selection accuracy was significantly lower. Therefore, to truly understand the influence of visual attention on autistic children's vocabulary acquisition, it is necessary to investigate how looking behaviour and learning accuracy inter-relate, and how children's looking during referent selection influences their retention. Here, we investigate the possibility that differences in visual attention during referent selection may reduce the quality or quantity of input feeding into mechanisms involved in encoding and retaining word-referent associations, resulting in the formation of fragile - or potentially inaccurate - memory representations that are vulnerable to rapid decay.

One factor that may influence autistic children's attention to stimuli, and subsequent retention of corresponding labels, is their semantic interests (Ackermann et al., 2020). To our knowledge, Rothwell et al. (2024) is the only study to investigate how autistic children's pre-existing interests relating to particular semantic categories influences their ability to learn new vocabulary. Autistic children with delayed language development and neurotypical children matched on receptive vocabulary completed a pair of referent selection tasks in which the novel target stimuli were high-interest (unfamiliar animals) or neutral-interest (unfamiliar objects). Retention was tested after delays of 5 minutes and 24 hours. A touch screen computer measured accuracy and response times at each word learning stage. The results showed that autistic children took significantly longer to correctly identify stimuli, particularly in the animal condition. However, autistic children demonstrated superior 5-minute retention accuracy of novel animal names compared to novel objects and, surprisingly, outperformed neurotypical children across conditions at 24-hour retention. These findings suggest that the autistic children displayed a speed-accuracy trade off: looking longer or more frequently at stimuli during referent selection may have facilitated their encoding of stronger word-referent associations, potentially reflecting an important relationship between low level visual attention and longer-term word learning accuracy.

For the first time, the present study investigated how autistic and neurotypical children's attention differs when learning words associated with stimuli of contrasting interest levels (animals vs. objects) and elucidates predictive relationships between visual attention and word learning accuracy. Two measures of visual attention were collected concurrently to the touch screen accuracy and response time data reported in Rothwell et al. (2024), from the same autistic and neurotypical participants. The visual attention measures presented in the present paper are proportion of time spent looking towards the target and number of looks towards the target. At referent selection, we expected autistic children to look longer towards unfamiliar objects regardless of whether they were intended targets – particularly in the animal condition – due to difficulties disengaging attention from salient stimuli (Sacrey et al., 2014). We also anticipated that autistic children might make more frequent looks to target stimuli due to greater curiosity and/or longer processing times required to generate correct responses (Venker et al., 2018). We expected that increased visual attention to targets would predict referent selection and retention accuracy across conditions, groups, and task stages, and that increased visual attention at referent selection would be associated with superior retention accuracy (Hilton et al., 2019). However, we were mindful of the possibility that between-population differences in visual attention may not necessarily translate to significant differences in word learning (Venker et al., 2021). Comparing autistic and neurotypical children's in-trial visual attention to stimuli, and how differences in visual attention influence referent selection and retention accuracy, is necessary to advance theoretical understanding of the mechanisms underpinning autistic children's vocabulary acquisition. Our results will have important implications for research methodology (by highlighting the extent to which looking truly reflects learning) and potentially signpost favourable conditions for interventions targeting receptive vocabulary development.

Method

The participants and methodology in the present study were the same as those reported in Rothwell et al. (2024).

Participants

Thirty-one children participated; 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 blinded for review (see Table 1). Autistic children were significantly older, $t(29) = -5.48$, $p < .001$, $d = 1.97$, than neurotypical children. All children had normal or corrected-to-normal colour vision and were monolingual native English speakers. All autistic participants were previously diagnosed by a qualified educational or clinical psychologist using standardised instruments (i.e. Autism Diagnostic Observation Scale and Autism Diagnostic Interview – Revised; Lord et al., 2002) and expert judgement.

Five participants were excluded from the study: one autistic participant who did not like animals, two neurotypical participants who could not complete the touch screen task, and two participants who did not complete both experimental conditions (one autistic and one neurotypical child).

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.

Standardised assessments

Childhood Autism Rating Scale Second Edition (CARS-2). Autism diagnoses were confirmed via the CARS-2 standard version, which was completed by children's classroom teachers, or the caregivers of eight neurotypical children who were tested in our Lab due to COVID-19 restrictions (Schopler et al., 2010). Higher raw scores on the CARS-2 indicate higher autistic traits. The CARS-2 has high internal consistency ($\alpha = .93$). Autistic children had significantly higher raw CARS-2 scores than the neurotypical children (autistic M = 34.70, SD = 10.23; neurotypical M = 16.78, SD = 2.56), $t(29) = -6.79$, $p < .001$, $d = 2.40$.

British Picture Vocabulary Scale Second Edition (BPVS-2). Children's receptive vocabulary was measured by the BPVS-2 (Dunn et al., 1997). Higher scores on the BPVS-2 indicate older age equivalents and therefore more advanced receptive vocabulary. The BPVS-2 has high internal consistency ($\alpha = .93$, split-half $r = .86$). Receptive vocabulary was used as our group matching criterion as it reflects children's ability to learn word-referent relationships (Bion et al., 2013). Raw scores were converted to receptive language age equivalents, which did not significantly differ between groups (autistic M age equivalent = 53.27 months, SD = 22.48; neurotypical M age equivalent = 60.31, SD = 27.44), $t(29) = 0.78$, $p = .44$.

Expressive Vocabulary Test Second Edition (EVT-2) & Mullen's Scales of Early Learning (MSEL). Children's expressive vocabulary was measured using form B of the EVT-

2 (Williams, 2007), or the expressive language module of the MSEL (Mullen, 1995) for children who scored below the baseline on the EVT-2. Higher scores on the MSEL and EVT-2 indicate older age equivalents and therefore more advanced expressive vocabulary. The internal consistency of the EVT-2 is excellent ($\alpha = .96$, *M* split-half $r = .93$). The split-half internal consistency median for the MSEL expressive language scale is also high ($r = .82$). Raw scores were converted to expressive language age equivalents, which did not significantly differ between groups (autistic *M* age equivalent = 48.47 months, *SD* = 27.70; neurotypical *M* age equivalent = 60.31 months, *SD* = 22.76), $t(29) = 1.30$, $p = .20$.

Leiter-3. Children's non-verbal intellectual abilities were measured using the Leiter-3 (Roid et al., 2013). Three autistic children did not complete the Leiter-3 due to school closures during the COVID-19 pandemic (autistic N = 12, neurotypical N = 16). Higher scores on the Leiter-3 demonstrate greater progress through the tasks, indicating increased intellectual abilities irrespective of age. The brief IQ version was used, which consists of four subscales, with good internal consistency: Figure Ground ($\alpha = .86$), Form Completion ($\alpha = .86$), Sequential Order ($\alpha = .95$), Classification and Analogies ($\alpha = .79$). Raw non-verbal IQ scores were transformed into validated scaled age-norms, and the neurotypical group's average age-norm was significantly higher than the autistic group's (autistic *M* = 77.67, *SD* = 11.73; neurotypical *M* = 101.38, *SD* = 7.84), $t(23) = 5.99$, $p < .001$, $d = 2.38$. Scaled IQ scores could not be calculated for three neurotypical children as they were below the age of three years. Raw scores on the Leiter-3 did not significantly differ between groups, suggesting that when age was not considered, their non-verbal cognitive abilities were similar at the time of testing (autistic *M* = 60.33, *SD* = 15.57; neurotypical *M* = 57.25, *SD* = 17.93), $t(26) = -0.48$, $p = .64$,

Conner's Teacher Rating Scale-15 (CTRS-15). To assess attentional behaviours, the CTRS-15 was completed by children's classroom teachers, or the caregivers of eight neurotypical children who were tested in our Lab due to COVID-19 restrictions (Pupura & Lonigan, 2009). Higher scores on the CTRS-15 indicate more extensive attentional difficulties. The CTRS-15 subscales significantly correlate with the previous CTRS scale which contains a larger set of items (CTRS-44; Gerhardstein et al., 2003), demonstrating psychometric similarity: Inattention (r = .92), Hyperactivity/Impulsivity (r = .94), Opposition (r = .96). Raw CTRS-15 scores did not significantly differ between groups (autistic M = 17.27, SD = 11.04; neurotypical M = 12.25, SD = 6.03), $t(29) = -1.58$, $p = .12$.

Repetitive Behaviour Questionnaire 2 (RBQ-2). Caregivers completed the RBQ-2 to assess the extent of children's restrictive and repetitive behaviours (Leekam et al., 2007). Higher scores on the RBQ-2 indicate increased restrictive and repetitive behaviours. This questionnaire has high internal consistency (α = .85). Autistic children's raw scores were significantly higher than neurotypical children's raw scores (autistic M = 43.87, SD = 8.37; neurotypical M = 27.00, SD = 5.80), $t(29) = -6.56$, p <.001, d = 2.34.

Animal interests questionnaire. To ensure that we recruited participants who liked animals (our high-interest stimuli), caregivers completed a questionnaire designed for this study assessing their children's animal interests (see Supplementary Materials). Higher scores on this measure indicate a greater interest in animals. The groups' raw scores did not significantly differ on this measure (autistic M = 23.93, SD = 5.55; neurotypical M = 23.31, SD = 2.80), $t(29) = -0.40$, $p = .69$.

(insert Table 1 here)

Materials

The study was administered via a touch screen computer running MATLAB. Audio stimuli for the word learning task included eight two-syllable novel words (manu, tanzer, boskot, virdex, toma, fiffin, chatten, modi). Visual stimuli were high-resolution colour photographs of 22 familiar objects, 4 unfamiliar objects, and 4 unfamiliar animals. Six familiar objects were presented during warm-up trials, and 16 familiar objects were displayed during referent selection trials across conditions. Novel words and novel object stimuli were selected from the NOUN database (Horst & Hout, 2016) and other academic sources. The names of the familiar objects were typically comprehensible by 16 months (Fenson et al., 1994) and were divided into two sets matched on mean comprehension age (13.5 months). Familiar objects in each set spanned a variety of non-animal semantic categories (e.g. vehicles, furniture) and sets were matched on frequency of semantic categories. Familiar objects within each set were divided into phonologically and visually distinct pairs and presented alongside an unfamiliar object or animal in referent selection trials (depending on condition). In every trial type, three pictures were presented side by side. Whilst detailed information about the study and methodology are reported here, for further information please see Rothwell et al. (2024).

Three cameras attached to the left, right, and centre of the computer recorded participants' visual attention and behaviour during the study. Recording was controlled by '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 computer 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 computer (1.2 megapixels), also recording at 30 frames-per-second. The red recording lights were obscured from participants using black tape to avoid distraction.

Procedure

The study took place in participants' own school, nursery, or blinded for review. The researcher administered the Leiter-3, BPVS-2, and EVT-2 or MSEL to participants over multiple sessions on different days. For analyses examining whether autistic and neurotypical children differed in their visual attention during referent selection and retention, the predictor variables were trial type (referent selection only, familiar/novel - within-subjects factor), condition (novel animals/novel objects - within-subjects factor) and population (neurotypical/autistic – between-subjects factor), and the outcome variables were looking time (proportion of looking) and looking frequency (number of looks). For analyses examining the relationship between visual attention and word learning accuracy, the predictor variables were looking time (proportion of looking), looking frequency (number of looks), condition (novel animals/novel objects - within-subjects factor) and population (neurotypical/autistic – between-subjects factor), and the outcome variables were referent selection and retention accuracy (correct/incorrect). The conditions were administered on different days and presentation order was counterbalanced. The word learning task was delivered via a customised touch screen computer and consisted of six stages: 1. Warm-up trials, 2. Referent selection trials, 3. Five-minute delay, 4. Retention trials, 5. 24-hour delay, 6. Retention trials (see Figure 1).

(insert Figure 1 here)

Warm up trials

At the beginning of the study, children completed three warm-up trials, where they were presented with images of three familiar objects in the left, middle, and right sections of the touch screen (order and location counterbalanced). Two seconds after the images appeared on the screen, 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, and the same instructions repeated up to six times until children provided a response. Responses were accepted only after the first label utterance, preventing children from selecting an image without hearing the requested label. If children responded incorrectly, the correct referent was highlighted by a green border and children could retry up to five times.

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

Eight referent selection trials directly followed the warm-up trials, which were in the same format except children did not receive feedback following their responses. In each condition, children viewed four sets of pictures (each containing one unfamiliar image and two familiar images) and learnt four novel words via fast mapping (Horst & Samuelson, 2008). Familiar stimuli were always objects, and novel stimuli were either animals (high-interest condition) or objects (neutral-interest condition). Each stimuli set was presented twice; on one trial the novel picture was requested, and on another trial a familiar picture was requested. Trial order was pseudo-randomised so that the same trial type (familiar name or novel name) was not presented on more than two trials sequentially. Stimuli positioning was pseudo-randomised across trials, ensuring that any target did not appear in the same location more than twice consecutively (left, middle, right).

5-minute delay

After completing the eight referent selection trials, children participated in a task unrelated to the experiment for five minutes to distract them from the stimuli (e.g. drawing, building with blocks).

Retention trials

After the five-minute delay, children completed one warm-up trial (as described above) to remind them of the task requirements. Eight retention trials followed (see Figure 1), with each of the four novel words serving as a target on two trials and a foil on four trials.

24-hour retention trials

After a 24-hour delay, children completed three warm-up trials to remind them how to complete the task. These were followed by another block of eight retention trials, which were presented in the same manner as the 5-minute retention trials, except 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 target stimuli on each trial. Children's visual fixations were coded frame-by-frame with a precision of 16.67 ms, 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. One-hundred-and-eighty-three videos were recorded, divided across referent selection, 5-minute retention, and 24-hour retention. Of these videos, 25% were reviewed by two independent coders. Inter-observer agreement was calculated per frame, yielding 98.08% for referent selection, 97.60% for 5-minute retention, and 97.23% for 24-hour retention. A custom MATLAB programme then calculated the primary dependent variables. These variables were calculated 233 ms after label onset to allow for saccade initiation latencies (Swingley, 2009).

We analysed two distinct measures of visual attention from our coded videos. ‘Proportion of looking towards the target’ was selected because this measure is commonly used in both the neurotypical and autism word learning literatures (e.g. Ackermann et al., 2020; Akechi et al., 2013) and is often interpreted as a measure of learning accuracy in the absence of an explicit behavioural response. It is calculated by dividing time spent looking towards the target by total time spent looking at all visible stimuli in the trial array. ‘Number of looks towards the target’ was also selected as autistic children are known to demonstrate greater exploration and detail-orientation towards stimuli that relate to their interests (e.g. Sasson et al., 2011), which could be interpreted as greater curiosity or information processing. The moment that a participant’s eyes directed towards the target was taken as the onset of a look, and the moment their eyes directed away from the target was taken as the offset of the look. Each individual look towards the target was counted.

Results

We examined whether autistic and neurotypical children differed in their visual attention during each stage of the word learning task (see Figure 2 for descriptive statistics). To elucidate the relationship between visual attention and word learning performance, we also investigated how variability in children’s in-trial visual attention predicted their response accuracy. Detailed analyses of referent selection and retention accuracy and response times are reported in Rothwell et al. (2024).

(insert Figure 2 here)

All models were conducted using the glmer and lmer functions from the lme4 package (Bates et al., 2015) using R Statistical Software (R Core Team, 2024). 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). Referent selection trial type was contrast coded as -0.5 (familiar) and 0.5 (novel). As dependent variables, trial-level accuracy at referent selection, 5-minute retention, and 24-hour retention were coded as 1 (correct) or 0 (incorrect). Proportion of time spent looking at the target stimuli on each trial was scored between 0 and 1. Number of looks to the target stimuli on each trial ranged from 0 to 14.

All models were built up sequentially, adding fixed effects individually 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 phases, or condition per participant for retention phases. As only final models are reported, please refer to the Supplementary Materials for full details of the model building sequences. If some models in a sequence were singular fitting or failed to converge, random effects were simplified until all models in the sequence successfully converged.

Referent selection

Linear mixed-effects models testing whether effects of population, condition, and trial type predicted variability in each visual attention measure during referent selection contained 496 data points.

Generalised linear mixed-effects models testing whether population, condition, and children's in-trial visual attention measures predicted their referent selection accuracy contained 491 data points, as five trials were excluded from autistic participants who provided ambiguous responses (they selected different stimuli simultaneously with both their head and their hand).

Proportion of time spent looking at the target

Proportion of time spent looking at the target 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 2). Neurotypical children looked significantly longer at targets than autistic children. The trial type x condition interaction was deconstructed by testing the effect of trial type on the animal and object conditions separately. Children looked significantly longer at the target on familiar trials than novel trials in the object condition (t = -5.45, p <.001), but proportional looking did not differ across trial types in the animal condition (t = -1.12, p = .26).

Referent selection accuracy was predicted by this visual attention measure (z = 7.59, p <.001; see Table 2). Across populations and conditions, as children's proportion of looking towards the target increased, so too did their referent selection accuracy.

(insert Table 2 here)

Number of looks towards the target

Number of looks towards the target stimuli was predicted by fixed effects of population (t = 3.20, p = .003) and trial type (t = 2.34, p = .026; see Table 3). Across conditions, autistic children made more looks towards the 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 visual attention measure x population interaction (z = -2.01, p = .044; see Table 3). This interaction was deconstructed by testing the effect of the visual attention measure 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).

(insert Table 3 here)

5-minute retention

Linear mixed-effects models testing whether effects of population and condition predicted variability in each visual attention measure at 5-minute retention contained 493 data points. Due to a technical error, three trials were removed for one neurotypical participant.

Generalised linear mixed-effects models testing whether population, condition, and children's in-trial visual attention measures at referent selection predicted their 5-minute retention accuracy contained 489 data points. Generalised linear mixed-effects models testing whether population, condition, and visual attention measures at 5-minute retention predicted 5-minute retention accuracy contained 488 data points. We excluded four trials due to a technical error, and four trials due to ambiguous responding.

Proportion of time spent looking at the target

Proportion of time spent looking at the target during 5-minute retention 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 visual attention measure (z = 11.21, p < .001; see Table 4). Across groups and conditions, children who looked proportionately longer at the target stimuli at 5-minute retention also responded more accurately at 5-minute retention.

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

(insert Table 4 here)

Number of looks towards the target

Number of looks towards the target during 5-minute retention was predicted by a population x condition interaction (t = 2.24, p = .033; see Table 5). 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 not the object condition (t = 1.30, p = .20).

Children's 5-minute retention accuracy was predicted by a visual attention measure x condition interaction (z = 2.60, p = .009; see Table 5). Across populations, children who looked more frequently towards the target stimuli during 5-minute retention 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 the target stimuli during referent selection responded with significantly greater accuracy at 5-minute retention (z = 2.20, p = .028; see Table 5).

(insert Table 5 here)

24-hour retention

Linear mixed-effects models testing whether effects of population and condition predicted variability in each visual attention measure at 24-hour retention contained 471 data points. Two autistic children in the animal condition and one neurotypical child in the object condition did not complete 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 population, condition, and children's in-trial visual attention measures at referent selection and 24-hour retention predicted their 24-hour retention accuracy contained 467 data points. We excluded five trials for autistic participants in the object condition due to non-completion (1), and ambiguity (4).

Proportion of time spent looking at the target

Proportion of time spent looking at the target during 24-hour retention 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 this visual attention measure ($z = 11.56$, $p < .001$; see Table 6). Across groups and conditions, as children's proportion of looking at the target stimuli increased during 24-hour retention, so did their 24-hour retention accuracy.

Children's 24-hour retention accuracy was predicted by a proportion of time spent looking at target stimuli during referent selection x population x condition interaction ($z = -2.04$, $p = .041$; see Table 6). This interaction was deconstructed by separately testing the effect of the visual attention measure 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 ($z = 2.72$, $p = .006$), but not the animal condition ($z = -0.83$, $p = .41$). However, this visual attention 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.

(insert Table 6 here)

Number of looks towards the target

Number of looks towards 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 7).

Children's 24-hour retention accuracy was predicted by a visual attention measure x population interaction ($z = -3.03$, $p = .002$; see Table 7). More frequent looks towards the target during 24-hour retention 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 the target stimuli during referent selection responded with significantly greater accuracy at 24-hour retention ($z = 3.11$, $p = .002$; see Table 7).

(insert Table 7 here)

Discussion

The present study investigated how autistic and neurotypical children's visual attention differed when learning words associated with high interest stimuli (animals) and low interest stimuli (objects) in Rothwell et al.'s (2024) task. We also tested whether differences in visual attention predicted variability in accuracy at referent selection, 5-minute retention, and 24-hour retention. Neurotypical children spent proportionally longer looking at target stimuli during referent selection than autistic children, and both populations looked longer at target stimuli during familiar trials than novel trials. However, autistic children looked more frequently towards targets than neurotypical children during all three word learning stages. At 5-minute retention, autistic children looked more frequently at targets in the animal condition than neurotypical children. Across groups and conditions, children's in-trial visual attention predicted accuracy at all three word learning stages. We also discovered that children's visual attention at referent selection predicted their 5-minute and 24-hour retention accuracy, suggesting that visual attention during encoding of new word-referent mappings directly influences the likelihood of memory consolidation (see Hilton et al., 2019 for similar results).

Children's visual attention during referent selection may be attributed to the differing cognitive demands associated with identifying targets for familiar and unfamiliar words (Preissler & Carey, 2005). It is possible that increased frequency of looks on novel trials reflected participants' need to check multiple stimuli to disambiguate the intended referent from competitor objects through successful implementation of word learning heuristics (de Marchena et al., 2011). Conversely, children may have 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. As our autistic sample was characterised by delayed language and intellectual development, these children may have looked more frequently towards target stimuli due to their need for greater input to support processing. While these differences in visual attention may not have had a detrimental impact on learning accuracy, they may explain why autistic children took significantly longer to generate correct responses during referent selection (i.e. a speed-accuracy trade off; see Rothwell et al., 2024).

Although neurotypical children looked proportionately longer towards targets than autistic children during referent selection, this difference did not persist across retention stages. Despite this difference in visual attention during initial encoding – which, in isolation, could be interpreted as evidence for reduced learning accuracy in autistic children – Rothwell et al.'s (2024) behavioural data show that the groups' referent selection accuracy did not significantly differ. These results demonstrate that, whilst autistic and neurotypical children's visual attention may vary during fast mapping, their likelihood of successfully identifying and retaining the meanings of novel words appears to be unaffected. Importantly, these findings indicate that population differences in word learning accuracy should not necessarily be inferred from contrasting profiles of visual attention. Indeed, autistic children's more frequent looks towards novel stimuli across both referent selection trial types in comparison to neurotypical children 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 (Rothwell et al., 2024). These findings also show that differences in time spent looking towards targets during referent selection do not necessarily persist for other word learning stages, highlighting the importance of studying word learning as a holistic system.

At 5-minute retention in the animal condition, Rothwell et al. (2024) showed that autistic children responded with significantly greater accuracy and outperformed vocabulary-matched neurotypical children. Moreover, after 24-hours, autistic children retained significantly more novel animal and object words than neurotypical children (Rothwell et al., 2024). Correspondingly, significant between-group conditional differences in visual attention emerged when testing children's retention of novel word meanings. Autistic children looked at animal targets significantly more frequently than neurotypical children at 5-minute retention and across conditions at 24-hour retention. As autistic children tend to process high-interest stimuli with greater focus and intensity (Elison et al., 2012), it may be that their interest in animals facilitated encoding of more robust word-referent representations that were less vulnerable to decay after 5-minutes. Overall, these patterns of results suggest that frequent looking may represent curiosity towards stimuli in autistic children. As predictive relationships were identified between both in-trial visual attention measures and response accuracy at 5-minute and 24-hour retention, it may be that group differences in visual attention reflected variability in memory consolidation.

Analysing relationships between visual attention and forced-choice accuracy represents an important methodological advancement that can generate 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, visual attention measures at each learning stage often contradicted one another and did not always align with direct accuracy measures. In terms of response accuracy, autistic children outperformed neurotypical children in the animal condition at 5-minute retention and across conditions at 24-hour retention (Rothwell et al., 2024). While population and condition differences in number of looks broadly complemented these between-group differences in accuracy, proportion of time spent looking at target stimuli did not significantly differ across populations or conditions at 5-minute or 24-hour retention. These findings suggest that proportional looking time and frequency of looks reflect distinct aspects of visual attention that vary in different ways between autistic and neurotypical populations and differ in how they relate to actual word learning outcomes. Of the two measures, our findings spotlight number of looks towards targets as a key measure – more frequent looks at referent selection predicted greater retention accuracy at 5-minutes and 24-hours, and population differences in this variable aligned with between-group differences in accuracy at both retention stages. Crucially, this study shows that differences in learning accuracy are not necessarily mirrored in visual attention, and we recommend caution when drawing conclusions about autistic children's word learning abilities exclusively from proportional looking times.

The current study also indicates how fast mapping and retention mechanisms are related within a unified word learning system. Whilst earlier studies with neurotypical children suggested that these word learning stages are distinct (e.g. Horst & Samuelson, 2008), our findings demonstrate a link between referent selection and overnight retention that is mediated by visual attention. Greater 24-hour retention accuracy was predicted by looking more frequently at targets during referent selection, suggesting that increased visual input supports consolidation of newly acquired word-referent associations into children's vocabularies (see McMurray et al., 2012 for similar findings). Additionally, differences in visual attention during referent selection may have contributed to differences between the groups' word learning accuracy. Autistic children's more frequent looks towards the target during encoding may have scaffolded their superior retention accuracy in comparison to neurotypical children. As such, interventions that increase visual fixations during learning could be employed to improve autistic children's vocabulary acquisition.

Naturally, we must consider the limitations of this study. As outlined in Rothwell et al. (2024), participant recruitment was hindered by the COVID-19 pandemic lockdown restrictions which were initiated mid-way through data collection. We therefore recommend that future studies combining behavioural and visual attention measures seek to replicate our findings with larger samples. As the study did not emulate a naturalistic environment, we cannot be certain that our samples' performance levels would be comparable under less controlled word learning conditions. Although autistic children's word learning may have benefitted from more frequent looks towards targets in our experiment, these extra looks might not be afforded in real world environments where processing times are restricted, and stimuli are presented more rapidly (Hartley et al., 2020). Thus, future research should explore autistic children's visual attention during word learning in a more demanding, naturalistic environment. Presenting children with more salient distractors, more complex and numerous stimuli, and faster paced learning environments would allow us to examine how autistic children's word learning is influenced by more naturalistic challenges.

It is also important to acknowledge that we did not directly assess children's interests in unfamiliar objects specifically. However, the unfamiliar objects were typical experimental stimuli that are often used in studies of this nature and were necessarily included to ensure comparability with previous research. As the familiar stimuli spanned a variety of non-animal categories that differed across trials (e.g. furniture, vehicles), we can be confident that pre-existing interests relating to these would not have systematically influenced children's responding. Moreover, observed differences in autistic children's attention to different semantic categories highlights the need for future studies to consider how participants' interests in stimuli may influence their word learning performance.

Overall, this study advances theoretical understanding of visual attention, and its relationship with accuracy, during novel word learning in autism and neurotypical development. Our results highlight that differences in visual attention do not necessarily reflect equivalent differences in word learning accuracy. Indeed, we revealed that autistic and neurotypical children matched on receptive vocabulary can differ in visual attention yet reach similar word learning outcomes (Rothwell et al., 2024). Our results therefore highlight the risk of drawing inaccurate conclusions about autistic children's learning from specific measures of looking time alone and show the importance of including multiple measures of learning outcomes and visual attention. Our study also discovered that increased visual attention during encoding at referent selection influences children's ability to retrieve novel word meanings at retention. This highlights the importance of designing word learning contexts that afford children optimal opportunities to attend to all vital information and thus improve chances of successful word learning. Our findings also emphasise the importance of studying word learning as a holistic system of inter-related stages.

References

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). https://doi.org/10.1111/desc.12915

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. https://doi.org/10.1016/j.rasd.2011.01.013

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. https://doi.org/10.1007/s10803-012-1555-3

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425787

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. https://doi.org/10.3389/fpsyg.2012.00088

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. https://doi.org/10.2307/1131924.

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

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. https://doi.org/10.1073/pnas.1113380109

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. https://doi.org/10.1016/j.cognition.2012.08.008

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

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. https://doi.org/10.1007/s10803-020-04771-2

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. https://doi.org/10.1016/j.cognition.2010.12.011.

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

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. https://doi.org/10.1016/j.rasd.2011.11.005

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. https://doi.org/10.1016/j.biopsych.2012.11.030

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. https://doi.org/10.2307/1166093.

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

Gerhardstein, R. R., Lonigan, J. L., Cukrowicz, K. C., & McGuffey, J.A. (2003). Factor structure of the Conners' Teacher Rating Scale–Short Form in a low-income preschool sample. *Journal of Psycho-educational Assessment, 21*, 223–243 https://doi.org/10.1177/073428290302100301

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

Hartley, C., & Allen, M. L. (2014). Generalisation of word-picture relations in children with autism and typically developing children**.** *Journal of Autism and Developmental Disorders, 44*, 2064-2071. https://doi.org/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. https://doi.org/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. https://doi.org/10.1016/j.cognition.2020.104265

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. https://doi.org/10.1016/j.jecp.2019.01.023

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. https://doi.org/10.1017/S030500091600057X

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. https://doi.org/10.1111/1540-5834.00090

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. *Behavior Research Methods*, *48*(4), 1393-1409. https://doi.org/10.3758/s13428-015-0647-3.

Horst, J. S., & Samuelson, L. K. (2008). Fast Mapping but Poor Retention by 24-Month-Old Infants. *Infancy, 13*(2), 128-157. https://doi.org/10.1080/15250000701795598

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. https://doi.org/10.1016/j.jecp.2019.104705

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. https://doi.org/10.1111/j.1469-7610.2007.01778.x

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

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

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

Mather, E., & Plunkett, K. (2012). The role of novelty in early word learning. *Cognitive science*, *36*(7), 1157-1177. https://doi.org/10.1111/j.1551-6709.2012.01239.x

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. https://doi.org/10.1037/a0029872

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

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

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

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. https://doi.org/10.1111/j.1467-8624.2007.01065.x.

Patael, S., & Diesendruck, G. (2008). Intentions help children learn meaningful rules. *Journal of Child Language, 35,* 221–234. https://doi.org/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. https://doi.org/10.3758/s13423-013-0466-4.

Preissler, M., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition, 97*(1), B13-B23. https://doi.org/10.1016/j.cognition.2005.01.008

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. https://doi.org/10.1080/15374410802698446.

Quine, W. V. O. (1960). *Word and object: An inquiry into the linguistic mechanisms of objective reference.* Cambridge, MA: MIT Press. https://doi.org/10.1068/p6639.

R Core Team. (2024). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Roid, G. H., Miller, L. J., Pomplun, M., & Koch, C. (2013). Leiter international performance scale (Leiter-3). Western Psychological Services, Los Angeles.

Rothwell, C., Westermann, G., & Hartley, C. (2024). How do Autistic and Neurotypical Children's Interests Influence their Accuracy During Novel Word Learning? *Journal of Autism and Developmental Disorders, 54*(9), 3301-3315. https://doi.org/10.1007/s10803-023-06066-8.

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. https://doi.org/10.1016/j.neubiorev. 2014.10.011.

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. https://doi.org/10.1007/s10803-010-1038-3

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.

Swingley, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of memory and language*, *60*(2), 252-269. https://doi.org/10.1016/j.jml.2008.11.003

Taxitari, L., Twomey, K. E., Westermann, G., & Mani, N. (2020). The Limits of Infants' Early Word Learning. *Language Learning and Development*, *16*(1), 1–21. https://doi.org/10.1080/15475441.2019.1670184

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. https://doi.org/10.1007/s10803-017-3098-0

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. https://doi.org/10.1002/aur.2038

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. https://doi.org/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. https://doi.org/10.1177/23969415221085476

Williams, K. T. (2007). Expressive Vocabulary Test. 2nd Edn. Minneapolis, MN: Pearson Assessments.

Yu, C., & Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental science*, *14*(2), 165-180. https://doi.org/10.1111/j.1467-7687.2010.00958.x

Zosh, J. M., Brinster, M., & Halberda, J. (2013). Optimal contrast: Competition between two referents improves word learning. *Applied Developmental Science, 17*(1), 20–28, https://doi.org/10.1080/10888691.2013.74842

Figure Captions

Figure 1. Examples of trial types in the word learning task

Figure 2. Mean visual attention measures for referent selection, 5-minute retention, and 24-hour retention, error bars show±1 SE

Figure 1 top

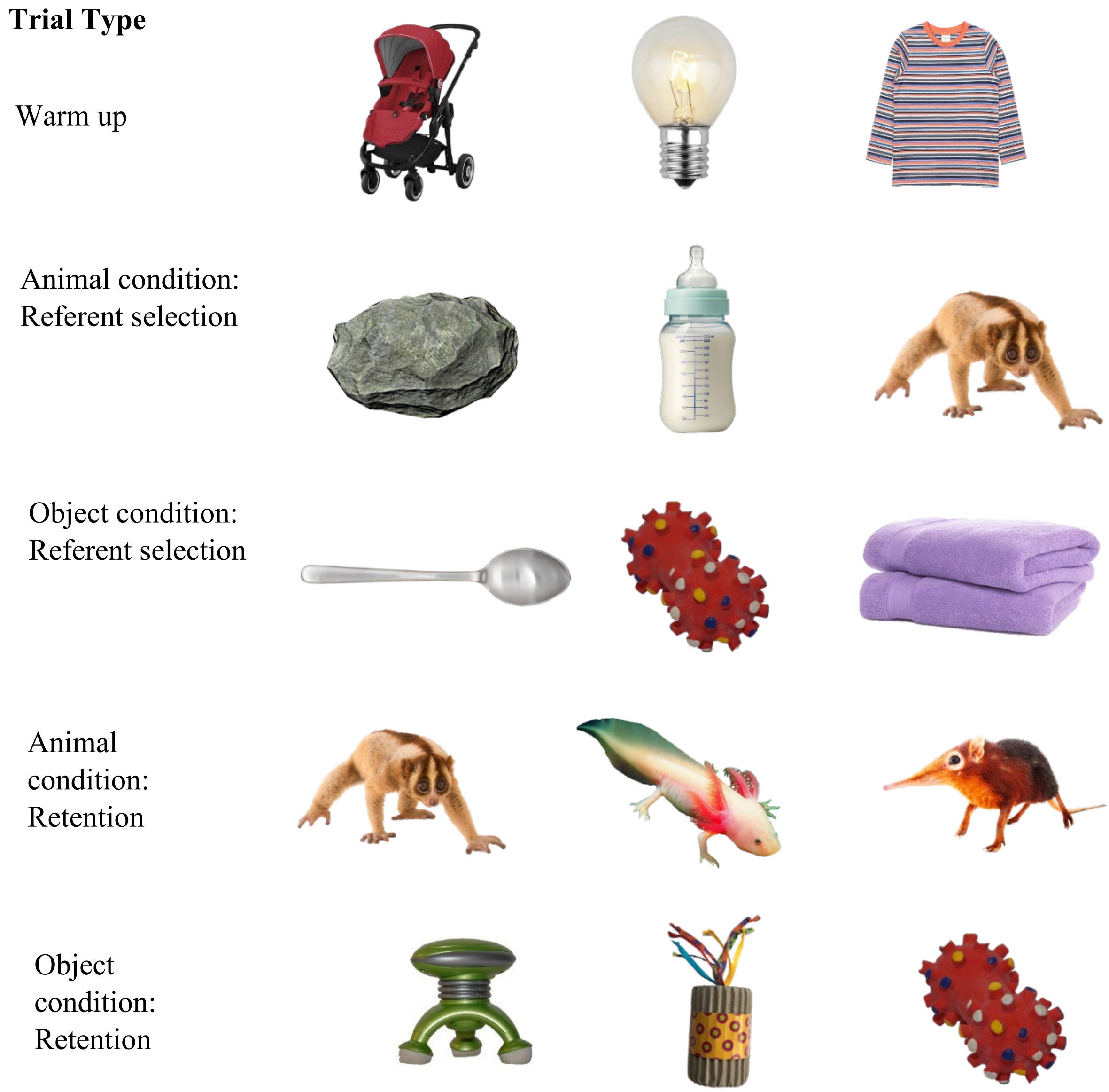
Trial Type
Warm up
Animal condition: Referent selection
Object condition: Referent selection
Animal condition: Retention
Object condition: Retention

Figure 2 top

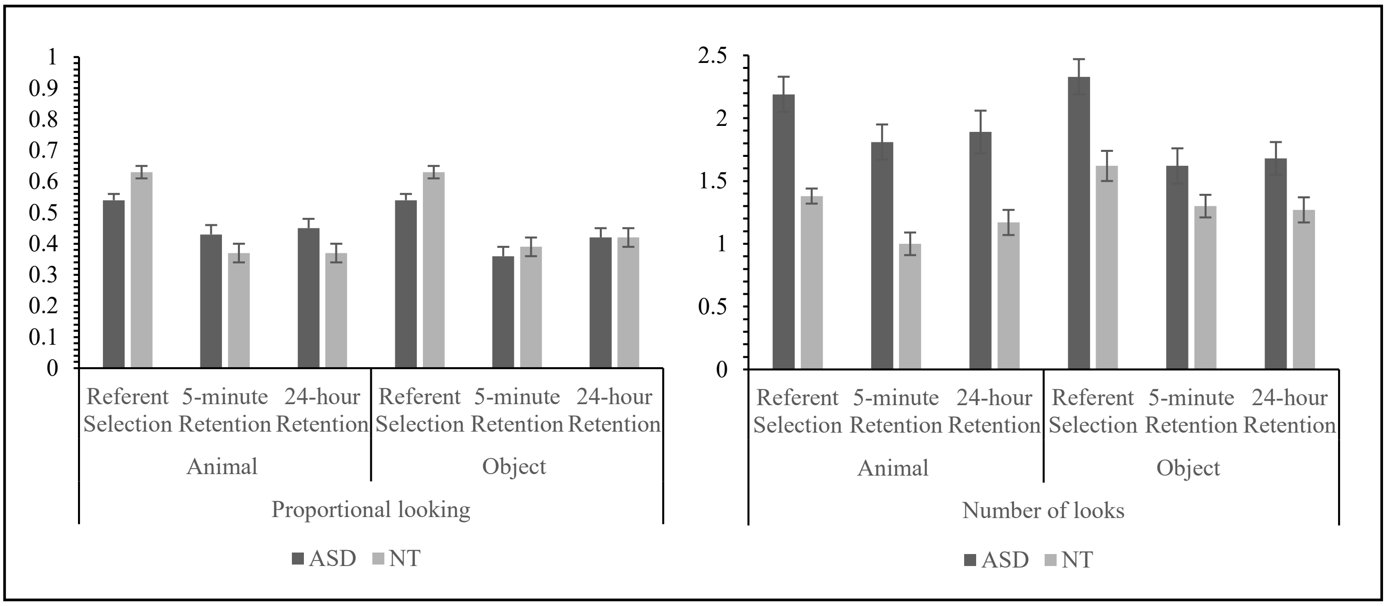
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Referent Selection
5-minute Retention
24-hour Retention
Referent Selection
5-minute Retention
24-hour Retention
Animal
Object
Proportional looking
ASD
NT
2.5
2
1.5
1
0.5
0
Referent Selection
5-minute Retention
24-hour Retention
Referent Selection
5-minute Retention
24-hour Retention
Animal
Object
Number of looks
ASD
NT

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[a]	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[a]	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. Participants in the present study are the same as those reported in Rothwell et al. (2024).

Table 2

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)	-1.71	0.60	-2.84	.004
	Proportion of time spent looking at the target	10.72	1.41	7.59	<.001
		AIC	**BIC**	**logLik**	**deviance**
		187.9	242.4	-80.9	161.9

Table 3

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

Table 4

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|)**
(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

Table 5

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

	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

Table 6

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

	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	7.19	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 Accuracy	(Intercept)	-0.43	0.28	-1.52	.13
	Prop. of time looking at target at referent selection	0.96	0.41	2.30	.021
	Population	0.55	0.54	1.03	.30
	Condition	0.22	0.50	0.44	.66
	Proportion x Population	0.47	0.83	0.57	.57
	Proportion x Condition	-0.72	0.84	-0.86	.39
	Population x Condition	2.47	1.00	2.47	.014
	Proportion x Population x Condition	-3.40	1.66	-2.04	.041
		AIC	**BIC**	**logLik**	**deviance**
		635.2	685.0	-305.6	611.2

Table 7

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 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	-5.05	<.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	-3.03	.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