Title
Visual Imagery Predicts Semantic Relatedness Judgements in Autistic but Not Neurotypical Adults

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors
Walker-Sharpe, Phoebe K
Twomey, Katherine E
Theakston, Anna
et al.

Publication Date
2026

Visual Imagery Predicts Semantic Relatedness Judgments in Autistic but Not Neurotypical Adults

Phoebe Walker-Sharpe[1], Anna Theakston[1], Samantha Durrant[1], Katherine Twomey (katherine.twomey@manchester.ac.uk)[1],

[1]University of Manchester, School of Psychology, Communication and Human Neuroscience

Abstract

Evidence suggests that the organization of semantic knowledge in autistic populations uses qualitatively different organization principles than neurotypical populations. Across two experiments, autistic and neurotypical adults rated the relatedness of abstract and concrete word pairs and completed the Vividness of Visual Imagery Questionnaire (VVIQ). There were no group differences in relatedness ratings; both groups rated abstract pairs as more related than concrete pairs. However, autistic participants reported significantly lower vividness of visual imagery than neurotypical participants. Critically, a VVIQ x group interaction emerged. Higher imagery scores predicted higher relatedness ratings only in the autistic group across both concept types. These findings demonstrate that equivalent semantic performance can arise through distinct processing pathways. Autistic individuals show perceptually grounded semantic processing that scales with imagery ability, while neurotypical individuals' semantic judgments seem to operate independently of visual imagery. This pattern reveals qualitative differences in how semantic knowledge is organized.

Keywords: autism, semantic relatedness, visual imagery

Experiment 1: Introduction

Word meaning derives from the relationship between linguistic expressions (i.e., words) and conceptual categories (i.e., concepts; Plag et al., 2015), which are organized within semantic networks as interconnected nodes (Quillian, 1967, 1969). The associative relationships between concepts facilitate the organization of semantic knowledge (Unger & Fisher, 2021).

Concreteness – the extent to which the concept refers to a perceivable referent (Brysbaert et al., 2013) – is an organizing principle of semantic knowledge (Vigliocco et al., 2018). Concrete concepts (e.g., *dog*) are directly experienced through sensory modalities and learnt through "word-to-world mapping" (Gleitman et al., 2005); abstract concepts (e.g., *justice*) are detached from sensory experience, lack single clearly perceivable referents (Borghi & Binkofski, 2014; Villani et al., 2024), and are learnt through sentence-to-world mappings (Vigliocco et al., 2018).

Concepts across the concreteness continuum exhibit small-world network properties – high local clustering with short paths between concepts (Lakhzoum et al., 2021). Within these small-world networks, the nature of associative relationships varies systematically. Concrete concepts form semantic neighborhoods based on shared perceptual and functional properties (McRae et al., 2005); for instance, *dog* and *cat* consistently share features like 'has fur', 'has four legs,' and 'is a pet,' creating substantial physical and taxonomic feature overlap. In contrast, abstract concepts show more variable, context-dependent, associations driven by more affective factors such as valence and arousal (Barsalou & Wiemer-Hastings, 2005; De Deyne et al., 2019) and lack the same featural content.

Neuropsychological evidence suggests concrete concepts rely on similarity-based organization, while abstract concepts depend more on associative relations (Crutch & Warrington, 2005, 2010) – a pattern supported by large-scale computational analyses (Hill et al., 2014). These organizational differences have processing consequences: abstract words take longer to process than concrete words in word naming and sentence completion tasks (Schwanenflugel & Stowe, 1989). Importantly, this body of research on semantic processing draws primarily from neurotypical populations. However, emerging evidence suggests that semantic processing may differ in autistic individuals.

Autism and Semantic Processing

The acquisition of concrete and abstract words in autistic children without language impairment is comparable to neurotypical controls in auditory lexical decision and verbal definition tasks (Norbury et al., unpublished data; cited in Vigliocco et al., 2018), and autistic children comprehend concrete and abstract word meanings, comparable to neurotypical controls, in the Stroop Paradigm (Eskes et al., 1990).

Yet, converging evidence suggests that, compared to a neurotypical sample, the organization of semantic knowledge in autism uses qualitatively different organizational principles (Gaffrey et al., 2007). Autistic children and adolescents demonstrate weaker associative relationships between lexical items in description tasks (Luo et al., 2016) and generate a higher proportion of unrelated associations in word association tasks (Battaglia, 2012). Autistic adults also

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

produce smaller semantic neighborhoods that are more restricted to specific subcategories in verbal fluency tasks (Ehlen et al., 2020) suggesting a more compartmentalized network structure. Autistic children and adolescents also produce semantic unconventionalities in semi-structured interviews (Zane & Grossman, 2024).

Neuroimaging evidence suggests enhanced activation in sensory and perceptual cortices during semantic tasks in autistic populations (Gaffrey et al., 2007), indicating greater reliance on perceptual grounding for conceptual representation. While semantic priming effects are preserved for simple common nouns, semantic performance is positively correlated with performance on non-verbal cognitive measures – a relationship absent in neurotypical controls (Toichi & Kamio, 2001) – suggesting reliance on different underlying strategies.

This perceptual processing style may be particularly well-suited to concrete concepts, which form semantic neighborhoods based on shared perceptual and functional properties, but less suited to abstract concepts that depend on variable, context-dependent linguistic-associative networks. This raises the possibility that whilst conceptual knowledge is preserved, semantic networks in autism may be structurally less precise or more diffuse than neurotypical populations. Rather than producing degraded associations – which would imply a failing or 'noisy' version of semantic associations – autistic individuals may instead apply distinct criteria for relatedness based on specific, rule-based, or idiosyncratic features. This suggests network organization that is structured differently, rather than one that is simply less functional.

No research to date has directly examined whether associative connection strengths differ between abstract and concrete concepts in autism. Given that these concept types rely on fundamentally different representational bases (perceptual versus linguistic-associative), understanding whether connection strength patterns differ across concept type is essential for characterizing the architecture of autistic semantic networks.

The Current Experiment

Experiment 1 has two goals. The first is to examine whether relatedness ratings differ between concrete and abstract word pairs. Semantic relatedness judgment captures functional associations between words rather than lexical relatedness, making it an appropriate measure of associative connection strength in semantic networks. Given that concrete concepts form stronger and more consistent associative connections based on shared perceptual and functional properties, whereas abstract concepts show more variable, context-dependent associations with less featural overlap, within-category relatedness should be higher for concrete than abstract pairs. We therefore predict that concrete word pairs will be rated as more related than abstract word pairs (H1).

The second goal is to determine whether relatedness ratings (a) differ between autistic and neurotypical groups and (b) whether these ratings are modulated by concreteness. Given evidence that autistic individuals may rely more heavily on visual information in semantic processing (Gaffrey et al., 2007), we predict a Group x Pair Type interaction, such that the difference between abstract and concrete relatedness ratings will be greater for the autistic group (H2) compared to neurotypical individuals.

Method

Both experiments were preregistered on OSF prior to data collection (https://osf.io/uyfv2).

Participants

Sample size was determined via power analysis in R Version 4.4.0 (R Core Team, 2024). The final sample included 80 adult participants (power = 82.20%, b =-9.62, $p < .001$): 40 autistic adults ($M_{\text{age}} = 34.7$, $SD_{\text{age}} = 13.1$, range = 18-64) with formal autism diagnoses and 40 neurotypical adults ($M_{\text{age}} = 37.6$, $SD_{\text{age}} = 12.8$, range = 19-63). All participants were monolingual English-speaking UK residents with no prior diagnoses affecting language or learning.

Participants were recruited online via Prolific (www.prolific.com), an online research platform with integrated screening tools. Autistic participants were identified via Prolific's integrated screening tool. Participants were included if they self-reported a formal clinical diagnosis of autism spectrum disorder made by a qualified medical specialist (psychiatrist, psychologist, or equivalent). Thirty-one participants withdrew during the study, and three participants were excluded for failing attention checks. These exclusions were additional to the final sample.

Design

The study used a 2 (Group type: autistic, neurotypical) x 2 (Pair Type: abstract, concrete) mixed design. The dependent variable was participants' relatedness rating for each word pair, measured on a continuous scale (0 - 100).

Materials

Lexical Stimuli Stimuli consisted of 120 noun pairs (60 abstract, e.g., *negation-humbleness*; 60 concrete, e.g., *moon-clothes*), controlled for length, age of acquisition (AoA), frequency and cosine similarity (a proxy for semantic relatedness). Words were selected from SUBTLEX-UK (van Heuven et al., 2014) and Brysbaert concreteness ratings (Brysbaert et al., 2014), constrained to UK English usage, AoA < 16 years and < 15 letters. Abstract words had concreteness ratings between 1.00-2.00; concrete words had ratings between 4.00-5.00.

Words were grouped into frequency bins (1.00-1.99, 2.00-2.99, 3.00-3.99, 4.00-5.00) and randomly paired within bins and word type. Cosine similarity scores were calculated using BERT embeddings (Devlin et al., 2019) to quantify semantic relatedness (0 = unrelated, 1 = identical). Pairs with low relatedness (<0.333) were excluded due to insufficient numbers of word pairs available.

The final set included 15 abstract and 15 concrete pairs per frequency bin, with half showing mid-relatedness (0.333-0.666) and half high-relatedness (>0.666) – 120 pairs in total. An independent-samples *t*-test confirmed no significant difference in cosine similarity between abstract pairs (M = 0.68, SD = 0.14) and concrete pairs (M = 0.68, SD = 0.14), $t(117.97) = 0.08$, $p = .93$, $d = 0.01$ (95% CI: [-0.04,0.06]).

Procedure

Participants completed the task online via Gorilla Experiment Builder (https://gorilla.sc). After providing informed consent and demographics, eligible participants completed three practice trials followed by 120 randomized experimental trials. Practice items were selected from non-noun word classes to avoid priming effects on the experimental stimuli, and were designed to illustrate weak, moderate, and strong semantic relationships (e.g., *stop–embarrass*, *gather–substitute*, and *walk–drive* respectively). Each trial presented the question "On a scale of 0-100 how related are the following words?" with the word pair below. A horizontal continuous slider scale appeared below (0 = weakly related, 100 = strongly related) for rating semantic relatedness. Six attention checks appeared throughout; participants were required to pass at least 2 of 6 for inclusion. A short break was provided every 30 trials. Mean completion time was 17 minutes. All participants received £3 compensation.

Results

All 80 participants completed 120 trials with no missing data. Split-half reliability was excellent ($r = .86$, p = <.001) and comparable across groups (autistic: $r = .91$; neurotypical: $r = .80$). Groups did not differ significantly in age ($t(78) = 1.02$, $p = .311$), gender distribution (Fisher's exact test, $p = .623$), qualification level ($t(75) = -1.25$, $p = .215$) or time taken to complete all trials ($t(74) = -0.39$, $p = .695$).

Relatedness Ratings

Abstract word pairs were rated as more related (M = 42.3, SD = 31.9) than concrete word pairs (M = 18.7, SD = 26.9), consistently across both groups (Figure 1).

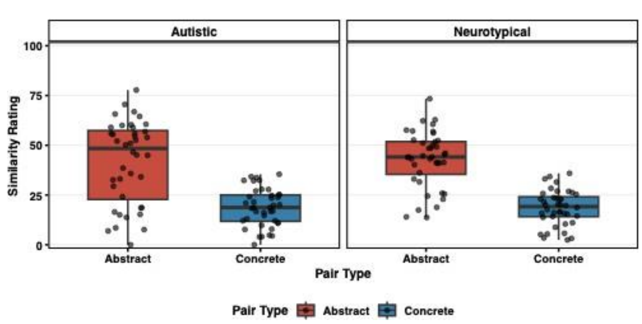

Figure 1. Experiment 1: Relatedness ratings by group and pair type. Individual points show participant means; boxes show medians and interquartile ranges.

Statistical Models

Linear mixed-effects models (LMEMs) were fitted using the lme4 package (Bates et al., 2015) with significance tests conducted using lmerTest (Kuznetsova et al., 2017) in R version 4.4.1 (R Core Team, 2024). Following the preregistered analysis plan, models included fixed effects of group (autistic vs. neurotypical), pair type (abstract vs. concrete) and their interaction, with random intercepts for participant. Statistical significance was assessed at $\alpha = .05$.

Model comparison via likelihood ratio tests revealed that pair type significantly improved model fit compared to the null model ($\chi^2(1) = 1701.94$, $p < .001$), confirming a significant main effect of pair type. Adding group did not improve fit ($\chi^2(1) = 0.03$, $p = .865$), and the group × pair type interaction did not significantly improve fit ($\chi^2(1) = 1.57$, $p = .211$). The full model showed a significant main effect of pair type ($\beta = .37$, $SE = .01$, $t = 43.17$, $p < .001$), with abstract pairs rated as more related than concrete pairs. The group × pair type interaction was not significant ($\beta = -.01$, $SE = .01$, $t = -1.25$, $p = .211$), indicating that the effect of pair type was comparable across groups.

Experiment 2: Introduction

Experiment 1 established that both groups rated abstract word pairs as more related than concrete word pairs. However, greater variability in autistic participants' abstract word ratings suggested potential individual differences in semantic processing. In addition, cosine similarity showed weak correlations with human ratings (rs = .19-.29), introducing measurement noise; inspection of the stimulus list revealed that BERT-derived similarity scores did not always align with human judgments of relatedness. Experiment 2 addressed both issues by selecting stimuli based on neurotypical participants' relatedness ratings from Experiment 1, ensuring stimulus relatedness reflected human semantic judgments rather than computational estimates.

Visual Imagery and Semantic Processing

Visual imagery – the generation of mental visual representations – contributes to conceptual processing by reactivating prior perceptual experiences (Bergen et al., 2007). Individuals with more vivid visual imagery show enhanced performance on semantic tasks, including metaphor comprehension (Stephan et al., 2022) and semantic verification (Kan et al., 2003), suggesting they access richer conceptual representations.

If conceptual knowledge is partially grounded in imagery-based simulations, then variations in imagery ability should predict the strength of associative relationships between concepts. Visual imagery may facilitate semantic processing for both concept types: by directly activating perceptual features for concrete concepts and supporting associative links between abstract concepts. Individuals with more vivid visual imagery may therefore access richer conceptual representations overall, strengthening perceived semantic relationships between word pairs.

Vividness of Visual Imagery in Autism

Autistic individuals report reduced vividness of visual imagery on self-report measures compared to neurotypical controls (King et al., 2024) and elevated rates of aphantasia – the absence of voluntary visual imagery (5-20% versus 0.7-3.9%; Dance et al., 2022; King et al., 2024; Zeman et al., 2020). In contrast, behavioral studies show preserved or enhanced imagery performance in some autistic individuals, including superior maintenance of visual patterns (Bled et al., 2024). This discrepancy between self-reported vividness and behavioral performance suggests substantial heterogeneity in autistic visual imagery. Individual differences in imagery could explain the absence of group differences in Experiment 1: if imagery ability predicts semantic judgments, within-group variability may obscure group-level patterns.

The Current Experiment

Experiment 2 retains the hypotheses from Experiment 1 regarding main effect of pair type (H1) and group x pair type interaction (H2) and has two additional goals. First, we examine whether autistic and neurotypical adults differ in vividness of visual imagery using the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973). We predict that the autistic group will report lower vividness of visual imagery than the neurotypical group (H3). Second, we test whether vividness of visual imagery predicts relatedness ratings and whether this relationship differs between groups. If conceptual knowledge is partially grounded in imagery-based simulations, then higher vividness should be associated with higher relatedness ratings overall (H4). Additionally, we predict a vividness of visual imagery x pair type interaction (H5): individuals with lower vividness of visual imagery may process both concept types similarly through linguistic associations, resulting in smaller concrete-abstract differences in relatedness ratings. To address methodological limitations, word pairs were selected using relatedness ratings from neurotypical participants in Experiment 1 rather than a computational measure, ensuring stimulus relatedness reflects human relatedness judgments.

Method

Participants

Sample size was determined to maintain equivalent statistical power to Experiment 1. The final sample included 240 adult participants: 120 autistic adults (M_{age} = 31.5, SD_{age} = 10.6, range = 19-61) with formal autism diagnoses and 120 neurotypical adults (M_{age} = 41.2, SD_{age} = 12.6, range = 19-65). All participants were monolingual English-speaking UK residents with no prior diagnoses affecting language or learning.

Participants were recruited using the same procedure as Experiment 1. One-hundred-and-one participants withdrew during the study; ninety-three additional participants were excluded for failing attention checks (minimum 2 of 3 required). These exclusions were additional to the final sample.

Design

The study used a 2 (Group: autistic, neurotypical) x 2 (Pair Type: abstract, concrete) mixed design with VVIQ as a continuous predictor. The dependent variable was relatedness ratings (0-100).

Materials

Semantic Relatedness Task Stimuli consisted of 40 noun pairs (20 abstract, e.g., *courage-ambition*; 20 concrete, e.g., *crab-bucket*) selected from Experiment 1: the 10 lowest and highest neurotypical-rated pairs per category, creating clearer within-category relatedness contrasts. The semantic relatedness task interface was identical to Experiment 1.

VVIQ The VVIQ (Marks, 1973) is a 16-item self-report questionnaire assessing imagery vividness across four scenarios (e.g., visualizing a natural scene). Participants rated image clarity on a reverse-coded 5-point scale for more intuitive scoring (1 = no image at all, 5 = perfectly clear and vivid), resulting in total scores from 16 (lowest vividness) to 80 (highest vividness), following recent practice (Wright et al., 2024; Dance et al., 2022; Milton et al., 2021).

Procedure

The procedure was identical to Experiment 1, with three exceptions. First, participants completed 40 experimental trials with one break after trial 20. Second, three attention checks were presented at trials 10, 20 and 30; participants were required to pass at least two checks. Third, following the semantic relatedness task, participants completed the VVIQ. Mean completion time was 12 minutes. All participants received £1.50 in compensation.

Results

All 240 participants completed 40 trials with no missing data. Split-half reliability was good (r = .61, p <.001) and comparable across groups (autistic: r = .69; neurotypical: r = .64). Groups differed significantly in age (t(231.05) = 6.41, p <.001) and in gender distribution (Fisher's exact test, p = .037) but not in qualification level (t(237.85) = -0.23, p = .815) or time taken (t(220.77) = 1.61, p = .109).

Relatedness Ratings

Replicating Experiment 1, abstract word pairs (autistic: M = 41.3, SD = 33.7; neurotypical: M = 42.2, SD = 33.7) were rated as more related than concrete word pairs in both groups (autistic: M = 27.2, SD = 32.9; neurotypical: M = 30.0, SD = 31.7). See Figure 2.

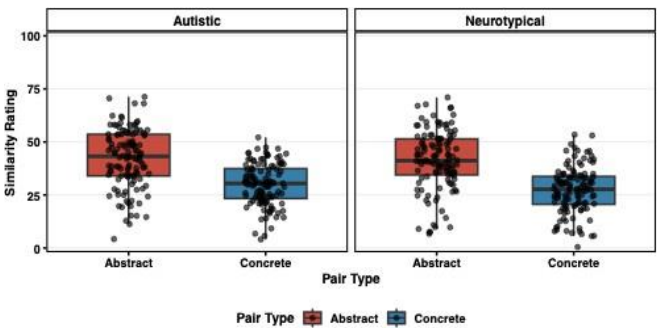

Figure 2. Experiment 2: Relatedness ratings by group and pair type. Both groups rated abstract pairs as more related than concrete pairs. Individual points show participant means; boxes show medians and interquartile ranges.

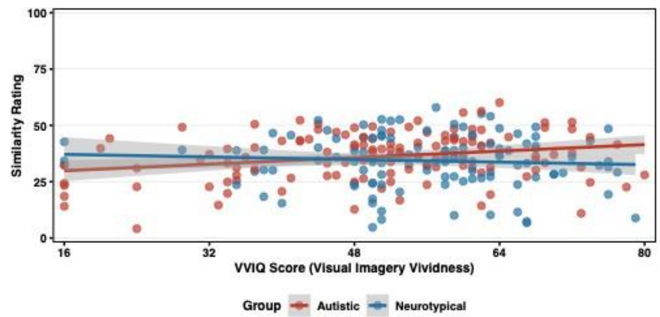

Figure 3. Experiment 2: VVIQ × Group interaction. Visual imagery vividness predicted relatedness ratings in autistic adults but not neurotypical adults. Individual points show participant means; lines show fitted regression slopes with 95% confidence intervals.

Visual Imagery

The neurotypical group reported significantly higher VVIQ scores (M = 55.9, SD = 11.4) than the autistic group (M = 50.2, SD = 15.2; t(221.11) = 3.28, p = .001).

Fixed Effects and Interaction Effects

Linear mixed-effects models (LMEMs) included fixed effects of group, pair type, VVIQ score, age, gender, and group x pair Type, VVIQ x pair type and VVIQ x group interactions, with by-participant random intercepts.

The model revealed a main effect of Pair Type (β = .20, SE = .01, t = 20.48, p < .001), with abstract pairs rated as more related than concrete pairs. The main effects of group (β = .01, SE = .02, t = -0.42, p = .672) and VVIQ (β = .03, SE = .02, t = 1.27, p = .207) were not significant. Age and Gender were included as covariates given significant group differences in demographics. Age was a significant negative predictor of relatedness ratings (β = -.07, SE = .02, t = -3.07, p = .002), with follow-up analyses revealing this effect was specific to the neurotypical group (β = -.09, SE = .03, t = -2.82, p = .006) and not significant in the autistic group (β = -.03, SE = .03, t = -1.22, p = .224). Gender did not significantly predict relatedness ratings ($\chi^2(1)$ = 9.09, p = .555). The Group × Pair Type interaction was also not significant (β = -.01, SE = .01, t = -1.15, p = .249); Bayes Factor analysis provided strong evidence against this interaction (BF_{01} = 8.62). The VVIQ × Pair Type interaction was also not significant (β = .01, SE = .01 t = 1.18, p =.237). However, the VVIQ × Group interaction was significant (β = 0.04, SE = 0.02, t = 2.06, p = .041).

Visual Imagery and Group Interaction

To understand the nature of the VVIQ x group interaction, separate models were fitted to each group. Results revealed that VVIQ significantly predicted relatedness ratings in the autistic group (β = 0.08, SE = 0.06, t = 2.84, p = .005), with higher vividness of visual imagery associated with higher relatedness ratings. This effect was not present in the neurotypical group (β = -0.01, SE = 0.09, t =-0.43, p = .668). See Figure 3.

General Discussion

The present studies examined whether autistic and neurotypical adults differ in semantic judgments for abstract and concrete word pairs, and whether visual imagery modulates these judgments. Across two experiments both groups rated abstract pairs as more related than concrete pairs with no group differences. In Experiment 2, however, a significant difference was observed in visual imagery across groups: the autistic group reported lower vividness than the neurotypical group. Critically, there was a significant VVIQ x group interaction, in that higher vividness of visual imagery predicted higher relatedness ratings in the autistic group across both concept types, with no equivalent relationship in the neurotypical group.

Contrary to predictions, abstract word pairs were rated as more related than concrete pairs. This unexpected pattern may reflect the broader associative flexibility of abstract concepts compared to the more constrained featural organization (relatedness) of concrete concepts (Crutch & Warrington 2005, 2010; Hill et al., 2014). Abstract concepts form wider-ranging connections across the semantic network and are used flexibly across diverse contexts (Barsalou & Wiemer-Hastings, 2005; De Deyne et al., 2019). This associative flexibility may create multiple pathways of connection between abstract concepts, producing stronger perceived relatedness. In contrast, concrete concepts' rich perceptual and functional features, while providing a basis for relatedness, may also make their differences appear more salient. The specific featural distinctions between concrete concepts may be more readily perceived than their shared features, leading to lower relatedness ratings despite their featural overlap. Thus, the constrained, feature-based organization of concrete concepts may highlight distinctions rather than commonalities in explicit relatedness judgments.

The absence of group differences in relatedness judgments is consistent with evidence of comparable acquisition and comprehension of concrete and abstract concepts in autistic and neurotypical individuals (Norbury et al., unpublished data, cited in Vigliocco et al., 2018). Yet, this comparable performance contrasts with evidence of distinct patterns of semantic associations observed in autism in generative tasks (e.g., Battaglia, 2012; Ehlen et al., 2020; Luo et al., 2016; Zane & Grossman, 2024). This may suggest differences in the nature of these tasks (i.e., generative vs. evaluative) may engage distinctly different processing mechanisms, given that generative tasks require unconstrained retrieval and production of semantic associations whereas evaluative tasks (i.e., relatedness judgements) provide the associative structure. Notably, tasks that tap into semantic representations through constrained paradigms – such as memory encoding tasks where autistic and neurotypical groups show comparable performance regardless of encoding level (Toichi & Kamio, 2002) – similarly suggest intact rather than degraded representations, indicating that autistic semantic differences may be specific to the demands of unconstrained generative retrieval.

The autistic group reported significantly lower vividness than the neurotypical group on the VVIQ, consistent with previous evidence of lower self-reported imagery vividness (King et al., 2024). Despite this, visual imagery significantly predicted relatedness ratings in the autistic group across concept types: those with higher vividness of visual imagery rated word pairs as more related across both abstract and concrete concepts – with no such relationship in the neurotypical group. This suggests that visual imagery may play a comparatively greater role in autistic semantic processing, supporting neuroimaging evidence of enhanced visual cortex activation during semantic tasks in autism (Gaffrey et al., 2007). This pattern is consistent with perceptually grounded semantic processing, such that word pairs are evaluated through richer perceptual simulations when visual imagery is more vivid. It should be noted, however, that the modest effect size of the VVIQ x group interaction suggests that the vividness of visual imagery is one contributor to individual differences in autistic semantic processing rather than a primary determinant, and replication is warranted. This evidence supports a difference-based rather than deficit-based account of autistic semantic processing: comparable performance at the group level emerges from qualitatively different underlying strategies.

These findings also inform the understanding of how visual imagery contributes to semantic processing in neurotypical groups. Whilst visual imagery facilitates performance in semantic verification and metaphor generation tasks (Kan et al., 2003; Stephan et al., 2022), it did not predict relatedness judgments. Importantly, the range of scores in each group was comparable, with each group recording similar minimum and maximum scores. This provides evidence against a threshold effect whereby visual imagery is only useful below a certain level. Rather, this discrepancy may reflect task-specific differences: tasks requiring active generation and integration of semantic relationships may benefit from visual imagery, whereas evaluative tasks may draw on semantic knowledge that operates independently of visual imagery.

Limitations and Future Directions

Several limitations warrant consideration. Autism diagnoses relied on self-reported confirmation of formal diagnosis via Prolific's screening tools rather than independent verification, which is a common limitation of online research designs. In the present studies, screening was restricted to participants self-reporting a formal clinical diagnosis made by a qualified medical specialist, excluding those self-identified without formal diagnosis or on diagnostic waiting lists – a more stringent criterion than some previous online studies (e.g., King et al., 2024).

Though the vividness of visual imagery was assessed using a widely used self-report measure (Marks, 1973), lower VVIQ scores in the autistic group may reflect differences in introspective reporting rather than genuine differences in imagery vividness, given evidence of discrepancies between objective performance and subjective self-report of internal states (Williams et al., 2023). Though VVIQ scores correlate with neural activity during imagery tasks (Fulford et al., 2018), whether this extends to autistic individuals remains unclear.

The relationship between visual imagery and semantic processing may also be modulated by inner speech (self-directed language used for conceptual manipulation) which was not measured in the current studies. Evidence regarding inner speech use in autism is mixed, with some studies suggesting reduced reliance on verbal mediation (Whitehouse et al., 2006) and others finding comparable inner speech use (Williams et al., 2008), making its role difficult to establish (see Williams et al., 2016).

At the stimulus level, the weak correlation between cosine similarity and human relatedness ratings in Experiment 1 suggests that distributional models may fail to capture aspects of semantic associations that underlie human relatedness judgments – a limitation with particular relevance for modelling autistic semantic cognition, given that vividness of visual imagery predicted relatedness judgments in the autistic group. Future work should examine the interplay between visual imagery, inner speech, and semantic processing in autism, alongside neuroimaging evidence of the neural correlates of the imagery-relatedness relationship observed.

Conclusion

Visual imagery predicted semantic relatedness judgments in autistic adults but not neurotypical adults, despite both groups producing equivalent ratings. This study is the first to provide evidence that comparable semantic performance may emerge from different processing strategies, suggesting that future characterizations of autistic semantic processing should focus less on whether semantic organization differs and more on how semantic knowledge is accessed.

Acknowledgements

This work was supported by the ESRC International Centre for Language and Communicative Development (LuCiD; ES/L008955) and a PhD studentship from the University of Manchester. This research was conducted collaboratively by a team that included autistic researchers. We thank all participants for their time and contribution to this research.

References

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher & R. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thought* (pp. 129-163).

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Battaglia, D. (2012). Word association and semantic priming in individuals with autism spectrum disorders. Doctoral Dissertation, City University, New York. https://www.proquest.com/openview/e7cec450bba83a361d a7f14717080e21/1?pq-origsite=gscholar&cbl=18750

Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive science*, *31*(5), 733-764. https://doi.org/10.1080/03640210701530748

Bled, C., Guillon, Q., Mottron, L., Soulières, I., & Bouvet, L. (2024). Visual mental imagery abilities in autism. *Autism Research*, *17*(10), 2064-2078. https://doi.org/10.1002/aur.3192

Borghi, A. M., & Binkofski, F. (2014). Words as social tools: An embodied view on abstract concepts. Springer.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness rating for 40 thousand generally known English words lemmas. Behavior Research Methods, 46(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*(3), 615-627. https://doi.org/10.1093/brain/awh349

Crutch, S. J., & Warrington, E. K. (2010). The differential dependence of abstract and concrete words upon associative and similarity-based information: Complementary semantic interference and facilitation effects. *Cognitive neuropsychology*, *27*(1), 46-71. https://doi.org/10.1080/02643294.2010.491359

Dance, C. J., Ipser, A., & Simner, J. (2022). The prevalence of aphantasia (imagery weakness) in the general population. *Consciousness and Cognition*, *97*, 103243. https://doi.org/10.1016/j.concog.2021.103243

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *BehaviorResearch methods*, *51*(3), 987-1006. https://doi.org/10.3758/s13428-018-1115-7

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)* (pp. 4171-4186).

http://wordvec.colorado.edu/pairwise_comparison.html

Ehlen, F., Roepke, S., Klostermann, F., Baskow, I., Geise, P., Belica, C., Ole Tiedt, H., & Behnia, B. (2020). Small semantic networks in individuals with autism spectrum disorder without intellectual impairment: a verbal fluency approach. *Journal of Autism and Developmental Disorders*, *50*(11), 3967-3987. https://doi.org/10.1007/s10803-020-04457-9

Eskes, G. A., Bryson, S. E., & McCormick, T. A. (1990). Comprehension of concrete and abstract words in autistic children. *Journal of autism and developmental disorders, 20*(1), 61-73. https://doi.org/10.1007/BF02206857

Gaffrey, M. S., Kleinhans, N. M., Haist, F., Akshoomoff, N., Campbell, A., Courchesne, E., & Müller, R. A. (2007). A typical participation of visual cortex during word processing in autism: An fMRI study of semantic decision. Neuropsychologia, 45(8), 1672-1684. https://doi.org/10.1016/j.neuropsychologia.2007.01.008

Fulford, J., Milton, F., Salas, D., Smith, A., Simler, A., Winlove, C., & Zeman, A. (2018). The neural correlates of visual imagery vividness: An fMRI study and literature review. *Cortex, 105*, 26–40. https://doi.org/10.1016/j.cortex.2017.09.014

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, *1*(1), 23–64. https://doi.org/10.1207/s15473341lld0101_4

Hill, F., Korhonen, A., & Bentz, C. (2014). A quantitative empirical analysis of the abstract/concrete distinction. Cognitive Science, 38(1), 162-177. https://doi.org/10.1111/cogs.12076

Kan, I. P., Barsalou, L. W., Olseth Solomon, K., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology*, *20*(3-6), 525-540. https://doi.org/10.1080/02643290244000257

King, R., Buxton, H., & Tyndall, I. (2024). Aphantasia and autism: An investigation of mental imagery vividness. *Consciousness and cognition*, *125*, 103749. https://doi.org/10.1016/j.concog.2024.103749

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13

Lakhzoum, D., Izaute, M., & Ferrand, L. (2021). *Semantic network analysis of abstract and concrete word associations.* [Preprint]. arXiv.

https://doi.org/10.48550/arXiv.2110.09096
Luo, S. X., Shinall, J. A., Peterson, B. S., & Gerber, A. J. (2016). Semantic mapping reveals distinct patterns in descriptions of social relations in adults with autism spectrum disorder. Autism Research, 9(8), 846-853. https://doi.org/10.1002/aur.1581
Marks, D. F. (1973). Visual imagery differences and eye movements in the recall of pictures. Perception & Psychophysics, 14(3), 407-412. https://doi.org/10.3758/BF03211175
McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior research methods*, *37*(4), 547-559. https://doi.org/10.3758/BF03192726
Milton F., Fulford, J., Dance, C., Gaddum, J., Heuerman-Williamson, B., Jones, K., Knight K., F., MacKisack, M., Winlove, C., & Zeman, A. (2021). Behavioral and neural signatures of visual imagery vividness extremes: Aphantasia versus hyperphantasia. *Cerebral cortex communications*, *2*(2), tgab035, https://doi.org/10.1093/texcom/tgab035
Plag, I., Arndt-Lappe, S., Braun, M., & Schramm, M. (2015). *Introduction to English linguistics*. Walter de Gruyter GmbH & Co KG.
Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral science*, *12*(5), 410-430. https://doi.org/10.1002/bs.3830120511
Quillian, M. R. (1969). The teachable language comprehender: A simulation program and theory of language. *Communications of the ACM*, *12*(8), 459-476. https://doi.org/10.1145/363196.363214
R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. Reading Research Quarterly, 114-126. https://doi.org/10.2307/748013
Stephan, E., Shidlovski, D., Shanetzki, M., Agajany, K., & Mashal, N. (2022). The role of vividness of imagery in metaphor generation. *European Journal of Social Psychology*, *53*(1), 212-229. https://doi.org/10.1002/ejsp.2898
Toichi, M., & Kamio, Y. (2001). Verbal association for simple common words in high-functioning autism. *Journal of Autism and Developmental Disorders*, *31*(5), 483-490. https://doi.org/10.1023/A:1012216925216
Unger, L., & Fisher, A. V. (2021). The emergence of richly organized semantic knowledge from simple statistics: A synthetic review. *Developmental Review, 60*, 100949. https://doi.org/10.1016/j.dr.2021.100949
Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly journal of experimental psychology*, *67*(6), 1176-1190. https://doi.org/10.1080/17470218.2013.850521
Vigliocco, G., Ponari, M., & Norbury, C. (2018). Learning and processing abstract words and concepts: Insights from typical and atypical development. Topics in Cognitive Science, 10(3), 533-549. https://doi.org/10.1111/tops.12347
Villani, C., Loia, A., & Bolognesi, M. M. (2024). The semantic content of concrete, abstract, specific, and generic concepts. *Language and Cognition*, *16*(4), 867-894. http://doi.org/10.1017/langcog.2023.64
Whitehouse, A. J., Maybery, M. T., & Durkin, K. (2006). Inner speech impairments in autism. *Journal of Child Psychology and Psychiatry*, *47*(8), 857-865. https://doi.org/10.1111/j.1469-7610.2006.01624.x
Williams, D., Happé, F., & Jarrold, C. (2008). Intact inner speech use in autism spectrum disorder: evidence from a short-term memory task. *Journal of Child Psychology and Psychiatry*, *49*(1), 51-58. https://doi.org/10.1111/j.1469-7610.2007.01836.x
Williams, D. M., Peng, C., & Wallace, G. L. (2016). Verbal thinking and inner speech use in autism spectrum disorder. *Neuropsychology Review*, *26*(4), 394-419. https://doi.org/10.1007/s11065-016-9328-y
Williams, Z. J., Suzman, E., Bordman, S. L., Markfeld, J. E., Kaiser, S. M., Dunham, K. A., Zoltowski, A. R., Failla, M. D., Cascio, C. J., & Woynaroski, T. G. (2023). Characterizing interoceptive differences in autism: A systematic review and meta-analysis of case–control studies. *Journal of autism and developmental disorders*, *53*(3), 947-962. https://doi.org/10.1007/s10803-022-05656-2
Wright, D. J., Scott, M. W., Kraeutner, S. N., Barhoun, P., Bertollo, M., Campbell, M. J., Walzing, B. M., Dahm., S. F., Esselaar, M., Frank, C., Hardwick, R. M., Fuelscher, I., Marshall, B., Hodges, N. J., Hyde, C., & Holmes, P. S. (2024). An international estimate of the prevalence of differing visual imagery abilities. *Frontiers in Psychology*, *15*, 1454107. https://doi.org/10.3389/fpsyg.2024.1454107
Zane, E. R., & Grossman, R. B. (2024). Normal but Different: Autistic Adolescents Who Score Within Normal Ranges on Standardized Language Tests Produce Frequent Linguistic Irregularities in Spontaneous Discourse. *Autism & Developmental Language Impairments*, *9*. https://doi.org/10.1177/23969415241283378
Zeman, A., Milton, F., Della Sala, S., Dewar, M., Frayling, T., Gaddum, J., Hattersley, A, Heuerman-Williamson, B., Jones, K., MacKisack, M. & Winlove, C. (2020). Phantasia–the psychological significance of lifelong visual imagery vividness extremes. Cortex, 130, 426-440. https://doi.org/10.1016/j.cortex.2020.04.003