Individual Differences in Visual Object Recognition: An Investigation Across Neurotypical and Neurodevelopmental Populations

Bahareh Jozranjbar

Thesis for the degree of Philosophiae Doctor

September 2023

School of Health Sciences

FACULTY OF PSYCHOLOGY

UNIVERSITY OF ICELAND

Individual Differences in Visual Object Recognition: An Investigation Across Neurotypical and Neurodevelopmental Populations

Bahareh Jozranjbar

Thesis for the degree of Philosophiae Doctor

Supervisor(s)

Heida Maria Sigurdardottir, Árni Kristjánsson

Doctoral committee

Heida Maria Sigurdardottir, Árni Kristjánsson, Randi Starrfelt

September 2023

School of Health Sciences
FACULTY OF PSYCHOLOGY
UNIVERSITY OF ICELAND

Einstaklingsmunur á sjónrænum hlutakennslum: Rannsókn á fólki með og án taugaþroskaraskana

Bahareh Jozranjbar

Ritgerð til doktorsgráðu

Leiðbeinandi/leiðbeinendur

Heida Maria Sigurdardottir, Árni Kristjánsson

Doktorsnefnd

Heida Maria Sigurdardottir, Árni Kristjánsson, Randi Starrfelt

September 2023

Heilbrigðisvísindasvið

SÁLFRÆÐIDEILD

HÁSKÓLI ÍSLANDS

ISBN: 978-9935-9764-0-6

ORCID: 0000-0002-1006-7626

Reykjavik, Iceland 2023

Ágrip

Þessi ritgerð snýst um sjónræn hlutakennsl, þá sérstaklega hvort og þá hvernig tilteknir slíkir hæfileikar tengjast innbyrðis, annars vegar hjá fólki með taugaþroskaraskanir og hins vegar í almennu þýði. Við rannsökum áhrif hlutagerðar, reynslu og sjónræns vinnsluminnis.

Samkvæmt rannsókn okkar á lesblindu á sumt lesblint fólk erfitt með að bera kennsl á hús en á ekki í sams konar erfiðleikum með að þekkja andlit. Þetta gengur gegn þeirri hugmynd að lesblinda sé aðeins bundin við erfiðleika í lestri. Frekari rannsókn okkar á almennu þýði ýtir enn frekari stoðum undir þá hugmynd að andlitskennsl séu í grundvallaratriðum ólík því að bera kennsl á orðleysur eða hús. Áhugavert var að lesblindir virtust nota sams konar aðferð í hlutaskynjun hvort sem um var að ræða einstaka sjónræna þætti hluta eða tengsl þátta, á meðan ólesblint fólk virtist beita ólíkum aðferðum eftir því hvort þættir eða tengsl þátta skiptu máli í hlutakennslum. Þessi munur á þáttum og tengslum var ekki jafngreinilegur í seinni rannsókn okkar á almennu þýði sem getur bent til að mismunandi verkefni geti kallað á ólíkar sjónskynjunaraðferðir.

Þriðja rannsókn okkar fjallaði um áhrif vinnsluminnis á sjónræn kennsl. Nákvæmni í vinnsluminni fyrir einfalda sjónræna þætti tengdist sérstaklega getu fólks til þess að muna uppbyggingu húsa. Þar sem slík tengsl voru veikari við hæfileika fólks til þess að muna sérfræðihluti bendir það til þess að fyrri þekking á hlutum geti minnkað álag á sjónrænt vinnsluminni. Ekki voru jafnsterk tengsl á milli vinnsluminnis og þess að muna einstaka sjónræna þætti húsa. Þetta gæti þýtt að það að muna uppbyggingu hluta auki minnisálag þar sem það krefst þess að fólk leggi á minnið bæði einstaka sjónræna þætti og innbyrðis afstöðu þeirra. Í ljósi þessara niðurstaðna getur verið að vandkvæði lesblindra við að þekkja hús í sjón tengist sjónrænni vinnsluminnisgetu þeirra.

Þessar rannsóknir auka skilning okkar á flóknu eðli sjónrænna hlutakennsla, bæði hjá fólki með og án taugaþroskaraskana, og sýna fram á áhrif reynslunnar og þátt sjónræns vinnsluminnis í því að þekkja hluti í sjón.

Lykilorð:

Lesblinda; Andlitskennsl; Hlutakennsl; Orðakennsl; Sjónrænt vinnsluminni

Abstract

This thesis probes the intricate nature of visual object recognition, especially how visual recognition skills are associated or dissociated in developmental disorders and the neurotypical population. Our primary areas of focus include the role of object types, the importance of visual expertise, and the impact of visual working memory.

In our investigation into developmental dyslexia, we discovered that some dyslexic readers struggled with recognizing houses, but did not face the same challenges with face recognition. This finding raises doubt about the commonly held belief that dyslexia only pertains to reading difficulties. Furthermore, our subsequent research with the neurotypical population solidified the idea that recognizing faces is fundamentally different from recognizing pseudowords or houses. Intriguingly, we found that dyslexic readers tend to adopt a unified strategy for processing visual objects, regardless of whether the task centers on an object's individual features or its overall configuration. In contrast, typical readers appeared to adjust their approach based on both featural and configural information. Yet, this differentiation was less pronounced in our second study with the neurotypical population, suggesting the potential influence of task demands on visual processing methods.

Our third study sought to understand the influence of visual working memory on these recognition patterns. We pinpointed a notable connection between visual working memory precision for simple visual elements and the ability to recall the overall structure of houses. As this connection was not found to the same degree for objects of expertise, such findings imply that familiarity with an object can potentially lessen the strain on visual working memory. This connection, however, was not present when memorizing the specific features of houses, indicating that recalling the overall structure might place more demands on memory due to the combined need to remember individual features and their relative positions. Upon revisiting our initial findings, we hypothesize that the challenges dyslexic readers face in recognizing houses might be related to their visual working memory capacities.

To conclude, this research sheds valuable light on the multifaceted nature of visual object recognition, highlighting its complexities, the role of expertise, and the contributions of visual working memory. Our conclusions enhance our understanding of visual recognition processes in both typical individuals and those with developmental challenges.

Keywords:

Dyslexia; Face recognition; Object recognition; Word recognition; Visual working memory

Acknowledgements

I must start by expressing my profound gratitude to my supervisors, Dr. Heida Maria Sigurdardottir and Dr. Árni Kristjánsson. Their steadfast support, guidance, and unwavering dedication have been the cornerstones of my doctoral journey. Beyond the confines of academic mentorship, the enriching experiences at the Icelandic Vision Lab conferences and the treasured gatherings they orchestrated not only added depth to my academic journey but also provided a sense of belonging and warmth, easing the challenge of being away from my homeland.

I also extend my thanks to Dr. Randi Starrfelt and Dr. Christian Gerlach. Their keen insights and thorough feedback were essential in elevating my research.

A special note of appreciation goes to my colleagues at the Icelandic Vision Lab. The collaborative spirit, discussions, and memories made in both gym office and Nýi Garður office have been invaluable. Each one of you has enriched this journey with unique insights, companionship, and lasting memories.

The nurturing research environment provided by the University of Iceland was pivotal. The comprehensive resources and vibrant academic atmosphere of the institution significantly bolstered my research endeavors.

I am grateful for the financial support from The Icelandic Research Fund (Grant No. 174013-051), the University of Iceland Research Fund, and the University of Iceland Teaching Fund. This backing was fundamental to my academic progress.

On a personal note, my heart brims with gratitude for my husband, Mohsen. His unwavering understanding and consistent support have been my bedrock. Likewise, my parents' endless encouragement has also been the wind beneath my wings, making this journey both achievable and significant.

Contents

List of Abbreviations

ADHD: Attention Deficit Hyperactivity Disorder

CDA: Contralateral Delay Activity

DBA: Developmental Body Agnosia

DD: Developmental Dyslexia

DOA: Developmental Object Agnosia

DP: Developmental Prosopagnosia

EEG: Electroencephalogram

FFA: Fusiform Face Area

IQ: Intelligence Quotient

MTM theory: Many-to-Many theory

O: object recognition ability

RSA: Representational Similarity Analysis

VLTM: Visual Long-Term Memory

VWM: Visual Working Memory

VWFA: Visual Word Form Area

List of Figures

List of Tables

List of Original Papers

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals (I, II, III):

I. Jozranjbar, B., Kristjánsson, Á., & Sigurdardottir, H. M. (2021). Featural and configural processing of faces and houses in matched dyslexic and typical readers. Neuropsychologia, 162, 108059.

II. Jozranjbar, B., Kristjánsson, Á., Starrfelt, R., Gerlach, C. & Sigurdardottir, H.M. (2023). Using representational similarity analysis to reveal category and process specificity in visual object recognition. Cortex, 166, 172-187.

III. Jozranjbar, B., Kristjánsson, Á., Gerlach, C., & Sigurdardottir, H. M. (submitted for publication) The impact of visual working memory constraints on object recognition.

Declaration of Contribution

Paper I: Bahareh Jozranjbar: Data collecting, Data analyzing, Writing- Original draft preparation, Writing- Reviewing and Editing. Árni Kristjánsson: Supervision, Writing- Reviewing and Editing. Heida Maria Sigurdardottir: Supervision, Writing- Reviewing and Editing.

Paper II: Bahareh Jozranjbar: Data collecting, Data analyzing, Writing- Original draft preparation, Writing- Reviewing and Editing. Árni Kristjánsson: Supervision, Writing- Reviewing and Editing. Randi Starrfelt: Writing- Reviewing and Editing. Christian Gerlach: Writing- Reviewing and Editing. Heida Maria Sigurdardottir: Supervision, Writing- Reviewing and Editing.

Paper III: Bahareh Jozranjbar: Data collecting, Data analyzing, Writing- Original draft preparation, Writing- Reviewing and Editing. Árni Kristjánsson: Supervision, Writing- Reviewing and Editing. Randi Starrfelt: Writing- Reviewing. Christian Gerlach: Writing- Reviewing and Editing. Heida Maria Sigurdardottir: Supervision, Writing- Reviewing and Editing.

1 Introduction

Imagine a situation where identifying familiar faces or reading basic words becomes an impossible task. This reflection underscores the critical role that visual cognition plays in our daily life. Despite our understanding of its importance, critical questions remain unresolved: What are the fundamental principles that govern visual object recognition? How are various visual recognition skills associated or separated? How do these factors influence the manifestation of certain developmental disorders? This thesis is dedicated to addressing these questions, focusing on the assessment of individual differences in visual object recognition, and conducting investigations across both neurotypical and neurodevelopmental populations.

While studying developmental disorders is important, drawing conclusions about typical cognitive development solely from these cases is challenging. According to some accounts (Bishop, 1997; D'Souza & Karmiloff-Smith, 2011), developmental deficits may partly reflect unique developmental pathways in the brains of individuals with such disorders. These accounts suggest that focusing only on developmental disorders to gain insights into typical cognition could overlook the dynamic and complex interactions of multiple developing systems that characterize the brain and its functional plasticity (Moses & Stiles, 2002). They emphasize that an infant's brain is initially highly interconnected, and specialization or modularization of neural networks develops over time. Hence, even when individuals with developmental disorders perform within the normal range on behavioral measures, they may be utilizing different cognitive and neurological processes. In other words, neurodevelopmental disorders may not accurately reflect the continuous process of relative modularization, which may be the endpoint of development, rather than its starting point (D'Souza & Karmiloff-Smith, 2011).

The human brain's ability to pick out a familiar face in a crowd or swiftly interpret written words is remarkable. How is this achieved? Two main accounts have been proposed as answers to this question. The domain-specific account suggests that specialized high-level visual mechanisms exist to process distinct types of visual stimuli such as faces, words, or locations (Kanwisher, 2000; Rhodes et al., 2004; Yovel & Kanwisher, 2004). In contrast, the domain-general account suggests that general high-level visual mechanisms are employed across various visual categories, albeit to varying extents (Behrmann & Plaut, 2013; Gauthier et al., 2014; Hills et al., 2015; Rice et al., 2020), and that they are shaped by the visual processing demands of a particular task (Sigurdardottir et al., 2021). The domain-general account posits that while general visual mechanisms operate across various visual categories, the core debate centers around their significance in perception or high-level vision. This perspective suggests that such mechanisms function across visual categories to differing degrees (Behrmann & Plaut, 2013; Gauthier et al., 2014; Hills et al., 2015; Rice et al., 2020), and they are tailored by the visual processing demands of distinct tasks (Sigurdardottir et al., 2021). It is pivotal to emphasize that these mechanisms are not limited to low-level visual deficits, which would naturally influence all visual categories, but reach into the nuances of sophisticated visual processing.

In the subsequent sections, we will delve more deeply into these two prominent accounts.

1.1 Domain specificity/generality debate

Within the study of object recognition, there is a distinction between domain-specific mechanisms, which apply to a narrow range of object categories (in the extreme, just one), vs. domain-general mechanisms, which pertain to a broader range of categories (in the extreme, all). As we progress into the subsequent sections, we will center our discussions around the innateness hypothesis, the neural representation hypothesis, and the processing style hypothesis. Then, we will delve into domain-general object identification ability, commonly referred to as 'o' or 'VG'.

1.1.1 Innate vs. Acquired

Faces play a fundamental role in our social interactions and are ubiquitous in our environment from infancy. In comparison, reading is a skill that emerges and is nurtured during later stages of human development. This stark contrast poses a question: Are our tendencies to recognize faces rooted in genetics? Or does the differentiation in processing faces compared to other objects arise from the type and frequency of our encounters with them, especially considering the prevalence of faces in a child's early environment? The 'innateness hypothesis' (Morton & Johnson, 1991) suggests that humans inherently possess the ability to recognize faces. This ability is considered to be present from birth, and further refined by experiences as one grows (Simion & Giorgio, 2015). Evidence supporting this hypothesis stems from studies demonstrating that newborns typically track faces preferentially over other stimuli in the first few months of life (Johnson et al., 1991; Morton & Johnson, 1991; Simion et al., 2007). This inclination for faces implies that certain brain regions or networks are designated for this function. Indeed, face recognition may be supported by distinct neural pathways in the brains of both humans and monkeys (Tsao & Livingstone, 2008). In contrast, reading is a skill acquired later in life, demanding intentional teaching and effortful learning to master. Considering the timeline of human evolution, written language is relatively new, having appeared about 5000 years ago (Parr, 2011). Therefore, the learning process for word perception is more comparable to the acquisition of specialized knowledge, such as recognizing fingerprints, or birds (Busey & Parada, 2010; Gauthier et al., 2000).

However, findings from electrophysiological studies on monkeys challenge the innateness hypothesis, raising the possibility that face recognition might be a learned rather than an inherent skill (Arcaro et al., 2017; Tanaka & Sengco, 1997). These studies suggest that sensory experiences and context, particularly during early development, have a profound impact on brain development and our ability to recognize faces. The emphasis on familiarity underscores the significance of early experiences and challenges the concept of innate face recognition. Additionally, visual words or text, which are undoubtedly learned to a degree, appear to engage a distinct neural network (Dehaene & Cohen, 2011). This underscores the idea that having a specialized neural pathway does not necessarily imply an innate origin. These open up a question of how the brain develops specificity for recognizing faces (Sunday & Gauthier, 2018; Young, 2018). Studies indicates that similar developmental processes occur during the acquisition of face processing in childhood and when adults acquire perceptual experience with a new visual category (Gauthier & Nelson, 2001).

Although there is ongoing debate about the innateness of face recognition, it is universally accepted that word recognition requires intentional, specialized learning. According to (Centanni et al., 2018), learning to recognize letters may induce changes in the fusiform gyrus which again could potentially decrease an area dedicated to face recognition, known as the fusiform face area (FFA). This aligns with the neuronal recycling hypothesis of Dehaene & Cohen, (2007), suggesting that learning to read might rewire neural pathways previously dedicated to face processing. This model emphasizes how the brain can adapt and repurpose certain neural pathways for different functions, providing insights into the nature of neuronal plasticity (Hernandez et al., 2019; Hernández et al., 2013).

1.1.2 Face and word processing specifications

Cohen et al., (2000) pinpointed a region within the left fusiform gyrus known as the 'Visual Word Form Area' (VWFA), emphasizing its importance for proficient reading, as it exhibited selectivity for written words. The specificity and purpose of the VWFA have been contested (Price & Devlin, 2011; Starrfelt & Gerlach, 2007), however, especially based on the recent evolution of written language. Researchers argue, based on the recycling hypothesis (Dehaene & Cohen, 2007), that cultural inventions such as reading repurpose evolutionarily older circuits. This suggests that the VWFA, while aiding in reading after literacy has been acquired, might also contribute to processing other visual stimuli (A. C. Vogel et al., 2014; see also Price & Devlin, 2003).This overlap hints at possible connections between reading and broader object recognition. Dehaene & Cohen, (2011) provided support for the idea that learning to read partially repurposes VWFA, a brain region originally evolved for recognizing objects and faces. Also, Dundas et al., (2013) posited a correlation between the emergence of face lateralization and reading competence. Similar to the VWFA, the nature of fusiform face area (FFA) located in the fusiform gyrus (Brodmann area 37) has been controversial. Recognized for its specialization in face recognition (Kanwisher et al., 1997), the FFA's functions stretch beyond faces to encompass tasks like word (Dehaene & Cohen, 2011) and object recognition (Rossion et al., 2012). While some underline its innate capabilities for processing biologically significant stimuli (McKone et al., 2012; Owen & Maratos, 2016), others point to evidence of a distributed network that overlaps especially for face processing ,and the influence of expertise in shaping neural configurations. For instance, areas within the face network, including FFA, might be influenced by visual primitives, responding to both faces and other round objects (Bao et al., 2020; Srihasam et al., 2014; Tsao et al., 2006; Yue et al., 2014). Studies in macaques and human infants further illuminate that face specialization materializes over time, suggesting that expertise in face processing, rather than inherent functional specificity, underpins the evolution of the FFA (Gauthier et al., 1999). The expertise hypothesis, championed by Gauthier and colleagues, frames the FFA as an element of a domain-general object recognition system, honed by expertise. This research aligns well with the many-to-many (MTM) theory (Behrmann & Plaut, 2013), which contends visual categories like faces and words are processed by distributed neural networks. These networks involve multiple nodes linked by structural connections and these networks may play a role in both face and word recognition, with a preference for face processing in the right hemisphere and word processing in the left.

1.1.3 Configural and featural processing

How do we process visual words compared to faces? While words are often two-dimensional, high-contrast visuals, faces present as detailed three-dimensional entities. This implies distinct perceptual analysis needs. Yet, given our expertise in recognizing both, it is intriguing to consider if the shared necessity for rapid, expert discrimination leads to overlapping processing mechanisms, despite their inherent contrasting nature. Featural and configural processing are recognized as distinct approaches for visual recognition (Lobmaier et al., 2010; Rossion et al., 2000).

Featural processing involves processing the individual features of a stimulus, whereas configural processing involves processing the relationships among different features (Maurer et al., 2002a). Featural processing is commonly linked with the identification of words or non-face objects (McKone et al., 2007; McKone & Robbins, 2007a; Pelli et al., 2003; Rossion, 2013; Pelli et al., 2003; Farah et al., 1998, 1995). In contrast, configural processing is considered to be pivotal for face recognition. This is underscored by the "processing style hypothesis" which posits that our visual system recognizes faces by focusing on the relationships between features, an approach known as configural processing (Diamond & Carey, 1986).

Configural processing encompasses three main elements: first-order relations, referring to the relative positioning of features; second-order relations, which deal with the distances between these features; and holistic processing, that perceive features as a coherent whole (Maurer et al., 2002a; Roberts et al., 2015). Holistic processing, that frequently regarded as the hallmark of face processing (McKone et al., 2007), which is supported by experimental evidence stemming from three standard paradigms: the inversion task, the part-whole task, and the composite task. The inversion task demonstrates that upside-down faces are harder to recognize than right-side-up ones, suggesting that upright faces are perceived as a whole, while inverted ones may rely on the encoding of individual features (Yin, 1969). The part-whole task shows that we recognize facial features, like eyes, more accurately when they are part of a full, upright face, implying that features are not encoded separately but are a part of a more complex representation (Tanaka & Farah, 1993). The composite task shows that when the top half of a face aligns perfectly with the bottom, even if participants are told to ignore the bottom half, their perception of the top half is influenced. But when misaligned or inverted, this influence weakens, suggesting that the parts of the face halves are encoded not separately but as a part of a more complex representation that encompasses the whole face – or large parts of it (Susilo et al., 2013; Young et al., 1987).

The role of holistic processing in visual word recognition is still under debate, with models like the interactive-activation model positing a hierarchical approach. This approach progresses from basic feature extraction such as lines and curves to letter identification, culminating in the recognition of entire words (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). In these models, while letters within words undergo parallel processing, the word is not considered an indivisible entity (Adelman et al., 2010). However, there is a potential enhancement in the recognition of familiar words by leveraging top-down semantic and linguistic insights on full-word representations, contributing to efficient visual word processing (McClelland & Rumelhart, 1981), though the scale of these influences is still a subject of discussion (Dehaene & Cohen, 2011; Price & Devlin, 2011; For a review, see Feizabadi et al., 2021).

The evidence delineates distinct roles at letter, word, and sentence-level processes in reading (Pelli & Tillman, 2007). implying that a level of whole-word processing in visual word recognition is plausible. Diverging from a linear letter-to-word model, some hybrid theories are suggesting simultaneous featural (letter-focused) and holistic (word-focused) pathways. Advocates of these theories are divided; some propose flexibility in method selection according to situational demands (Allen et al., 1995), while others argue for the distinct contributions of letter and word-based information in visual word identification (Besner & McCann, 2016).

The notion of "quasi-parallel" processing of letters is introduced to describe the rapid efficiency of visual word processing (Cohen, 2003; Cohen et al., 2008), paralleling concepts in face processing expertise, particularly the automatization of the parallel processing of facial components (Richler et al., 2011). Recent definitions of "holistic" in visual word processing research have adopted a milder interpretation, representing a compulsory attention to every component of an object (Wong et al., 2012), or mandatory encoding of/attention to every letter in a word (Ventura et al., 2020). This stance leans towards top-down lexical influences over bottom-up perceptual ones (Ventura et al., 2017, 2020), aligning with the perspective of visual word processing being reliant on rapid parallel letter processing with lexical modulation, rather than processing the word as an inseparable entity. Consequently, this interactive hierarchical perspective opens up possibilities for whole-word effects that might influence visual word recognition, even if the processing is not strictly holistic, raising questions regarding the similarities in the experimental effects between visual word and whole object processing in faces.

There exists some evidence indicating that the three principal indicators of holistic processing—the inversion effect, the part-whole effect, and the composite effect—are also discernible in word processing.

Koriat & Norman (1985) found that words tend to be perceived holistically, especially when they are in their regular orientations. This was referred to as the 'word inversion effect'. However, when words are rotated close to 180 degrees, the holistic reading pattern declines, and reading shifts towards recognizing individual features of the word. Interestingly, the speed of recognition remains constant for words of different lengths from 0° up to around 60°. Yet as the degree of word rotation increases beyond 60°, recognition times increase. When the rotation exceeded roughly 120°, no additional complications arose with further disorientation. All demonstrated consistent length effects, suggesting a letter-by-letter reading method or feature-based reading (Koriat & Norman, 1985).

The inversion effect is not confined to just faces and words. Experts on specific dog breeds struggle to identify those breeds when inverted, just as in the face-inversion effect and word-inversion effect (Diamond & Carey, 1986; Yin, 1969). Experts identifying individual budgerigar birds, for example, had difficulty recognizing both inverted faces and birds, while novices only had issues with faces (Campbell & Tanaka, 2018). These results support the idea that the face-inversion effect might also apply to expert recognition of visually similar objects, although this is contested by some research (for example, Robbins & McKone, 2007).

In visual word processing, the word superiority effect is akin to the whole-face advantage or part-whole effect (Feizabadi et al., 2021). The word superiority effect shows that individual letters are recognized more efficiently when placed within real words as opposed to when they appear alone or within a nonsensical sequence of letters (Reicher, 1969). Interestingly, real words may not be necessary to generate an advantage in identifying letters. There is also a pseudo-word superiority effect. A pseudo-word superiority effect exists, where a letter in a pronounceable, yet pseudoword is more easily recognized than in a non-pronounceable string (Baron & Thurston, 1973). Further, a study involving Italian, a language with greater orthographic regularity than English or French, demonstrated advantages for both real and pseudo-words (Ripamonti et al., 2018). In a study by Wong et al., (2011), the 'composite word effect' was observed. Participants were asked to match either the right or left halves of a pair of words, ignoring the opposing halves. In congruent trials, both halves of the word pairs matched, while in incongruent trials, one half matched while the other did not. The results indicated that participants performed better during congruent trials, suggesting that they were inadvertently influenced by the halves they were supposed to ignore. Notably, this effect was more noticeable for words that participants were highly familiar with, such as those from their primary language as opposed to their secondary language, and with real words rather than pseudowords.

The "expertise hypothesis" proposes that, given adequate practice and familiarity, configural processing can be applied to objects beyond just faces (Carey, 1992; Diamond and Carey, 1986; Gauthier and Tarr, 1997; Meadows, 1974). Expertise in a particular category often manifests as increased configural processing for that category (Bukach et al., 2006a; Gauthier & Bukach, 2007a; Gauthier & Tarr, 2002a; Wong et al., 2011, 2019). Some studies indicate that proficient readers tend to utilize configural processing more than those who are less skilled at reading (Ventura et al., 2020; Wong et al., 2011, 2012, 2019). However, configural processing might not necessarily serve as an indicator of broad visual expertise. The influence of expertise on varying processes might depend on the amount of information each process offers (Hsiao & Cottrell, 2009; L. Zhang & Cottrell, 2005) and the specific requirements of a task (Zhou et al., 2012). Expertise may primarily enhance configural processing when configurations form the key information for identifying those categories. However, when individual features are essential for recognition, expertise might then strengthen the proficiency of processing these features for specific categories. For example, Zhou et al. (2012) found that art students, experts in drawing faces, exhibit diminished holistic face processing. This coud be because, when these students sketch faces, they focus intently on the features of the face. Similarly, Tso et al. (2014) observed that Chinese readers with limited writing experience showed more holistic processing. In contrast, proficient Chinese character writers displayed a decreased holistic approach. The assumption is that the act of writing refines one's ability to discern characters based on their distinct features. This finding is consistent with their earlier study (Van Yip Tso et al., 2012), where Chinese first graders, that were still new to writing, exhibited enhanced holistic processing in recognizing Chinese characters compared to their non-Chinese-speaking counterparts who were not exposed to the local Chinese curriculum. As these Chinese students advanced in grades, their holistic processing effect diminished.

1.2 Domain-general mechanism (*o/VG)*

Research on object recognition has traditionally emphasized domain-specific mechanisms, especially in the context of face recognition (Duchaine & Nakayama, 2006). These mechanisms are believed to apply to a restricted number of object categories. In contrast, domain-general mechanisms apply to a broader range of object categories. Most of the earlier research focused on the difference between face and object recognition, often operating under the assumption that a universal mechanism was responsible for the recognition of all non-face objects. However, this perspective is being increasingly challenged. For example, Gerlach, (2009) introduces pre-semantic account of category-effects (PACE), a theory explaining the brain's distinct processing of object categories like natural versus artifacts at pre-semantic levels. PACE outlines visual object recognition in two phases: first, the integration of shape components into detailed shape descriptions, and second, matching these descriptions with stored representations in long-term memory. PACE argues that structural similarity between objects affects both shape configuration and selection in opposite ways. High structural similarity is beneficial for shape configuration but harmful for selection, and vice versa. The effects of structural similarity also depend on specific task requirements, such as perceptual differentiation and stimulus degradation. Also, distinct neural pathways have been identified for recognizing animals versus tools (Chao et al., 2002), large objects versus small ones (Konkle & Oliva, 2012a), and curvilinear objects as opposed to rectilinear ones (Nasr et al., 2014; Yue et al., 2014). Parallel to these neural distinctions, behavioral differences across various object categories have been observed. Intriguingly, these behavioral observations often mirror the variations seen in neural responses (M. A. Cohen et al., 2014, 2015). Furthermore, the correlation strength among object recognition tests appears to fluctuate, suggesting that the role of domain-general mechanisms might be different depending on the object category (McGugin et al., 2012). Given these complexities, it is natural to question the significance of domain-general mechanisms in individual object recognition capabilities. It might be that the correlations observed in object recognition tests reflect broader cognitive attributes such as IQ or attentiveness rather than any inherent capability related to object recognition. However, an increasing body of research has pointed towards the existence of a domain-general ability for object recognition that operates independently of general intelligence and other cognitive factors. This capability displays correlations across various visual tasks and categories (Hendel et al., 2019; Richler et al., 2019a).

For example, individuals who excel at recognizing certain visual objects, such as faces, also perform well in recognizing other objects, like cars (Geskin & Behrmann, 2018). Richler et al. (2019) found that about 89% of the performance variance related to new object categories could be explained by a common factor, "o," suggesting a widespread object identification capacity across categories. Furthermore, fingerprint examiners not only outperform novices in fingerprint-comparison (i.e., domain-specific; Busey & Vanderkolk, 2005; Tangen et al., 2011), but also on face-comparison tasks (i.e., domain-general; Phillips et al., 2018). Notably, Hendel et al. (2019) found that "super-recognizers," individuals with extraordinary face recognition skills, outperformed controls across all visual categories, including words, objects, and faces. Super-recognizers also score above average on primate-face and fingerprint-comparison tasks (Towler et al., 2021). This result implies that the mechanisms bolstering superior face processing in super-recognizers could also enhance their performance in other visual categories, reinforcing the notion of a general factor in the visual domain. To describe this phenomenon, Hendel et al., (2019) introduced the term "factor VG." Adding to this, Maratos et al. (2022) found that familiar stimuli, whether words or objects, share some common underlying cognitive mechanisms, suggesting a domain-general object recognition mechanism. Overall, the accumulated evidence pointing towards a universal object identification capability (o or VG) suggests the possibility of a broad visual recognition skill in expert groups.

1.3 Developmental disorders: Domain generality/specificity debate

Researchers have explored specific developmental disorders, such as developmental dyslexia and developmental prosopagnosia, in order to understand the underlying processes involved. These studies have often emphasized the specificity of these disorders, investigating if deficits in one domain, like face recognition, might influence another, such as reading development. Developmental dyslexia is primarily characterized by significant difficulties with reading despite normal intellectual capacity and adequate educational opportunities (Shaywitz, 1998). Developmental prosopagnosia, on the other hand, is marked by pronounced issues in face recognition, attributable to underdeveloped visual mechanisms (Duchaine & Nakayama, 2006). Both disorders occur in the absence of brain injury (Duchaine & Nakayama, 2006; Shaywitz, 1998). Recognition problems in dyslexia and prosopagnosia might be confined to words and faces respectively, supporting domain-specific accounts (Dehaene & Cohen, 2011; Kanwisher, 2000; Kleinschmidt & Cohen, 2006; Rhodes et al., 2004; Robotham & Starrfelt, 2017; Yovel & Kanwisher, 2004). However, they might also extend to other visual categories, suggesting at least some common underlying factors (Gauthier et al., 2014; Hills et al., 2015; Sigurdardottir et al., 2015, 2018, 2019).

1.4 The Boundaries of Recognition Impairments in Dyslexia and Prosopagnosia

While dyslexic readers have been shown to have visual recognition problems (Brachacki et al., 1995; Huestegge et al., 2014; Jozranjbar et al., 2021; Mayseless & Breznitz, 2011; Sigurdardottir et al., 2015), their visual processing deficits may not generalize to all object types (Gabay et al., 2017; Sigurdardottir et al., 2018), making the specificity of word recognition problems in dyslexia and their relationship to face and object recognition problems an open question. The association between dyslexia and face recognition has yielded mixed results. While some studies have reported impaired face recognition in dyslexic readers (Collins et al., 2017; Gabay et al., 2017; Sigurdardottir et al., 2015, 2018, 2019), other studies have contradicted this finding (Holmes & McKeever, 1979; Jozranjbar et al., 2021; Rüsseler et al., 2003; Smith-Spark & Moore, 2009). Kühn et al. (2020) posited that some individuals with dyslexia have difficulties with face recognition, while others do not, which may reflect the heterogeneous nature of dyslexia (Á. Kristjánsson & Sigurdardottir, 2023).

Regarding developmental prosopagnosia (DP), an extensive review of 238 cases by Geskin & Behrmann, (2018) found that only around 20% of DP cases exhibited normal object recognition abilities. This led to their hypothesis that DP might be attributed to a broader cognitive—perceptual deficit that is not limited to face recognition alone. Notably, the same review concluded that word recognition was not impaired in DP. However, a contrasting perspective is provided by Gray & Cook, (2018). They suggest that developmental agnosias, whether related to faces (DP), objects (developmental object agnosia; DOA), or bodies (developmental body agnosia; DBA), might be distinct neurodevelopmental conditions. These conditions can often appear concurrently, especially as DOA and DBA incidences seem to be elevated in those with DP. Gray & Cook, (2018) argue that such overlaps might be due to shared genetic or environmental factors. Hence, while many individuals with DP might also struggle with other recognition challenges, there exists a group that faces specific challenges related to face recognition alone. Note, however, that this co-occurrence is not necessarily inevitable, and some individuals may have selective difficulty with faces (Gray & Cook, 2018). But often, visual object recognition is impacted in developmental prosopagnosia, though the severity may not match that of face recognition impairment (Gerlach et al., 2016).

The potential disadvantage in word recognition among individuals with developmental prosopagnosia (DP) continues to be a subject of debate. Early studies suggested that DP does not significantly impact reading speeds when compared to neurotypical group (Burns et al., 2017; Rubino et al., 2016; Starrfelt et al., 2018). However, the relatively small average sample size of around 10 DP individuals in these studies could have constrained their statistical power to discern subtle differences.

Addressing this statistical limitation, Burns & Bukach, (2021) combined data from these preceding studies. Their analysis unveiled a correlation where DP individuals exhibited marginally slower reading speeds compared to the neurotypical group. Yet, this correlation was questioned by Gerlach & Starrfelt, (2022), as they failed to replicate Burns and Bukach's result.

1.5 Featural vs. Configural Processing in dyslexia and prosopagnosia

The link between visual word, face, and object processing might be shaped by the nature of the tasks and stimuli involved, requiring either similar or distinct types of visual processing. This perspective suggests possibility that the perceptual differences found in developmental dyslexia and prosopagnosia might be more related to specific processes rather than distinct categories. There is debate over what the key processes involved in reading and face recognition are, where some suggest that word recognition relies largely on features (Pelli et al., 2003), while configural processing is a key component of face recognition (McKone et al., 2007; McKone & Robbins, 2007a; Rossion, 2013). However, a counterargument suggests that configural face processing might depend on domain-general mechanisms, which are tapped into for object perception through a process of accumulating visual expertise (e.g., Bukach et al., 2006; Gauthier & Bukach, 2007; Gauthier & Tarr, 2002; but see McKone & Robbins, 2007). It has been suggested that gaining expertise in a visual category could enhance configural processing for objects within that category (Wong et al., 2009). Therefore, if individuals with dyslexia and prosopagnosia struggle with acquiring visual expertise (Barton et al., 2019; Sigurdardottir et al., 2017), their recognition issues could be rooted in configural processing impairments.

Studies on developmental prosopagnosia have produced inconsistent results regarding its association with configural face processing deficits. Some reports highlight impaired configural processing in prosopagnosia (Avidan et al., 2011; Liu & Behrmann, 2014; Palermo et al., 2011), while others contradict this finding (Biotti et al., 2017; Le Grand et al., 2006; Susilo et al., 2010; Ulrich et al., 2017). In some instances, individuals with prosopagnosia performed worse than control groups on both featural and configural face processing tasks (Gerlach et al., 2016; Le Grand et al., 2006; Yovel & Duchaine, 2006).

Meanwhile, dyslexic readers have demonstrated configural processing comparable to controls for face recognition (N. Brady et al., 2021; Sigurdardottir et al., 2015, 2019), or even increased for word recognition (N. Brady et al., 2020; R. V. Tso et al., 2021; R. V. Y. Tso et al., 2020). However, Conway et al., (2017) found that college students with dyslexia, while sensitive to orthographic regularity, might rely more on featural processes than their non-dyslexic peers. The diverging results may be due to different manipulations of configural processing, such as the distinction between second-order configural processing and holistic processing (Maurer et al., 2002b; Roberts et al., 2015).

1.6 Challenges in Drawing Conclusions from Developmental Disorders

Drawing inferences about typical cognitive development or architecture from developmental disorders like dyslexia and prosopagnosia can pose challenges. Some perspectives, such as those provided by Bishop (1997) and D'Souza & Karmiloff-Smith (2011), propose that these developmental deficits are manifestations of unique brain developmental paths in the affected individuals. An exploration into these disorders aiming to understand normal cognition might potentially overlook the dynamic complexity of the developing brain, marked by the interplay of numerous emergent systems and functional plasticity (Moses & Stiles, 2002). These accounts highlight that the brain of an infant initially exhibits a high degree of interconnectedness, with neural networks becoming increasingly specialized or modularized over time. It is therefore conceivable that individuals with developmental disorders might employ different cognitive and neurological processes even when their behavioral scores align with the typical range, marking an atypical developmental path. Neurodevelopmental disorders might not mirror a continuous process of modularization. Indeed, modularity could signify the end state of development, rather than its point of origin (D'Souza & Karmiloff-Smith, 2011). Consequently, face processing in people with prosopagnosia might start and continue to develop abnormally. Starrfelt and Robotham (2018) posit that the absence of a distinction between face and object recognition in prosopagnosia does not necessarily have implications for the cognitive architecture of the face processing system in typically developed adults. Furthermore, any findings related to developmental dyslexia must be considered with caution. Evidence of temporal, motor, attentional, auditory, and visual dysfunction (De Martino et al., 2001; Farmer & Klein, 1995; Giofrè et al., 2019; Goswami, 2011; Kristjánsson & Sigurdardottir, 2023; Norton et al., 2015; Reid, 2018; Valdois et al., 2004; Ziegler et al., 2010) suggest that dyslexia is a diverse deficit. As a result, investigating variations in face, word, and object recognition in neurotypical populations could provide valuable insights that complement studies of neurodevelopmental conditions.

1.7 VWM and visual object recognition

The distinction observed between expert categories and the recognition of other less familiar objects may be attributed, in part, to the capabilities of our visual working memory (VWM). VWM is a crucial cognitive system that provides us with the capacity to temporarily store and manipulate visual information (Alvarez & Cavanagh, 2004a; Awh et al., 2007a; A. Kristjánsson, 2006; Luck & Vogel, 1997a; E. K. Vogel et al., 2001a; W. Zhang & Luck, 2008a). The exploration of VWM has primarily been done by research focused on simple, parametrized object classes such as geometric shapes and colours. However, the impact of VWM's limitations on the more complex stimuli found in the real world, which are not merely combinations of basic features, remains less thoroughly investigated.

Recent studies have demonstrated that VWM is not simply a static store; it is flexibly adapted based on the nature of stimuli, expertise, and the strategy deployed in processing information. The meaningful perception of stimuli might prompt the use of extra resources. This can be influenced by factors like the actual size of objects (Konkle & Oliva, 2012b; Long et al., 2016a, 2019a), anticipated related object classes (Kaiser et al., 2015a; O'Donnell et al., 2018a), and proficiency in specific categories (Curby et al., 2009a; Curby & Gauthier, 2007a; Xie & Zhang, 2017a). By encompassing both behavioral and EEG data, Asp et al. (2021) found that the ability to derive meaning from a stimulus significantly bolsters its retention in visual working memory. For instance, the same stimulus is more effectively recalled when it is perceived as a face than when it is not. Also, Brady et al. (2016) revealed that with extended encoding time, VWM's capacity for real-world objects improved. Contrarily, such a benefit was absent when participants encoded simpler stimuli like colors. They attributed this to the extra conceptual information that real-world objects have over simple stimuli. However, there is an alternative perspective suggesting that the Visual Long-Term Memory (VLTM) system, known for its substantial capacity, might play a role in encoding real-world objects (Cowan, 1988; Quirk et al., 2020). Similarly, Asp et al. (2021) reported an amplified contralateral delay activity (CDA) – a neural indicator believed to signify active storage in VWM, for recognizable images versus unrecognizable images. Asp et al. (2021) hypothesized that extended encoding time was especially advantageous for real-world objects. This prolonged time might allow for more comprehensive processing of these items, potentially tapping into prior knowledge about them, which can be utilized to maintain them in active memory. Such an approach would be less relevant for simpler stimuli, like colors, as highlighted by Brady et al. (2016). However, some studies found that extended encoding time enhances working memory performance for both basic stimuli and real-world objects (Li et al., 2020a; Quirk et al., 2020). Brady & Störmer, (2020) criticized inconsistent foil selections across studies, noting that using highly distinct foils for colors and random ones for objects, as was done in Li et al., (2020) and Quirk et al., (2020), can skew results. They emphasized that the similarity between a target and its foil directly impacts memory performance. To accurately gauge VWM capacity, foils should be notably different from targets, otherwise the diminished performance is influenced by foil similarity rather than VWM capacity. In a notable innovation, Brady & Störmer, (2020) employed deep convolutional neural networks to enhance foil dissimilarity, finding that when foils are optimally selected, memory retention is typically stronger for objects than colors.

The literature presents differing perspectives on the impact of stimulus familiarity on VWM. While some studies, like Jackson & Raymond, (2008) and Ngiam et al. (2019), suggest that accumulated experiences with visual categories throughout our lives can bolster the capacity or precision of our VWM for these categories, other studies, such as Chen et al. (2006), Olson & Jiang, (2004), and Pashler, (1988), suggest that familiarity with object classes does not necessarily translate to enhanced VWM performance. Xie & Zhang, (2017, 2018) found no significant boost in memory capacity when using Pokémon figures, finding only effects upon memory consolidation. Contrarily, using alphabet letters, Ngiam et al. (2019) observed enhanced memory capacity resulting from stimulus familiarity. They further argued that the familiarity gap between first-generation (known) and recent-generation (unknown) Pokémon might be too subtle to influence memory capacity. To elucidate, Ngiam et al. (2019) contrasted familiar English letters with unfamiliar novel characters, marking a stark difference in familiarity, which led to the observed variation in memory capacity.

Differences in VWM capacity when confronted with real-world object classes and simple object classes may also arise due to distinct processing strategies, such as featural and configural processing. Simple object classes are generally processed featurally, but real-world objects, with their complex and relatively inseparable features, may be stored and remembered more holistically. Expertise may foster the development of efficient visual processing strategies, such as configural processing (Bukach et al., 2006a; Gauthier & Bukach, 2007a; Gauthier & Tarr, 2002a; but see McKone & Robbins, 2007a). Starrfelt et al. (2013) found that the Word Superiority Effect (WSE), especially for simple short words, is evident in vocal reaction times. A part of this superiority likely arises from the accelerated visual processing of words as opposed to individual letters. This aligns with past observations of the WSE and the idea that top-down connections might boost the processing of letters within words. In contrast, single letter processing might depend more on bottom-up signals. Curby et al. (2009) suggest that configural processing allows for the merging of various object features into "chunks" or units, reducing the demand on VWM. Consequently, this holistic approach enables experts to optimize the inherently limited storage capacity of the VWM system. Nevertheless, Norris & Kalm, (2018) cautioned that the definition of a 'chunk' remains somewhat nebulous, with some defining it as compressed codes (T. F. Brady et al., 2009a), and others viewing it as a cue for retrieving information from long-term memory (Huang, 2011a; Hulme et al., 1991a, 1997a; Jones & Farrell, 2018a; Kahneman et al., 1992a). Regardless of the approach, both agree that the contents of a chunk cannot be accessed directly but must be retrieved from long-term memory. This retrieval process can result in a decrease in response speed, despite the increased number of remembered features (Huang & Awh, 2018a).

The distinction between featural and fully integral representations may not be strict; they might exist on a continuum. For instance, Brady et al. (2013) found that distinct features of real-world object classes were forgotten at different rates, implying they were not compressed into a single information unit. Seen from this perspective, a VWM chunk could be viewed as a collection of highly integrated features, likely due to learned associations, which can be distinguished from less associated features. However, it is important to note that the effectiveness of chunking can vary based on task requirements (Jackson & Raymond, 2008a; Jozranjbar et al., 2023). Markov et al. (2021) reported that features of real-world object classes are represented independently and are not always stored as fully bound units. In summary, visual working memory is influenced by numerous factors. These include the nature of the stimuli, familiarity, processing strategy, expertise, and the possibility of chunking information. However, the exact mechanisms, representations, and influences remain an active area of exploration.

2 Aims

This thesis aims to examine the fundamental principles governing visual object recognition, investigating how certain visual recognition skills are associated or distinct and the role these relationships play in the onset of developmental disorders, specifically dyslexia.

In our investigation of developmental dyslexia, we tested whether dyslexic readers have difficulties in non-word visual objects. We further assessed if these potential recognition difficulties are limited to distinct visual categories, like faces or houses, and whether they are restricted to specific processes, such as featural or configural processing. Given that individuals with developmental disorders might exhibit visual recognition abilities akin to the neurotypical population but may operate on different cognitive and neurological grounds, a singular focus on these disorders might not capture the full spectrum of brain development intricacies. This understanding led us to also incorporate neurotypical individuals in our investigations. We aim to understand whether certain visual categories, such as faces, houses, or words, and processes like featural or configural, rely on overlapping or distinct recognition mechanism. Additionally, we study visual working memory (VWM), aiming to determine its significance and potential constraints on these visual categories and processes, investigating if VWM's role varies across these categories or processes. Moreover, we investigate the relationship between visual working memory (VWM) precision for simple features and memory accuracy for featural or configural information of unfamiliar faces, houses, and pseudowords. We were particularly interested in whether the associations of VWM differ based on types of categories or processes.

In sum, this thesis combines findings from both dyslexia group and neurotypical groups, along with the dynamics of VWM, to offer a comprehensive insight into visual object recognition. Through this multifaceted approach, our goal is to elucidate the intricate interactions among diverse visual categories and processes and comprehend how elements like VWM may account for the interconnectedness between these categories.

2.1 Paper I: Featural and Configural Processing of Faces and Houses in Matched Dyslexic and Typical Readers (Neuropsychologia)

The first paper in this thesis investigates the role of high-level visual processing in dyslexic readers and the possible category- or process-specific visual problems they might face. By focusing on face and house processing, this research aims to address the following questions: (a) Do dyslexic readers have problems with recognizing non-word visual objects? (b) Are these potential problems confined to specific visual categories (like faces or houses)? (c) Are the problems faced by dyslexic readers tied to specific processes (such as featural or configural)?

2.2 Paper II: Using representational similarity analysis to reveal category and process specificity in visual object recognition (Cortex)

The primary objective of the second study is to explore the organizing principles of visual object processing. More specifically, it aims to discern whether different visual categories (like faces, houses, and words) and visual manipulations (such as featural, configural) depend on unique or shared mechanisms. It probes whether some categories rely on a specific mechanism while others employ a common, domain-general mechanism. The study also explores whether superior performance in one category is linked to inferior performance in another, indicating a potential tradeoff, and whether the featural versus configural process is universal or category-specific. This research contributes to the understanding of individual differences in face, word, and object recognition in neurotypical populations, providing insights that complement studies of neurodevelopmental conditions.

2.3 Paper III: The impact of visual working memory constraints on object recognition (Under review)

The aim of the third study is to fill a gap in previous research on visual working memory (VWM), which has predominantly targeted simple stimulus classes, leaving the impact of VWM constraints on complex stimuli less explored. Using an individual differences methodology, we examine the relationship between VWM precision for simple stimuli and VWM accuracy for featural and configural information in unfamiliar houses, faces, and pseudowords. We attempt to ascertain the extent to which memory accuracy for these objects is linked to, and potentially explained by, the VWM precision as independently assessed for simple stimuli.

3 Materials and methods

Across all three papers, we employed a similar set of stimuli for faces and houses. While the first study exclusively used face and house stimuli, the subsequent papers expanded the stimuli set to include pseudowords. Consistently, we utilized a measurement of behavioral accuracy across all studies. This metric enabled us to evaluate how well participants could retain visual information in their short-term memory. The format for object recognition task was a delayed match-to-sample with two alternative forced choices in all three papers. In the first study, the task was not structured in blocks. Participants were shown a sample, which hinted at the type of object they would encounter next. However, the processing style manipulation remained undisclosed until the match and foil were presented. This arrangement made it possible for participants to potentially adjust their strategies based on the stimulus category and not the processing information. To address this, our later studies introduced a more structured approach by dividing the task into clear blocks. This ensured that participants were informed of the type of combinations they would see next, like a featural face or a configural pseudoword, for instance.

From an analytical standpoint, our work predominantly relied on Representational Similarity Analysis (RSA) and Bayesian Multiple Regression Analysis. RSA, which was central to the first two papers, offers an in-depth exploration of data patterns. RSA involves the comparison of observed data patterns (reference models) with predicted data patterns (conceptual models). The reference models were generated from correlation matrices of accuracy derived from different stimuli and conditions. These matrices serve as reference points and are set against conceptual models that hypothesize potential data patterns, emphasizing aspects like stimulus and process. Through RSA, we have extracted patterns that could otherwise go unnoticed, aiding in understanding the interdynamics between various task conditions and their underlying cognitive strategies. Bayesian multiple regression Analysis, which was featured in the last two papers, enriched our analytical framework by offering a probabilistic perspective. This method is pivotal for quantifying uncertainties about our primary parameters.

In the following sections, I will elaborate on the specific materials and methodologies underpinning each paper.

3.1 Paper I: Featural and Configural Processing of Faces and Houses in Matched Dyslexic and Typical Readers (Neuropsychologia)

We recruited 34 people who reported a previous dyslexia diagnosis and 34 matched self-reported typical readers. Participants provided information about previous diagnoses, including dyslexia, ADHD, dyscalculia, autism, hearing impairment, and other language problems. Three questionnaires were administered: the Icelandic version of the Adult Reading History Questionnaire (ARHQ-Ice), and Behavioral Evaluation Questionnaire for Adults I and II (assesses symptoms of ADHD). Participants then performed a visual task measuring configural and featural processing of faces and houses, a lexical decision task evaluating the effect of word length, and their reading was assessed with the IS-FORM and IS-PSEUDO tests.

3.2 Paper II: Using representational similarity analysis to reveal category and process specificity in visual object recognition (Cortex)

The study consisted of two experimental sessions. In the first session, participants (N = 97) evaluated their ability to recognize visual words and faces using a 5-point Likert scale and provided demographic details. In the second session, participants then took part in a visual recognition task, in which they were required to remember featural or configural information for faces, houses, and pseudowords.

3.3 Paper III: The impact of visual working memory constraints on object recognition (Under review)

Participants were recruited from a pool of people who had completed two sessions of the second study (see section 3.2). Those who passed attention checks in session 2 were invited to participate in session 3, which involved a visual working memory task with simple stimuli (oriented lines).

4 Results

The studies' results are explained in more detail in the papers attached to the thesis, and here we briefly describe the most important results of each of the studies and the main conclusions from each.

4.1 Paper I: Featural and Configural Processing of Faces and Houses in Matched Dyslexic and Typical Readers (Neuropsychologia)

In this study comparing dyslexic and typical readers in object recognition tasks, both groups had similar reaction times, but dyslexic readers were less accurate at recognizing houses, both featurally and configurally. However, no significant differences in accuracy were observed between the groups in featural or configural processing of faces. A 2 x 2 x 2 repeated measures ANOVA was conducted to analyze the interaction of group, stimulus, and process. The main effect of the group was not significant, and there were no main effects of stimulus and process. The interaction of stimulus and process was significant, showing that featural processing of houses was less accurate than featural processing of faces, and configural processing of houses was more accurate than configural processing of faces. No other significant interactions were found. The lack of main effects and interactions with the group factor was unexpected, but we suggested that visual recognition problems may be influenced by educational level. We found that group differences were larger for houses compared to faces, but only for participants with lower educational levels.

Univariate methods can be limited in detecting informative data patterns, as they focus only on group averages. To address this issue, we used representational similarity analysis (RSA) to assess the structure of information representation in dyslexic and typical readers. We developed reference models for both groups and created three conceptual models based on predicted patterns for stimuli, processes, and difficulty levels. Stimulus type (face vs. house) was the dominant pattern for both groups, but the representation of processes (featural vs. configural) varied between them. Typical readers' performance was more consistent when the process was constant, while dyslexic readers seemingly relied on a single process that was indistinguishable for supposed featural and configural trials. This pattern was observed for both faces and houses.

4.2 Paper II: Using representational similarity analysis to reveal category and process specificity in visual object recognition (Cortex)

This study used representational similarity analysis (RSA) to estimate the separability of visual object recognition mechanisms. Six conceptual models were preregistered based on possible predicted patterns: category-specific, cost, fine-tuning, process, holistic expertise, and time models. The holistic expertise model was excluded due to unreliable data. Five remaining models and three additional ones (face specialization, word specialization, and combination models) were considered.

1. Category model: Assumes greater correlation between performance in same-category blocks than between-category blocks.
2. Process model: Assumes performance in blocks with the same manipulation (featural or configural) should best predict performance in other blocks with the same manipulation.
3. Cost model: Assumes some category specificity and negative links between word and face processing.
4. Fine-tuning model: Assumes some category specificity and positive links between word and face processing.

5. Time model: Assumes performance drifts or fluctuations over time, with positive correlation between performance in adjacent blocks.
6. Face model: Assumes that faces tap into domain-specific processes, that are distinct from house and word processing.
7. Word model: Assumes that words tap into domain-specific processes that are distinct from face and house processing.
8. Combination model: Assumes high association between blocks only when they share both category and process.

These models were compared to the reference model to determine the best fit for the observed data. The reference model is a 24 x 24 correlation matrix where each cell represents the correlation of the accuracy of two blocks of the visual recognition task. The study used Bayesian multiple regression to identify which model, or combination of models, best predicted the observed data. The best-fitting model included the time model, the face model, and the combination model. These predictors were found to play a role in producing the observed data patterns.

Using Bayesian model averaging, extreme evidence for the inclusion of the time and combination conceptual models was found, while moderate evidence supported the inclusion of the face model. Other models, such as process, category, cost, and word models, were less relevant. The results strongly support the combination model, which assumes high association between blocks only when they share both category and process. An exploratory multidimensional scaling on the reference model showed data separation in accordance with the face model and certain category-process combinations clustering together, in alignment with the combination model.

4.3 Paper III: The impact of visual working memory constraints on object recognition (Under review)

In paper III, a Bayesian multiple regression analysis was conducted to determine the best model predicting visual working memory (VWM) precision using the accuracy of the six conditions (featural and configural faces, houses and pseudowords) as predictors. The results show that the best model for predicting VWM precision of simple lines was the one with configural house accuracy as the only predictor. Bayesian model averaging confirmed the importance of configural house performance for predicting VWM precision. However, no evidence was found to suggest that models incorporating accuracy for featural faces, houses, and words, as well as configural faces and words, predict VWM precision for simple stimuli more effectively than models that do not include these predictors.

5 Discussion

Further detailed discussion of the studies described here is can be found in the papers attached to the thesis.

5.1 Paper I: Featural and Configural Processing of Faces and Houses in Matched Dyslexic and Typical Readers (Neuropsychologia)

A major focus of the research in this thesis was to determine whether dyslexic readers have difficulty recognizing non-word visual objects, and if such challenges are confined to written words (domain-specific) or are also applicable to other types of visual stimuli (domain-general). The results indicated that dyslexia-related visual processing issues are not limited to the realm of written words but can permeate other visual domains. Interestingly, there were hints that difficulties did not apply uniformly across all categories of visual stimuli, with dyslexic participants struggling more with the recognition of houses compared to faces. This aligns with previous research suggesting that dyslexic readers may have difficulties recognizing non-face objects (Brachacki et al., 1995; Huestegge et al., 2014; Sigurdardottir et al., 2015). Nevertheless, contrasting findings have emerged from other studies, suggesting possible variations due to different methodologies or the influence of external factors like participants' education levels (Gabay et al., 2017; Sigurdardottir et al., 2018). Also, there is a lack of consensus in existing literature regarding impairments in face processing among dyslexic readers. Some research reports have found such impairments (Collins et al., 2017; Gabay et al., 2017; Sigurdardottir et al., 2015, 2018, 2019), while other studies have not (Brachacki et al., 1995; Holmes & McKeever, 1979; Rüsseler et al., 2003; Smith-Spark & Moore, 2009).

The observed dyslexic readers' struggles with non-face object recognition, specifically with houses but not faces, challenge the visual expertise hypothesis in dyslexia (Lieder et al., 2019; Sigurdardottir et al., 2017, 2018), which posits that recognition difficulties should escalate with more frequently encountered visual categories. The study acknowledges the complexity and heterogeneity of object recognition issues in dyslexia, advising against generalizing its findings to all non-face object recognition. More research incorporating a broader range of visual categories is needed to provide a comprehensive understanding of object recognition challenges in dyslexic readers.

The study also aimed to identify whether the visual problems in dyslexia were confined to featural or configural processing. The investigation yielded no significant group differences in mean accuracy for either type of processing. However, representational similarity analyses (RSA) revealed distinct recognition strategies between dyslexic and non-dyslexic readers. While typical readers appeared to deploy separate featural and configural processes for object identification, dyslexic readers seemed to rely on a single process for the same task, indicating that dyslexic readers have more access to a domain-general mechanism than specialized mechanisms for distinguishing between features and configurations. This could be experience-dependent, reflecting problems with acquiring expertise (Brachacki et al., 1995; Cao et al., 2019; Lieder et al., 2019; Sigurdardottir et al., 2017). However, the lack of a process by education interaction in the RSA results suggests that processing differences between dyslexic and typical readers may not be related to differences in reading experience. A further possibility is that dyslexia might arise from fewer available intact visual processes. Clarifying these visual processes is a promising avenue for future research.

5.2 Paper II: Using representational similarity analysis to reveal category and process specificity in visual object recognition (Cortex)

This study aimed to examine whether visual object recognition relies on dissociable or shared mechanisms by evaluating participants' recognition of unfamiliar faces, houses, and words based on features or configurations. The study found that visual recognition accuracy for one object category often predicted performance for another, suggesting a shared component for object recognition, as proposed by previous studies like Hendel et al. (2019) and Richler et al. (2019). However, this was not universally true. In line with previous studies that argued for the uniqueness of face processing (Kanwisher et al., 1997; Robbins & McKone, 2007; Tanaka & Farah, 1993; Young et al., 2013), the authors found evidence suggesting that face recognition employs specific mechanisms different from those supporting word and house processing.

The study also explored the role of visual expertise and found that it did not necessarily moderate the shared variance between object categories. For instance, recognition accuracy for faces and words (common objects of expertise) was not more correlated with each other than with accuracy for houses (generally not an expertise category). The study concluded that visual expertise for faces and words may therefore be relatively independent.

Additionally, the study addressed the idea of separable processes of features and configurations. Some blocks where configural information was manipulated did not necessarily predict performance in other configural blocks and vice versa for featural blocks. This led to a proposal that some combinations of manipulation (featural/configural) and category (unfamiliar houses/words/faces) are "special", possessing distinct variance not shared with other combinations.

The study had some limitations. The inherent differences between different object categories could have affected the task demands. The decision to use pseudowords instead of real words for practical reasons was acknowledged as a potential shortcoming. Some of the manipulations between categories were also different, which might have influenced the results.

Despite these limitations, the study concluded that these results provide important insights into the mechanisms of object recognition and their potential relevance for neurodevelopmental conditions, such as developmental prosopagnosia, developmental dyslexia, and developmental object agnosia. This may also be relevant for conditions such as autism and Williams syndrome, which show specific face processing patterns. The results also indicate the potential for representational similarity analyses (RSA) as a valuable tool for understanding visual representations in neurodevelopmental conditions.

5.3 Paper III: The impact of visual working memory constraints on object recognition (under review)

In this study, we explored the relationship between visual working memory (VWM) precision for simple features and memory accuracy for featural or configural information of unfamiliar faces, houses, and pseudowords.

Our initial hypothesis was based on the idea that VWM precision might show a stronger association with memory for less familiar object categories such as houses, especially when remembering their individual features may demand encoding a greater amount of information. Contrastingly, we believed that expertise object classes like faces and words, which may require encoding fewer bits or chunks, would impose a lesser strain on VWM. Consequently, the memory of these object classes would not be strongly associated with VWM for simple stimuli.

All factors in the object recognition task, including memory accuracy for features in houses, faces, and pseudowords, and the configuration of houses, faces and pseudowords, were associated with VWM precision. Accuracy for these objects may therefore in part reflect individual variances in VWM precision or other indirectly measured facets by the VWM task, like attention and visual acuity. Upon analysis, the results furthermore demonstrated a pronounced link between VWM precision for simple stimuli and memory accuracy for configural information in houses.

These findings highlight the varying demands placed on VWM when processing non-expert categories, which appear to be contingent on the type of information. Featural processing, which involves encoding individual features independently, may require fewer cognitive resources than configural processing. The latter necessitates the integration of multiple features and their spatial arrangements into a unified representation, potentially exerting more strain on VWM when remembering configurations for non-expert categories.

Therefore, the findings from our study contribute valuable insights to our understanding of VWM's operation, particularly in the context of different types of object categories and the varying cognitive demands associated with featural and configural processing.

6 Conclusions

This thesis examines the functional architecture of visual object recognition, focusing on the association and dissociation between certain visual recognition skills in individuals with developmental disorders comparing this with neurotypical populations. Specifically, we investigated various influential factors, including the nature of objects, the significance of visual expertise, and the demands placed on visual working memory.

In our first study on developmental dyslexia, we found that certain dyslexic readers faced difficulty in recognizing houses. These findings challenge the notion that dyslexia' is confined to reading. Intriguingly, when it came to face recognition, dyslexic readers displayed no significant difficulties compared to a control group. Such an observation bolsters the theory that face recognition operates through a distinct mechanism. Arguably, if dyslexic readers have problems with acquiring visual expertise in general, they would have demonstrated issues with face recognition as well, given that both faces and words are objects of expertise. Our subsequent study with the neurotypical population affirmed that face recognition mechanisms differ from those required for pseudowords or houses. This reinforced the notion that mere familiarity or expertise with certain objects does not influence the shared variance across expert objects like faces and words.

Additionally, our observations suggest that dyslexic readers employ a single strategy when processing objects, irrespective of focusing on their featural or configural information. However, typical readers showed distinct behaviors based on featural and configural processing. Interestingly, this distinction was not apparent in our second study that sampled the neurotypical population. This difference in results may arise from the distinct task demands. In study 2, participants appeared to modify their approach based on both the processing method and category for each block. Contrary to study 1, where participants discerned the category (either face or house) during unblocked trials without clarity on the processing type, study 2 presented blocked trials, giving participants insight into the forthcoming category and processing type. This suggests that they adopted specific strategies for each combination instead of using a more generalized featural or configural method, as was likely observed with typical readers in study 1.

In our third study, our objective was to discern whether variations in visual memory could provide clarity on the observed differences. Here, we found a robust correlation between VWM precision for basic visual stimuli and the capability to recall configural information about houses. This observation posits that being acquainted with an object might alleviate the demands on VWM. However, this correlation was absent for featural information of houses. This result indicates that configural processing of non-expert categories could be more taxing on VWM than featural processing, possibly due to the necessity of remembering both the features and their interrelationships. In contrast, when an individual has significant experience with a certain object category, they could form integrated clusters of information that enable swift recognition and processing of the object as a whole. Revisiting our initial study, it seems plausible that the difficulties that dyslexic readers encounter in recognizing houses might stem from discrepancies in their VWM capabilities. Also, the "single process" dyslexic readers employ for both featural and configural information could potentially be attributed to general VWM. This suggests that for dyslexic readers, VWM precision may play a more pivotal role in object perception compared to typical readers. Further research is warranted to delve deeper into this hypothesis.

In conclusion, we explored the intricacies of visual object recognition, investigating the association/dissociation between recognizing certain visual categories in dyslexic and typical readers. Our studies revealed that dyslexic readers face challenges recognizing houses, yet they have no such difficulties with faces. In our subsequent research with the neurotypical population, we identified distinct patterns in recognizing faces compared to pseudowords or houses. Dyslexic readers consistently approach visual objects with a unified strategy, whereas neurotypical readers adapt their method based on featural and configural cues. However, this distinction became less evident in our second study with the neurotypical group, hinting at the modifying role of task demand. Moreover, our analysis of visual working memory emphasized its significance in recognizing the configural aspects of houses. Though this linkage was not as strong for objects of expertise, our data suggest that familiarity with an object might reduce the demands on visual working memory. Interestingly, this correlation was absent when focusing on the specific features of houses, implying that recalling an object's holistic structure could be more memory-intensive, as it accounts for both individual features and their spatial relationships. We speculate that the challenges dyslexic readers face may be attributed to constraints in their visual working memory.

References

Adelman, J. S., Marquis, S. J., & Sabatos-DeVito, M. G. (2010). Letters in Words Are Read Simultaneously, Not in Left-to-Right Sequence. *Psychological Science*, *21*(12), 1799–1801. https://doi.org/10.1177/0956797610387442

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), Article 2.

Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R., & Livingstone, M. S. (2017). Seeing faces is necessary for face-domain formation. *Nature Neuroscience*, *20*(10), 1404–1412. https://doi.org/10.1038/nn.4635

Asp, I. E., Störmer, V. S., & Brady, T. F. (2021). Greater Visual Working Memory Capacity for Visually Matched Stimuli When They Are Perceived as Meaningful. *Journal of Cognitive Neuroscience*, *33*(5), Article 5. https://doi.org/10.1162/jocn_a_01693

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia*, *49*(9), Article 9.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), Article 7.

Bao, P., She, L., McGill, M., & Tsao, D. Y. (2020). A map of object space in primate inferotemporal cortex. *Nature*, *583*(7814), 103–108. https://doi.org/10.1038/s41586-020-2350-5

Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology*, *4*(2), Article 2. https://doi.org/10.1016/0010-0285(73)90012-1

Barton, J. J., Albonico, A., Susilo, T., Duchaine, B., & Corrow, S. L. (2019). Object recognition in acquired and developmental prosopagnosia. *Cognitive Neuropsychology*, *36*(1–2), Article 1–2.

Behrmann, M., & Plaut, D. C. (2013). Distributed circuits, not circumscribed centers, mediate visual recognition. *Trends in Cognitive Sciences*, *17*(5), Article 5.

Besner, D., & McCann, R. S. (2016). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In *Attention and performance XII* (pp. 201–219). Routledge.

Biotti, F., Wu, E., Yang, H., Jiahui, G., Duchaine, B., & Cook, R. (2017). Normal composite face effects in developmental prosopagnosia. *Cortex*, *95*, 63–76.

Bishop, D. V. (1997). Cognitive neuropsychology and developmental disorders: Uncomfortable bedfellows. *The Quarterly Journal of Experimental Psychology Section A*, *50*(4), Article 4.

Brachacki, G. W., Nicolson, R. I., & Fawcett, A. J. (1995). Impaired recognition of traffic signs in adults with dyslexia. *Journal of Learning Disabilities*, *28*(5), Article 5.

Brady, N., Darmody, K., Newell, F., & Cooney, S. M. (2020). *Holistic processing of words and faces in dyslexia*.

Brady, N., Darmody, K., Newell, F. N., & Cooney, S. M. (2021). Holistic processing of faces and words predicts reading accuracy and speed in dyslexic readers. *Plos One*, *16*(12), Article 12.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*(4), Article 4. https://doi.org/10.1037/a0016797

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2013). Real-world objects are not represented as bound units: Independent forgetting of different object details from visual memory. *Journal of Experimental Psychology: General*, *142*(3), Article 3. https://doi.org/10.1037/a0029649

Brady, T. F., & Störmer, V. S. (2020a). *Comparing memory capacity across stimuli requires maximally dissimilar foils: Using deep convolutional neural networks to understand visual working memory capacity for real-world objects* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/25t76

Brady, T. F., & Störmer, V. S. (2020b). *The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/kzvdg

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences*, *113*(27), Article 27. https://doi.org/10.1073/pnas.1520027113

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences*, *10*(4), Article 4.

Burns, E. J., Bennetts, R. J., Bate, S., Wright, V. C., Weidemann, C. T., & Tree, J. J. (2017). Intact word processing in developmental prosopagnosia. *Scientific Reports*, *7*(1), Article 1.

Burns, E. J., & Bukach, C. M. (2021). Face processing predicts reading ability: Evidence from prosopagnosia. *Cortex*, *145*, 67–78.

Busey, T. A., & Parada, F. J. (2010). The nature of expertise in fingerprint examiners. *Psychonomic Bulletin & Review*, *17*(2), 155–160. https://doi.org/10.3758/PBR.17.2.155

Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research*, *45*(4), 431–448.

Campbell, A., & Tanaka, J. W. (2018). Inversion Impairs Expert Budgerigar Identity Recognition: A Face-Like Effect for a Nonface Object of Expertise. *Perception*, *47*(6), 647–659. https://doi.org/10.1177/0301006618771806

Cao, X., Yang, Q., Zhong, P., & Chen, C. (2019). The characteristics of face configural effect in illiterates and literates. *Acta Psychologica*, *201*, 102951.

Centanni, T. M., Norton, E. S., Park, A., Beach, S. D., Halverson, K., Ozernov-Palchik, O., Gaab, N., & Gabrieli, J. D. (2018). Early development of letter specialization in left fusiform is associated with better word reading and smaller fusiform face area. *Developmental Science*, *21*(5), e12658. https://doi.org/10.1111/desc.12658

Chao, L. L., Weisberg, J., & Martin, A. (2002). Experience-dependent modulation of category-related cortical activity. *Cerebral Cortex*, *12*(5), Article 5.

Chen, D., Yee Eng, H., & Jiang, Y. (2006). Visual working memory for trained and novel polygons. *Visual Cognition*, *14*(1), Article 1. https://doi.org/10.1080/13506280544000282

Cohen, L. (2003). Visual Word Recognition in the Left and Right Hemispheres: Anatomical and Functional Correlates of Peripheral Alexias. *Cerebral Cortex*, *13*(12), 1313–1333. https://doi.org/10.1093/cercor/bhg079

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area. *Brain*, *123*(2), Article 2. https://doi.org/10.1093/brain/123.2.291

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *NeuroImage*, *40*(1), Article 1. https://doi.org/10.1016/j.neuroimage.2007.11.036

Cohen, M. A., Konkle, T., Rhee, J. Y., Nakayama, K., & Alvarez, G. A. (2014). Processing multiple visual objects is limited by overlap in neural channels. *Proceedings of the National Academy of Sciences*, *111*(24), Article 24. https://doi.org/10.1073/pnas.1317860111

Cohen, M. A., Nakayama, K., Konkle, T., Stantić, M., & Alvarez, G. A. (2015). Visual Awareness Is Limited by the Representational Architecture of the Visual System. *Journal of Cognitive Neuroscience*, *27*(11), Article 11. https://doi.org/10.1162/jocn_a_00855

Collins, E., Dundas, E., Gabay, Y., Plaut, D. C., & Behrmann, M. (2017). Hemispheric organization in disorders of development. *Visual Cognition*, *25*(4–6), Article 4–6.

Conway, A., Brady, N., & Misra, K. (2017). Holistic word processing in dyslexia. *Plos One*, *12*(11), Article 11.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*(2), 163.

Curby, K. M., & Gauthier, I. (2007). A visual short-term memory advantage for faces. *Psychonomic Bulletin & Review*, *14*(4), Article 4.

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), Article 1.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, *56*(2), Article 2.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), Article 6.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), Article 2.

D'Souza, D., & Karmiloff-Smith, A. (2011). When modularization fails to occur: A developmental perspective. *Cognitive Neuropsychology*, *28*(3–4), Article 3–4.

Duchaine, B. C., & Nakayama, K. (2006). Developmental prosopagnosia: A window to content-specific face processing. *Current Opinion in Neurobiology*, *16*(2), Article 2.

Dundas, E. M., Plaut, D. C., & Behrmann, M. (2013). The joint development of hemispheric lateralization for words and faces. *Journal of Experimental Psychology: General*, *142*(2), Article 2.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 628–634. https://doi.org/10.1037/0096-1523.21.3.628

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is" special" about face perception? *Psychological Review*, *105*(3), Article 3.

Feizabadi, M., Albonico, A., Starrfelt, R., & Barton, J. J. S. (2021). Whole-object effects in visual word processing: Parallels with and differences from face recognition. *Cognitive Neuropsychology*, *38*(3), 231–257. https://doi.org/10.1080/02643294.2021.1974369

Gabay, Y., Dundas, E., Plaut, D., & Behrmann, M. (2017). Atypical perceptual processing of faces in developmental dyslexia. *Brain and Language*, *173*, 41–51.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition*, *103*(2), Article 2.

Gauthier, I., McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Van Gulick, A. E. (2014). Experience moderates overlap between object and face recognition, suggesting a common ability. *Journal of Vision*, *14*(8), Article 8.

Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. *Current Opinion in Neurobiology*, *11*(2), Article 2.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*(2), Article 2.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), Article 2.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform'face area'increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*(6), Article 6.

Gerlach, C. (2009). Category-specificity in visual object recognition. *Cognition*, *111*(3), 281–301. https://doi.org/10.1016/j.cognition.2009.02.005

Gerlach, C., Klargaard, S. K., & Starrfelt, R. (2016). On the relation between face and object recognition in developmental prosopagnosia: No dissociation but a systematic association. *PloS One*, *11*(10), Article 10.

Gerlach, C., & Starrfelt, R. (2022). Face processing does not predict reading ability in developmental prosopagnosia: A commentary on Burns & Bukach (2021). *Cortex*.

Geskin, J., & Behrmann, M. (2018). Congenital prosopagnosia without object agnosia? A literature review. *Cognitive Neuropsychology*, *35*(1–2), Article 1–2.

Gray, K. L. H., & Cook, R. (2018). Should developmental prosopagnosia, developmental body agnosia, and developmental object agnosia be considered independent neurodevelopmental conditions? *Cognitive Neuropsychology*, *35*(1–2), Article 1–2. https://doi.org/10.1080/02643294.2018.1433153

Hendel, R. K., Starrfelt, R., & Gerlach, C. (2019). The good, the bad, and the average: Characterizing the relationship between face and object processing across the face recognition spectrum. *Neuropsychologia*, *124*, 274–284. https://doi.org/10.1016/j.neuropsychologia.2018.11.016

Hernandez, A. E., Claussenius-Kalman, H. L., Ronderos, J., Castilla-Earls, A. P., Sun, L., Weiss, S. D., & Young, D. R. (2019). Neuroemergentism: A framework for studying cognition and the brain. *Journal of Neurolinguistics*, *49*, 214–223. https://doi.org/10.1016/j.jneuroling.2017.12.010

Hernández, M., Martin, C. D., Barceló, F., & Costa, A. (2013). Where is the bilingual advantage in task-switching? *Journal of Memory and Language*, *69*(3), 257–276. https://doi.org/10.1016/j.jml.2013.06.004

Hills, C. S., Pancaroglu, R., Duchaine, B., & Barton, J. J. (2015). Word and text processing in acquired prosopagnosia. *Annals of Neurology*, *78*(2), Article 2.

Holmes, D. R., & McKeever, W. F. (1979). Material specific serial memory deficit in adolescent dyslexics. *Cortex*, *15*(1), Article 1.

Hsiao, J. H., & Cottrell, G. W. (2009). Not all visual expertise is holistic, but it may be leftist: The case of Chinese character recognition. *Psychological Science*, *20*(4), Article 4.

Huang, L. (2011). Familiarity does not aid access to features. *Psychonomic Bulletin & Review*, *18*(2), Article 2. https://doi.org/10.3758/s13423-011-0052-6

Huang, L., & Awh, E. (2018). Chunking in working memory via content-free labels. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-017-18157-5

Huestegge, L., Rohrßen, J., van Ermingen-Marbach, M., Pape-Neumann, J., & Heim, S. (2014). Devil in the details? Developmental dyslexia and visual long-term memory for details. *Frontiers in Psychology*, *5*, 686.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*(6), Article 6. https://doi.org/10.1016/0749-596X(91)90032-F

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), Article 5. https://doi.org/10.1037/0278-7393.23.5.1217

Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), Article 3.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Jones, T., & Farrell, S. (2018). Does syntax bias serial order reconstruction of verbal short-term memory? *Journal of Memory and Language*, *100*, 98–122. https://doi.org/10.1016/j.jml.2018.02.001

Jozranjbar, B., Kristjansson, A., & Sigurdardottir, H. M. (2021). Featural and configural processing of faces and houses in matched dyslexic and typical readers. *Neuropsychologia*, 108059.

Jozranjbar, B., Kristjánsson, Á., Starrfelt, R., Gerlach, C., & Sigurdardottir, H. M. (2023). Using representational similarity analysis to reveal category and process specificity in visual object recognition. *Cortex*, S0010945223001387. https://doi.org/10.1016/j.cortex.2023.05.012

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), Article 2. https://doi.org/10.1016/0010-0285(92)90007-O

Kaiser, D., Stein, T., & Peelen, M. V. (2015). Real-world spatial regularities affect visual working memory for objects. *Psychonomic Bulletin & Review*, *22*(6), Article 6. https://doi.org/10.3758/s13423-015-0833-4

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*(8), Article 8.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), Article 11.

Kleinschmidt, A., & Cohen, L. (2006). The neural bases of prosopagnosia and pure alexia: Recent insights from functional neuroimaging. *Current Opinion in Neurology*, *19*(4), Article 4. https://doi.org/10.1097/01.wco.0000236619.89710.ee

Konkle, T., & Oliva, A. (2012a). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, *74*(6), 1114–1124.

Konkle, T., & Oliva, A. (2012b). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, *74*(6), Article 6.

Koriat, A., & Norman, J. (1985). Reading rotated words. *Journal of Experimental Psychology. Human Perception and Performance*, *11*(4), 490–508. https://doi.org/10.1037//0096-1523.11.4.490

Kristjánsson, A. (2006). Simultaneous priming along multiple feature dimensions in a visual search task. *Vision Research*, *46*(16), 2554–2570.

Kristjánsson, Á., & Sigurdardottir, H. M. (2023). The Role of Visual Factors in Dyslexia. *Journal of Cognition*, *6*(1), 31. https://doi.org/10.5334/joc.287

Kühn, C. D., Gerlach, C., Andersen, K. B., Poulsen, M., & Starrfelt, R. (2020). Face recognition in developmental dyslexia: Evidence for dissociation between faces and words. *Cognitive Neuropsychology*, 1–9.

Le Grand, R., Cooper, P. A., Mondloch, C. J., Lewis, T. L., Sagiv, N., de Gelder, B., & Maurer, D. (2006). What aspects of face processing are impaired in developmental prosopagnosia? *Brain and Cognition*, *61*(2), Article 2.

Li, X., Xiong, Z., Theeuwes, J., & Wang, B. (2020). Visual memory benefits from prolonged encoding time regardless of stimulus type. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(10), Article 10. https://doi.org/10.1037/xlm0000847

Lieder, I., Adam, V., Frenkel, O., Jaffe-Dax, S., Sahani, M., & Ahissar, M. (2019). Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia. *Nature Neuroscience*, *22*(2), Article 2.

Liu, T. T., & Behrmann, M. (2014). Impaired holistic processing of left-right composite faces in congenital prosopagnosia. *Frontiers in Human Neuroscience*, *8*, 750.

Lobmaier, J. S., Bölte, J., Mast, F. W., & Dobel, C. (2010). Configural and featural processing in humans with congenital prosopagnosia. *Advances in Cognitive Psychology*, *6*, 23.

Long, B., Konkle, T., Cohen, M. A., & Alvarez, G. A. (2016). Mid-level perceptual features distinguish objects of different real-world sizes. *Journal of Experimental Psychology: General*, *145*(1), Article 1. https://doi.org/10.1037/xge0000130

Long, B., Moher, M., Carey, S., & Konkle, T. (2019). Real-world size is automatically encoded in preschoolers' object representations. *Journal of Experimental*

Psychology: Human Perception and Performance, *45*(7), Article 7. https://doi.org/10.1037/xhp0000619

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), Article 6657.

Maratos, F. A., Chu, K., Lipka, S., Stupple, E. J. N., & Parente, F. (2022). Exploring pattern recognition: What is the relationship between the recognition of words, faces and other objects? *Cognitive Processing*. https://doi.org/10.1007/s10339-022-01111-3

Markov, Y. A., Utochkin, I. S., & Brady, T. F. (2021). Real-world objects are not stored in holistic representations in visual working memory. *Journal of Vision*, *21*(3), Article 3. https://doi.org/10.1167/jov.21.3.18

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002a). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002b). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), Article 6.

Mayseless, N., & Breznitz, Z. (2011). Brain activity during processing objects and pseudo-objects: Comparison between adult regular and dyslexic readers. *Clinical Neurophysiology*, *122*(2), Article 2.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375–407. https://doi.org/10.1037/0033-295X.88.5.375

McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Gauthier, I. (2012). The Vanderbilt Expertise Test reveals domain-general and domain-specific sex effects in object recognition. *Vision Research*, *69*, 10–22.

McKone, E., Crookes, K., Jeffery, L., & Dilks, D. D. (2012). A critical review of the development of face recognition: Experience is less important than previously believed. *Cognitive Neuropsychology*, *29*(1–2), 174–212. https://doi.org/10.1080/02643294.2012.660138

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences*, *11*(1), Article 1.

McKone, E., & Robbins, R. (2007). The evidence rejects the expertise hypothesis: Reply to Gauthier & Bukach. *Cognition*, *103*(2), Article 2.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, *98*(2), 164–181. https://doi.org/10.1037/0033-295X.98.2.164

Moses, P., & Stiles, J. (2002). The lesion methodology: Contrasting views from adult and child studies. *Developmental Psychobiology*, *40*(3), Article 3. https://doi.org/10.1002/dev.10031

Nasr, S., Echavarria, C. E., & Tootell, R. B. (2014). Thinking outside the box: Rectilinear shapes selectively activate scene-selective cortex. *Journal of Neuroscience*, *34*(20), Article 20.

Ngiam, W. X. Q., Khaw, K. L. C., Holcombe, A. O., & Goodbourn, P. T. (2019). Visual working memory for letters varies with familiarity but not complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(10), Article 10. https://doi.org/10.1037/xlm0000682

Norris, D. G., & Kalm, K. (2018). *What's in a chunk? Chunking and data compression in verbal short-term memory* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/2st4j

O'Donnell, R. E., Clement, A., & Brockmole, J. R. (2018). Semantic and functional relationships among objects increase the capacity of visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(7), Article 7. https://doi.org/10.1037/xlm0000508

Olson, I. R., & Jiang, Y. (2004). Visual short-term memory is not improved by training. *Memory & Cognition*, *32*(8), Article 8. https://doi.org/10.3758/BF03206323

Owen, S., & Maratos, F. A. (2016). Recognition of subtle and universal facial expressions in a community-based sample of adults classified with intellectual disability: Facial expression recognition in ID adults. *Journal of Intellectual Disability Research*, *60*(4), 344–354. https://doi.org/10.1111/jir.12253

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia*, *49*(5), Article 5.

Parr, L. A. (2011). The evolution of face processing in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1764–1777. https://doi.org/10.1098/rstb.2010.0358

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), Article 4. https://doi.org/10.3758/BF03210419

Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature*, *423*(6941), Article 6941.

Pelli, D. G., & Tillman, K. A. (2007). Parts, Wholes, and Context in Reading: A Triple Dissociation. *PLoS ONE*, *2*(8), e680. https://doi.org/10.1371/journal.pone.0000680

Phillips, P. J., Yates, A. N., Hu, Y., Hahn, C. A., Noyes, E., Jackson, K., Cavazos, J. G., Jeckeln, G., Ranjan, R., Sankaranarayanan, S., Chen, J.-C., Castillo, C. D., Chellappa, R., White, D., & O'Toole, A. J. (2018). Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *Proceedings of the National Academy of Sciences*, *115*(24), 6171–6176. https://doi.org/10.1073/pnas.1721355115

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage*, *19*(3), Article 3.

Price, C. J., & Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*(6), 246–253. https://doi.org/10.1016/j.tics.2011.04.001

Quirk, C., Adam, K. C. S., & Vogel, E. K. (2020). No Evidence for an Object Working Memory Capacity Benefit with Extended Viewing Time. *Eneuro*, *7*(5), Article 5. https://doi.org/10.1523/ENEURO.0150-20.2020

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*(2), 275–280. https://doi.org/10.1037/h0027768

Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of Cognitive Neuroscience*, *16*(2), Article 2.

Rice, G. E., Kerry, S. J., Robotham, R. J., Leff, A., Ralph, M. A. L., & Starrfelt, R. (2020). *Revealing the graded spectrum of visual perceptual function following posterior cerebral artery stroke*.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic Processing Predicts Face Recognition. *Psychological Science*, *22*(4), Article 4. https://doi.org/10.1177/0956797611401753

Richler, J. J., Tomarken, A. J., Sunday, M. A., Vickery, T. J., Ryan, K. F., Floyd, R. J., Sheinberg, D., Wong, A. C.-N., & Gauthier, I. (2019a). Individual differences in object recognition. *Psychological Review*, *126*(2), Article 2. https://doi.org/10.1037/rev0000129

Richler, J. J., Tomarken, A. J., Sunday, M. A., Vickery, T. J., Ryan, K. F., Floyd, R. J., Sheinberg, D., Wong, A. C.-N., & Gauthier, I. (2019b). Individual differences in object recognition. *Psychological Review*, *126*(2), Article 2. https://doi.org/10.1037/rev0000129

Ripamonti, E., Luzzatti, C., Zoccolotti, P., & Traficante, D. (2018). Word and pseudoword superiority effects: Evidence from a shallow orthography language. *Quarterly Journal of Experimental Psychology*, *71*(9), 1911–1920. https://doi.org/10.1080/17470218.2017.1363791

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, *103*(1), Article 1. https://doi.org/10.1016/j.cognition.2006.02.008

Roberts, D. J., Ralph, M. A. L., Kim, E., Tainturier, M.-J., Beeson, P. M., Rapcsak, S. Z., & Woollams, A. M. (2015). Processing deficits for familiar and novel faces in patients with left posterior fusiform lesions. *Cortex*, *72*, 79–96.

Robotham, R. J., & Starrfelt, R. (2017). Face and word recognition can be selectively affected by brain injury or developmental disorders. *Frontiers in Psychology*, *8*, 1547.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, *21*(2), Article 2.

Rossion, B., Dricot, L., Devolder, A., Bodart, J.-M., Crommelinck, M., Gelder, B. de, & Zoontjes, R. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *12*(5), Article 5.

Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Defining face perception areas in the human brain: A large-scale factorial fMRI face localizer analysis. *Brain and Cognition*, *79*(2), 138–157. https://doi.org/10.1016/j.bandc.2012.01.001

Rubino, C., Corrow, S. L., Corrow, J. C., Duchaine, B., & Barton, J. J. (2016). Word and text processing in developmental prosopagnosia. *Cognitive Neuropsychology*, *33*(5–6), Article 5–6.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, *89*(1), 60–94. https://doi.org/10.1037/0033-295X.89.1.60

Rüsseler, J., Johannes, S., & Münte, T. F. (2003). Recognition memory for unfamiliar faces does not differ for adult normal and dyslexic readers: An event-related brain potential study. *Clinical Neurophysiology*, *114*(7), Article 7.

Shaywitz, S. E. (1998). Dyslexia. *New England Journal of Medicine*, *338*(5), Article 5.

Sigurdardottir, H. M., Arnardottir, A., & Halldorsdottir, E. T. (2021). Faces and words are both associated and dissociated as evidenced by visual problems in dyslexia. *Scientific Reports*, *11*(1), Article 1.

Sigurdardottir, H. M., Arnardottir, A., Halldorsdottir, E. T., Omarsdottir, H. R., & Valgeirsdottir, A. S. (2019). *Faces and words are both associated and dissociated: Evidence from visual problems in dyslexia*.

Sigurdardottir, H. M., Danielsdottir, H. B., Gudmundsdottir, M., Hjartarson, K. H., Thorarinsdottir, E. A., & Kristjánsson, Á. (2017). Problems with visual statistical learning in developmental dyslexia. *Scientific Reports*, *7*(1), Article 1.

Sigurdardottir, H. M., Fridriksdottir, L. E., Gudjonsdottir, S., & Kristjánsson, Á. (2018). Specific problems in visual cognition of dyslexic readers: Face discrimination deficits predict dyslexia over and above discrimination of scrambled faces and novel objects. *Cognition*, *175*, 157–168.

Sigurdardottir, H. M., Ívarsson, E., Kristinsdóttir, K., & Kristjánsson, Á. (2015). Impaired recognition of faces and objects in dyslexia: Evidence for ventral stream dysfunction? *Neuropsychology*, *29*(5), Article 5.

Simion, F., & Giorgio, E. D. (2015). Face perception and processing in early infancy: Inborn predispositions and developmental changes. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00969

Simion, F., Leo, I., Turati, C., Valenza, E., & Dalla Barba, B. (2007). How face specialization emerges in the first months of life. In *Progress in Brain Research* (Vol. 164, pp. 169–185). Elsevier. https://doi.org/10.1016/S0079-6123(07)64009-6

Smith-Spark, J. H., & Moore, V. (2009). The representation and processing of familiar faces in dyslexia: Differences in age of acquisition effects. *Dyslexia*, *15*(2), Article 2.

Srihasam, K., Vincent, J. L., & Livingstone, M. S. (2014). Novel domain formation reveals proto-architecture in inferotemporal cortex. *Nature Neuroscience*, *17*(12), 1776–1783. https://doi.org/10.1038/nn.3855

Starrfelt, R., & Gerlach, C. (2007). The visual what for area: Words and pictures in the left fusiform gyrus. *Neuroimage*, *35*(1), Article 1.

Starrfelt, R., Klargaard, S. K., Petersen, A., & Gerlach, C. (2018). Reading in developmental prosopagnosia: Evidence for a dissociation between word and face recognition. *Neuropsychology*, *32*(2), Article 2.

Starrfelt, R., Petersen, A., & Vangkilde, S. (2013). Don't words come easy? A psychophysical exploration of word superiority. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00519

Sunday, M. A., & Gauthier, I. (2018). Face expertise for unfamiliar faces: A commentary on Young and Burton's "Are we face experts?" *Journal of Expertise*, *1*(1), 35–41.

Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., Pidcock, M., Dawel, A., Jeffery, L., & Wilson, C. E. (2010). Face recognition impairments despite normal holistic processing and face space coding: Evidence from a case of developmental prosopagnosia. *Cognitive Neuropsychology*, *27*(8), Article 8.

Susilo, T., Rezlescu, C., & Duchaine, B. (2013). The composite effect for inverted faces is reliable at large sample sizes and requires the basic face configuration. *Journal of Vision*, *13*(13), 14–14. https://doi.org/10.1167/13.13.14

Tanaka, J. W., & Farah, M. J. (1993). Parts and Wholes in Face Recognition. *The Quarterly Journal of Experimental Psychology Section A*, *46*(2), Article 2. https://doi.org/10.1080/14640749308401045

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, *25*(5), Article 5.

Tangen, J. M., Thompson, M. B., & McCarthy, D. J. (2011). Identifying fingerprint expertise. *Psychological Science*, *22*(8), 995–997.

Towler, A., Dunn, J. D., Martínez, S. C., Moreton, R., Eklöf, F., Ruifrok, A., Kemp, R., & White, D. (2021). *Diverse routes to expertise in facial recognition*.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A Cortical Region Consisting Entirely of Face-Selective Cells. *Science*, *311*(5761), 670–674. https://doi.org/10.1126/science.1119983

Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of Face Perception. *Annual Review of Neuroscience*, *31*(1), 411–437. https://doi.org/10.1146/annurev.neuro.30.051606.094238

Tso, R. V., Au, T. K., & Hsiao, J. H. (2014). Perceptual expertise: Can sensorimotor experience change holistic processing and left-side bias? *Psychological Science*, *25*(9), Article 9.

Tso, R. V., Chan, R. T., Chan, Y., & Lin, D. (2021). Holistic processing of Chinese characters in college students with dyslexia. *Scientific Reports*, *11*(1), Article 1.

Tso, R. V. Y., Chan, R. T. C., & Hsiao, J. H. (2020). Holistic but with reduced right-hemisphere involvement: The case of dyslexia in Chinese character recognition. *Psychonomic Bulletin & Review*, 1–10.

Ulrich, P. I., Wilkinson, D. T., Ferguson, H. J., Smith, L. J., Bindemann, M., Johnston, R. A., & Schmalzl, L. (2017). Perceptual and memorial contributions to developmental prosopagnosia. *Quarterly Journal of Experimental Psychology*, *70*(2), Article 2.

Van Yip Tso, R., Au, T. K., & Hsiao, J. H. (2012). *Writing facilitates learning to read in Chinese through reduction of holistic processing: A developmental study*. *34*(34).

Ventura, P., Fernandes, T., Leite, I., Almeida, V. B., Casqueiro, I., & Wong, A. C.-N. (2017). The Word Composite Effect Depends on Abstract Lexical Representations But Not Surface Features Like Case and Font. *Frontiers in Psychology*, *8*, 1036. https://doi.org/10.3389/fpsyg.2017.01036

Ventura, P., Fernandes, T., Pereira, A., Guerreiro, J. C., Farinha-Fernandes, A., Delgado, J., Ferreira, M. F., Faustino, B., Raposo, I., & Wong, A. C.-N. (2020). Holistic word processing is correlated with efficiency in visual word recognition. *Attention, Perception, & Psychophysics*, *82*(5), 2739–2750. https://doi.org/10.3758/s13414-020-01988-2

Vogel, A. C., Petersen, S. E., & Schlaggar, B. L. (2014). The VWFA: it's not just for words anymore. *Frontiers in Human Neuroscience*, *8*, 88.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), Article 1. https://doi.org/10.1037/0096-1523.27.1.92

Wong, A. C.-N., Bukach, C. M., Hsiao, J., Greenspon, E., Ahern, E., Duan, Y., & Lui, K. F. H. (2012). Holistic processing as a hallmark of perceptual expertise for nonface categories including Chinese characters. *Journal of Vision*, *12*(13), 7–7. https://doi.org/10.1167/12.13.7

Wong, A. C.-N., Bukach, C. M., Yuen, C., Yang, L., Leung, S., & Greenspon, E. (2011). Holistic processing of words modulated by reading experience. *PloS One*, *6*(6), Article 6.

Wong, A. C.-N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects: Becoming a Ziggerin expert—But which type? *Psychological Science*, *20*(9), Article 9.

Wong, A. C.-N., Wong, Y. K., Lui, K. F., Ng, T. Y., & Ngan, V. S. (2019). Sensitivity to configural information and expertise in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(1), Article 1.

Xie, W., & Zhang, W. (2017a). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory & Cognition*, *45*(4), Article 4. https://doi.org/10.3758/s13421-016-0679-7

Xie, W., & Zhang, W. (2017b). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory & Cognition*, *45*(4), Article 4. https://doi.org/10.3758/s13421-016-0679-7

Xie, W., & Zhang, W. (2018). Familiarity Speeds Up Visual Short-term Memory Consolidation: Electrophysiological Evidence from Contralateral Delay Activities. *Journal of Cognitive Neuroscience*, *30*(1), 1–13. https://doi.org/10.1162/jocn_a_01188

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*(1), Article 1.

Young, A. W. (2018). Faces, people and the brain: The 45th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology*, *71*(3), 569–594. https://doi.org/10.1177/1747021817740275

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational Information in Face Perception. *Perception*, *16*(6), Article 6. https://doi.org/10.1068/p160747

Young, A. W., Hellawell, D., & Hay, D. C. (2013). Configurational information in face perception. *Perception, 42(11), 1166-1178.*

Yovel, G., & Duchaine, B. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience*, *18*(4), Article 4.

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, *44*(5), Article 5.

Yue, X., Pourladian, I. S., Tootell, R. B., & Ungerleider, L. G. (2014). Curvature-processing network in macaque visual cortex. *Proceedings of the National Academy of Sciences*, *111*(33), Article 33.

Zhang, L., & Cottrell, G. W. (2005). *Holistic processing develops because it is good*. 2428–2433.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), Article 7192.

Zhou, G., Cheng, Z., Zhang, X., & Wong, A. C.-N. (2012). Smaller holistic processing of faces associated with face drawing experience. *Psychonomic Bulletin & Review*, *19*(2), Article 2.

Original Publications

This Thesis has been written based on the following papers:

Paper I: Jozranjbar, B., Kristjánsson, Á., & Sigurdardottir, H. M. (2021). Featural and configural processing of faces and houses in matched dyslexic and typical readers. Neuropsychologia, 162, 108059.

Paper II: Jozranjbar, B., Kristjánsson, Á., Starrfelt, R., Gerlach, C. & Sigurdardottir, H.M. (2023). Using representational similarity analysis to reveal category and process specificity in visual object recognition. Cortex, 166, 172-187.

Paper III: Jozranjbar, B., Kristjánsson, Á., Gerlach, C., & Sigurdardottir, H. M. (submitted for publication) The impact of visual working memory constraints on object recognition.

Paper I

Featural and Configural Processing of Faces and Houses in Matched Dyslexic and Typical Readers

Bahareh Jozranjbar[1*], Árni Kristjánsson[1,2] & Heida Maria Sigurdardottir[1]

[1]Icelandic Vision Laboratory
Department of Psychology
University of Iceland

[2] School of Psychology
National Research University Higher School
of Economics, Moscow, Russian Federation

Abstract

While dyslexia is typically described as a phonological deficit, recent evidence suggests that ventral stream regions, important for visual categorization and object recognition, are hypoactive in dyslexic readers who might accordingly show visual recognition deficits. By manipulating featural and configural information of faces and houses, we investigated whether dyslexic readers are disadvantaged at recognizing certain object classes or using particular visual processing mechanisms. Dyslexic readers found it harder to recognize objects (houses), suggesting that visual problems in dyslexia are not completely domain-specific. Face recognition accuracy was equivalent in the two groups. Lower recognition accuracy for houses was also related to reading difficulties even when accuracy for faces was kept constant, which could indicate a specific relationship between visual word processing and visual processing of non-face objects. Representational similarity analyses (RSA) revealed that featural and configural processes were clearly separable in typical readers, which was not the case for dyslexic readers who appear to rely on a single process. This was not restricted to particular visual categories, occurring for both faces and houses. We speculate that reading deficits in some dyslexic readers reflect their reliance on a single process for object recognition.

Keywords: dyslexia; reading; face recognition; object recognition; high-level vision

Highlights

- Dyslexia is not exclusively restricted to difficulties in visual word processing, as some dyslexic readers show a deficiency in object recognition.
- Representational Similarity Analysis (RSA) suggests that dyslexic readers rely on a single visual process (featural or configural) regardless of task demands.
- Dyslexic readers' failure to use different processes could contribute to their reading difficulty, as efficient reading requires both featural and configural processing.

Introduction

Developmental dyslexia is a reading disorder that occurs despite normal intellectual capacity, adequate educational opportunities and intact sensory abilities (Shaywitz, 1998). A century ago, developmental dyslexia was considered a visual memory deficit (Hinshelwood, 1896; Morgan, 1896). In the 1970s, the focus of dyslexia research moved from visual deficits to impaired phonological processing (Vellutino et al., 2004). While phonological impairments should not be underestimated, the role of phonological factors in reading depends on the orthographic depth of languages and is less important in transparent orthographies (Norton et al., 2015; Ziegler et al., 2010). More recently, evidence has accumulated for temporal, motor, attentional, auditory, and visual dysfunction in dyslexia (De Martino et al., 2001; Farmer & Klein, 1995; Giofrè et al., 2019; Goswami, 2011; Norton et al., 2015; Reid, 2018; Valdois et al., 2004; Ziegler et al., 2010) suggesting that dyslexia is a heterogeneous deficit.

Researchers continue to debate whether dyslexic readers have difficulties with high-level visual processing and, if so, whether their problem is limited to specific visual categories or processes. Recent behavioral and neuroscientific research suggests that dyslexic readers have problems with high-level visual processing (Collins et al., 2017; Gabay et al., 2017; Sigurdardottir, Arnardottir, et al., 2019; Sigurdardottir et al., 2018; Sigurdardottir et al., 2015). High-level vision is typically considered to involve later stages of the ventral visual stream that analyze the physical properties of objects and surfaces in the environment. Low-level vision refers to earlier visual stages and is characterized by processing physical properties of the retinal image (Cox, 2014).

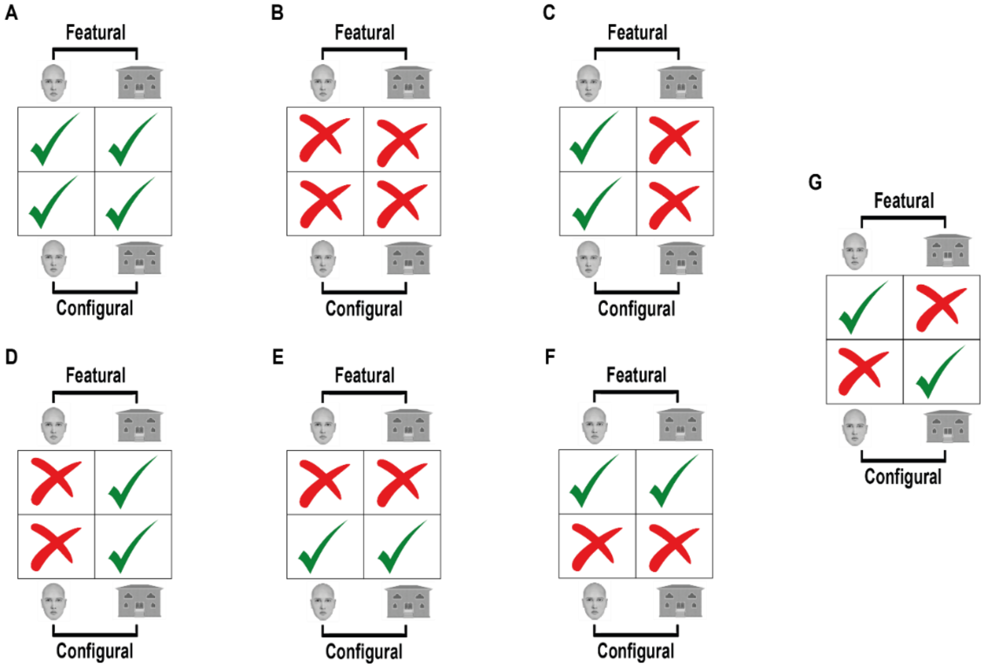


Figure 1. A visual summary of possibilities regarding featural or configural processing of faces and houses in dyslexic readers. A. Domain-specific; B. Domain-general; C. Deficit for non-face objects; D. Deficit for objects of expertise, or for faces and words as a result of mutual structural and functional dependency of these two categories; E. Process-specificity in featural processing difficulty; F. Process-specificity in configural processing difficulty; G. Process-specificity in the process most important for recognizing objects of a particular visual category. Checks indicate intact abilities, whereas crosses indicate deficient abilities.

A meta-analysis of functional imaging studies of reading and reading-related tasks in children and adults with dyslexia uncovered hypoactivity of high-level ventral stream regions – in or close to the left fusiform gyrus (Richlan et al., 2011). Notably, these hypoactive regions seem to partly correspond to the so-called visual word form area (VWFA, see e.g. Cohen et al., 2002; Dehaene & Cohen, 2007, 2011), potentially reflecting a visual recognition deficit for words. These areas – including the VWFA – are however not only activated during visual word tasks but also during different types of visual object tasks (Price & Devlin, 2003; Starrfelt & Gerlach, 2007; Vogel et al., 2014). Therefore, potential visual recognition deficits in dyslexic readers might be restricted to visual words (Domain-specific; figure 1A) or generalize across categories (Domain-general; figure 1B). Domain-specific accounts assume that specialized cognitive functions support the processing of specific types of visual stimuli (Kanwisher, 2000; Rhodes et al., 2004; Yovel & Kanwisher, 2004) such as words, faces or places. Domain-general accounts postulate that general mechanisms operate on several visual categories but to different degrees (Behrmann & Plaut, 2013; Gauthier et al., 2014; Hills et al., 2015; Rice et al., 2020).

Face processing has been at the center of the domain generality / specificity debate. If face processing requires domain-specific mechanisms (e.g. Kanwisher, 2000; Rhodes et al., 2004; Yovel & Kanwisher, 2004) then visual word processing problems in dyslexia should be completely independent of any problems in face processing (see Robotham & Starrfelt, 2017); word processing deficits could however generalize to other non-face objects if visual words do not depend on domain-specific mechanisms (figure 1C). According to other accounts, word processing problems in dyslexia should specifically generalize to faces, as these accounts postulate a mutual structural and functional dependency of faces and words through specific restrictions on neural and cognitive development (figure 1D; Behrmann & Plaut, 2019; Dehaene et al., 2010; Plaut & Behrmann, 2011). Additionally, both words and faces are objects of expertise (e.g. Gauthier & Nelson, 2001; Gauthier et al., 2000; Ventura et al., 2019; Wong & Gauthier, 2007) and problems with gaining visual expertise in dyslexia could lead to specific connections between word and face processing (Lieder et al., 2019; Sigurdardottir et al., 2017; but see Sigurdardottir, Hjartarson, et al., 2019). This would predict a specific deficiency for word and face processing but not for other visual categories in dyslexia (figure 1D).

Several studies have found no significant differences in the performance of dyslexic and typical readers in face recognition (Brachacki et al., 1994; Holmes & McKeever, 1979; Rüsseler et al., 2003). Although Smith-Spark and Moore (2009) found no overall differences in naming speed or accuracy for familiar faces between dyslexic and typical readers, they discovered that typical readers were faster at naming familiar faces that they learnt earlier in life than dyslexic readers. Several other studies have on the other hand reported face processing problems for dyslexic readers (Collins et al., 2017; Gabay et al., 2017; Sigurdardottir, Arnardottir, et al., 2019; Sigurdardottir et al., 2018; Sigurdardottir et al., 2015). Additionally, more general visual recognition problems for dyslexic readers have been found. Sigurdardottir et al. (2015) reported difficulties in recognizing exemplars of other complex non-word familiar visual categories such as different birds, butterflies, cars, houses, or planes. Brachacki et al. (1995) found that adults with dyslexia were worse than typical readers at distinguishing between real and fake traffic signs. Huestegge et al. (2014) also reported more detail-related errors in visual long-term memory of children with dyslexia than their matched controls. However, Gabay et al. (2017) and Sigurdardottir et al. (2018) did not observe object recognition deficits for dyslexic readers for cars and novel objects. The results from the current literature is therefore mixed.

Differences in processing?

One potential reason for the mixed findings is that visual problems in dyslexia are process-specific, and some tasks may specifically tap into a visual process that is deficient in dyslexic readers. Featural and configural processing have been identified as separable approaches to visual recognition (Lobmaier et al., 2010; Rossion et al., 2000), and it is possible that dyslexic readers have a deficiency in one but not the other. The term 'featural processing' refers to the processing of the basic features of a stimulus (e.g., letters in a word) while 'configural processing' involves encoding or interpreting associations between the features of a stimulus (Maurer et al., 2002). Three types of configural processing have been described: (1) first order relations – referring to the basic configuration of features (e.g., the relative position of letters in a word); (2) second-order relations – perceiving the spatial relationship among features (e.g., distance between letters); and (3) holistic processing – perceiving the features as a Gestalt (e.g., compulsory attention to all parts of a word) (Maurer et al., 2002; Roberts et al., 2015).

Featural processing is generally assumed to be important for word recognition (Farah et al., 1998; Martelli et al., 2005; Pelli et al., 2003) while configural processing is often considered the hallmark of face processing (McKone et al., 2007) and less important for word recognition (Farah et al., 1998). Sigurdardottir et al. (2015) demonstrated intact holistic processing of faces in dyslexic readers. As face and object recognition problems for dyslexic readers were still found, this may reflect problems with featural processing in dyslexia. If dyslexia is process-specific, difficulties with featural processing of non-word objects (regardless of object category) would be expected as word processing is assumed to depend predominantly on featural information (figure 1E). Interestingly, Sigurdardottir, Arnardottir, et al. (2019) found that dyslexic readers had problems with featural processing of faces while global form processing of faces – likely reflecting configural processing – was intact. Tso et al. (2020) found that dyslexic readers had stronger holistic processing and weaker left side biases for Chinese characters than controls. As expertise in Chinese characters is associated with decreased holistic processing and a stronger left side bias (an indication of right hemisphere lateralization; Hsiao & Cottrell, 2009; Tso et al., 2014), Tso et al. (2014) proposed that the right hemisphere can engage in either holistic or part-based/featural processing based on task demands, and that dyslexic readers might have a problem with right hemisphere featural processing.

Gaining visual expertise with a category is however more often associated with increased configural processing for the category (e.g. Bukach et al., 2006; Gauthier & Bukach, 2007; Gauthier & Tarr, 2002; but see McKone & Robbins, 2007). Wong et al. (2011; 2019) found that word recognition in expert readers is comparable to other domains of perceptual expertise as it relies on configural processing. Configural processing of words is involved in fast parallel reading (Ventura et al., 2019) and as reading skills improve, letter-by-letter reading changes to parallel word reading (Grainger et al., 2012; Grainger & Ziegler, 2011). With improved reading abilities, the VWFA also becomes more sensitive to common combinations of letters, presumably reflecting reading expertise (Binder et al., 2006). Wong et al. (2011) showed that word processing is more likely to be configural for native than second-language readers, and for words compared to pseudowords. Visual recognition deficits in dyslexia may therefore also manifest as impairments in fast and accurate configural processing of words and objects (figure 1F).

The literature on featural vs. configural processing is dyslexia is somewhat mixed. For example, Brady et al. (2020) suggested that dyslexia is associated with increased configural (holistic) processing of words, while Conway et al. (2017) suggested that holistic word processing in dyslexia is impaired and that dyslexic readers may read more analytically than typical readers. Franceschini et al. (2017) also found that while typical readers process global or configural information before local or featural information, dyslexic children prioritize local processing above global processing. Franceschini et al. (2017) also reported that a lack of normal hierarchical global-to-local visual processing in pre-readers is predictive of future reading problems (see also Franceschini et al., 2021).

Potential configural processing deficits in dyslexic readers can possibly be attributed to their difficulty with acquiring expertise, as studies have shown that dyslexic readers have a problem with visual learning (e.g. Lieder et al., 2019; Sigurdardottir et al., 2017; but see Brachacki et al., 1995; Sigurdardottir, Hjartarson, et al., 2019). However, configural processing may not be a marker of general visual expertise and how expertise impacts different processes could depend on how much information each process conveys (Hsiao & Cottrell, 2009; Zhang & Cottrell, 2005) and on task demands (Zhou et al., 2012). Expertise may only lead to enhanced processing of configurations if this is the most diagnostic information for identity. When features are crucial for recognition, then expertise may improve the efficiency of featural processing for those categories. Zhou et al. (2012), for instance, found that art students who are experts in face drawing have decreased holistic face processing, since when art students draw faces, they need to attend to face features. Also, Tso et al. (2014), discovered that Chinese readers with limited writing experience had higher holistic processing, but Chinese readers who could write characters fluently had reduced holistic processing, presumably because writing improves the ability to perceive the characters analytically. Therefore, if the process shift is experience-dependent, we may expect dyslexic readers to have problems with processes that are important for that specific category if they have trouble with acquiring visual expertise (figure 1, example G). For example, configural processing is dominant for face recognition, therefore dyslexic readers who cannot gain expertise may have limited configural processing abilities for faces.

Current Aims

Given the ambiguity of existing empirical findings, we manipulated configural vs. featural information of two visual categories, faces and houses, to address the following questions: (a) Do dyslexic readers have problems with recognizing non-word visual objects? (b) Are such potential problems specific to particular visual categories (faces or houses)? (c) Are problems in dyslexic readers process-specific (featural or configural)?

Method

Participants

We recruited 68 participants, or 34 pairs matched on gender, age (±5 years), and educational levels. Thirty-four participants reported a previous diagnosis of dyslexia (21 women; mean age: 37.2 years, range 18-62) and 34 were self-reported typical readers (21 women; mean age: 36.8, range 18—67). The stopping rule for data collection was to either test all volunteers for the study up to a limit of 40 matched pairs, or to stop data collection when no new matched pairs could be recruited within a particular extended time period, whichever came first.

In each group, 6 people had completed the first level of schooling (high school), 15 the second level (gymnasium), 8 the third level (undergraduate degree), and 5 had completed the fourth level (graduate degree). All participants reported normal or corrected-to-normal vision. Participants were recruited using varied means, e.g. advertisements on social media and radio. All were native Icelandic speakers. Participants received a gift certificate at a local shopping mall (value: 3000 ISK, approximately $25) for participation.

For our main analyses, three participant pairs were excluded, leaving 31 matched pairs (see section Method: Verification of group classification). With an alpha = .05, power = 0.80, and 31 participants in each group, minimum detectable effect size (MDE) of paired Cohen's d was 0.52, and MDE of Pearson's r was 0.35 (GPower 3.1; Faul et al., 2007), considered to be medium effect sizes.

Procedure

The study was approved by the Icelandic National Bioethics Committee (ID 14-027) and reported to the Icelandic Data Protection Authority. Participants gave their informed consent. The experiment took place in a well-lit and quiet room. All tasks and questionnaires were computerized and presented using PsychoPy (Peirce et al., 2019) on an Asus monitor (60 Hz, resolution 2560 x 1440 pixels). Viewing distance was approximately 57 cm. The stimuli were presented on a white background. Verbal instructions were prerecorded, and participants were instructed to listen to them attentively.

Participants sat in front of the monitor and started by answering questions about previous diagnoses (including dyslexia, attention deficit hyperactivity disorder (ADHD), dyscalculia, autism, hearing impairment, and language problems other than dyslexia). Dyscalculia and ADHD are two common dyslexia comorbidities (Germanò et al., 2010; Wilson et al., 2015). There is substantial evidence that people with autism have poor face recognition (Griffin et al., 2021), so we wanted to make sure that they were not the driver of the face recognition results of our object recognition task. We asked about hearing impairments as they might possibly affect reading development (Moeller et al., 2007) and understanding of verbal instructions.

Three questionnaires were administered: The Adult Reading History Questionnaire (ARHQ-Ice), and Behavioral Evaluation Questionnaire for Adults I and II. Questions were displayed individually and consecutively, in written form, but were also read to the participants via headphones. Next, participants performed a visual task to measure their configural and featural processing of faces and houses. Following this, they completed a lexical decision task that evaluated the effect of word length (further description and results of this task can be found in Supplementary Materials), and finally, their reading was assessed with the IS-FORM and IS-PSEUDO reading tests.

Adult Reading History Questionnaire (ARHQ-Ice)

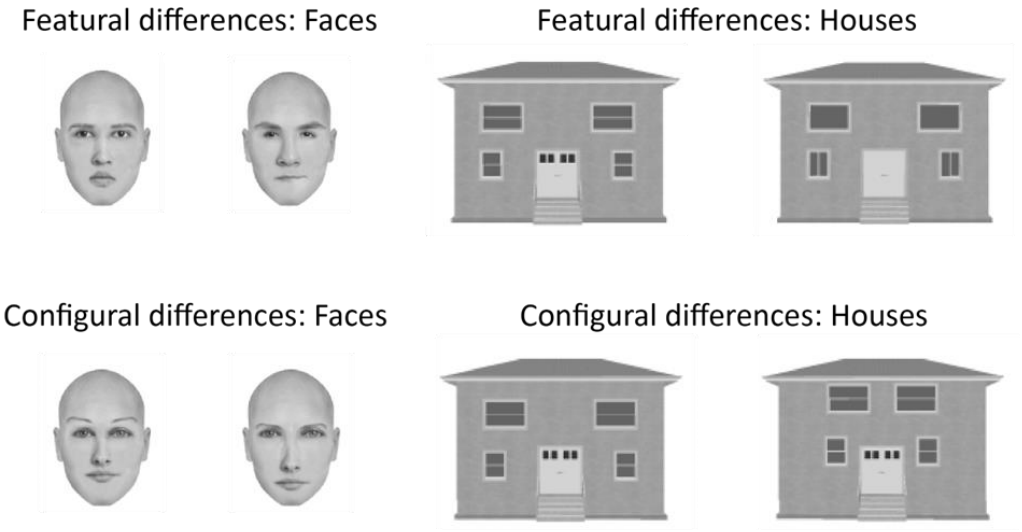
The original ARHQ is a 23-item self-report questionnaire (responses made on a 5-point Likert scale) developed by Lefly and Pennington (2000). We used the Icelandic version of the ARHQ (ARHQ-Ice), a valid and reliable test for assessing the reading difficulties of adults (Bjornsdottir et al., 2014). As suggested by Bjornsdottir et al. (2014), question number 15 in the Icelandic version was excluded from analysis because of truncated range. Scores on ARHQ-Ice range from 0 to 1, where higher scores indicate more difficulties in reading. The suggested cut-off score for dyslexia screening is 0.43 or higher (Bjornsdottir et al., 2014).

Behavioral Evaluation Questionnaire for adults I and II

The Behavioral Evaluation Questionnaire for adults I and II was used to evaluate ADHD symptoms according to DSM-IV criteria. These self-report questionnaires are reliable and valid tools for screening for ADHD (Magnússon et al., 2006). After listening to instructions, participants answered questions on a four-point Likert scale. The first questionnaire was used to evaluate ADHD symptoms in the last six months and the second one measured ADHD symptoms during childhood from ages 5 to 12 years. The total scores on each questionnaire can vary from 0 to 54, where scores above 25.8 on the childhood ADHD measure, and scores above 32.5 on the current ADHD measure, are considered indicators of ADHD.

Object Recognition Task

The object recognition task was administered to measure featural and second-order configural processing of faces and houses. Second-order configural processing (where second-order relations are manipulated, see section Introduction: Differences in processing?) will from now on be referred to simply as configural processing. Stimuli were part of a larger set with different featural and configural information developed by Collins et al. (2012; we thank Jane E. Joseph for providing the stimuli). In our task, there were 192 trials for each combination (featural faces, configural faces, featural houses and configural houses).



Featural differences: Faces

Featural differences: Houses

Configural differences: Faces

Configural differences: Houses

Figure 2. Examples of faces and houses with no common features or configurations (difficulty level: 0 shared features/configurations).

For each sample face/house image, there was a pair of images: a match and foil. The match was identical to the sample, but the foil was different featurally or configurally with 0, 1, 2 or 3 features/configurations in common with the sample (difficulty level; see figure 2). For example, if difficulty level was 1, the sample and foil shared 1 feature (e.g., same eyes or door) or 1 configuration (e.g., same space between eyes or windows) while if difficulty level was 2, they shared 2 features (e.g. identical eyes and nose or identical door and windows) or 2 configurations (e.g. distance between the two eyes, and between the eyes and the nose, or between the two windows, and between the windows and the door).

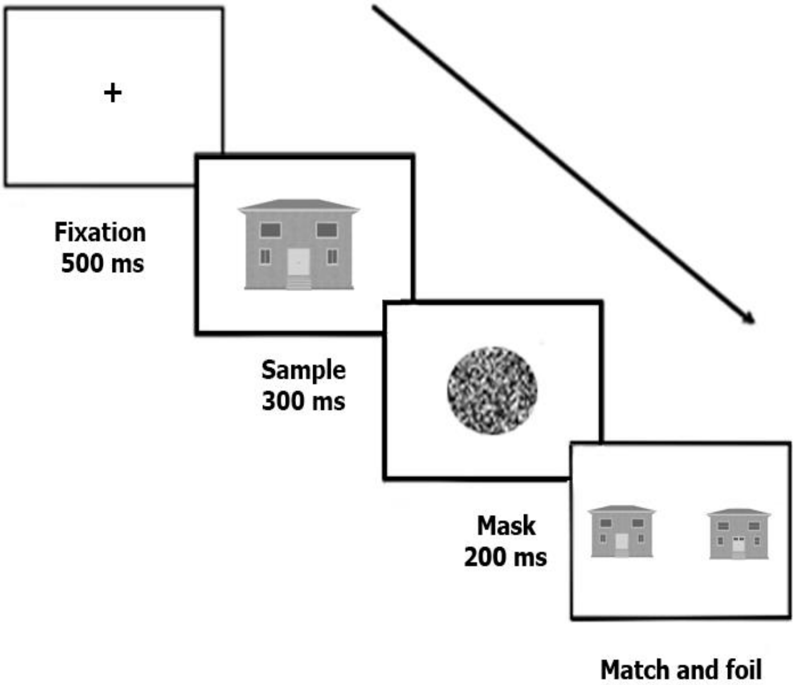


Figure 3. The experimental design. On each trial, a sample image (face or house) appeared at screen center followed by match and foil images displayed simultaneously to the left and right of screen center. The match image was identical to the sample image, but the foil image was different either featurally or configurally at four possible levels (difficulty levels 0-3). Match and foil images remained onscreen until the participants indicated with a button press which of the images was identical to the sample image.

Participants were asked to respond quickly while minimizing errors. After listening to instructions, participants completed a short practice test followed by the main object recognition task (figure 3). Both were two-alternative forced-choice delayed match-to-sample tasks. Trials had the same randomized order for each participant. In the practice test, 20 stimulus-pairs of simple geometrical objects were used (e.g. circle, square), and participants received feedback about the correctness of each response.

Each trial in the main task started with a 500 ms fixation cross at screen center. After the fixation cross disappeared, a sample image (face or house) appeared at screen center for 300 ms followed by a circular random dot mask (~8° in diameter) displayed at the same location for 200 ms. Subsequently, match and foil images appeared simultaneously approximately 3° to the left and right of screen center. Sample images were roughly 4° x 4° while foil and match images were 2.5° x 2.5°. The match and foil images were smaller to minimize the use of low-level template matching. Match and foil images remained onscreen until participants indicated with a button press which image was identical to the sample image. The "Z" button corresponded to the left image and the "M" button to the right image (both marked with yellow stickers). After response, the stimuli disappeared, and the next trial started after a 500 ms inter-trial interval (fixation cross around 0.4° × 0.4°). No feedback was given. After every 50 trials, participants were given a break, and pressed the space bar when ready to continue.

Our main focus was on accuracy even though we additionally assessed reaction times, as previous studies have shown that group differences between dyslexic and typical readers are found for accuracy (Sigurdardottir et al., 2018; Sigurdardottir et al., 2015).

IS-FORM and IS-PSEUDO Reading Tests

The IS-FORM and IS-PSEUDO reading tests (Sigurdardottir et al., 2015; 2017) were used to measure reading ability. The tests were computerized and included three lists consisting of 128 common Icelandic word forms, 128 uncommon Icelandic word forms, and 128 pseudowords. Participants were informed by recorded instructions that they had to read the (pseudo)words of each list out loud as fast as possible while keeping errors to a minimum. Participants pressed the space bar and a countdown began, thereafter they read a short practice test out loud. Following this, the researchers left the room, and the three lists were presented consecutively following a countdown in each case, and participants read each list out loud. Two main factors were analyzed in these tests: reading speed (number of words or pseudowords read per minute) and reading accuracy (percentage of correctly read words or pseudowords).

Statistical Analysis and Results

Data Analysis Overview and Exclusion/Inclusion

Before calculating mean response times (RTs), trials with RTs deviating by more than three standard deviations from the mean of each participant for each condition of the object recognition task (e.g., featural processing of faces with difficulty level of 1) and lexical decision task (e.g., 4 letters) were excluded; other trials were included, regardless of accuracy. D_{av} refers to a paired Cohen's d which is calculated as the difference between the means of each group, all divided by the standard deviation of the paired difference scores. The paired difference, and thus the D_{av} and the confidence intervals (*CI*), are always based on typical reader scores minus the paired dyslexic reader scores. Due to recording failure, three participants (one dyslexic reader and two typical readers) had missing values for reading speed and accuracy for the IS-FORM lists. To impute the missing values, we first estimated the relationship between ARHQ-Ice and reading speed and accuracy for the IS-FORM/IS-PSEUDO lists from a previous larger independent dataset with linear regression (ARHQ-Ice as an independent variable and reading speed and accuracy as dependent variables). Subsequently, we used the regression coefficients from this regression model to estimate imputed values for missing reading speed and accuracy values of the current study from ARHQ-Ice scores. Average reading speed (words per minute) and average reading accuracy (percent of correctly read words) were then calculated across the three reading lists for each participant. For additional analyses excluding these participants, see Supplementary Information, Correlations of Object Recognition, Reading Speed and Reading Accuracy After Excluding Participants with Missing Reading Scores. ARHQ, reading speed and reading accuracy were used to verify proper group assignment using binary logistic regression, see Results: Verification of Group Classification. Our main analyses were based on accuracy measures from the object recognition task, and included comparisons of group means, comparison with reading measures, and representational similarity analysis (RSA), as detailed below. Alpha levels were set to 0.05 and all statistical tests were two-sided.

We compared object recognition of matched dyslexic and typical readers with paired t-tests to see whether featural face, featural house, configural face, and configural house accuracy differed between these two groups. Then, we used a 2 x 2 x 2 repeated measures ANOVA to assess any interactions, with group (dyslexic and typical readers), stimulus (faces and houses), and process (featural and configural) as factors, and accuracy of object recognition as the dependent measure. We calculated zero-order correlations to estimate the association between performance on the object recognition task (accuracy for the four subtasks, as well as total accuracy of faces, houses, featural, and configural processing) and the three measures of reading performance (ARHQ-Ice, reading speed, reading accuracy). We additionally calculated partial correlations to estimate the specificity of such associations.

We also used representational similarity analyses (RSA) to compare correlational matrices of object recognition tasks (reference models) to predicted data patterns (conceptual models). Three examples of conceptual models (stimuli, processes, and difficulty levels) are shown in figure 7, panel B. The stimuli model will fit the data well if individuals mainly differ in their ability to discriminate/recognize faces vs. houses. The processes model fits the data well if individuals mainly differ in their ability to discriminate/recognize objects by the use of featural vs. configural processing. The difficulty level model will fit the data well if individuals mainly differ in their ability to discriminate/recognize objects of different difficulty levels.

Reading Ability and History of Reading Problems

Dyslexic readers reported a greater history of reading problems than typical readers on the ARHQ-Ice. They also read less accurately and more slowly than typical readers on the IS-FORM and IS-PSEUDO reading tests (table 1). In what follows, "reading speed" and "reading accuracy" respectively, refer to the average speed, i.e. (pseudo)word forms read per minute regardless of accuracy, and average percent of correctly read (pseudo)word forms across the IS-FORM common word forms, IS-FORM uncommon word forms, and IS-PSEUDO pseudoword forms. For Cumming estimation plot of Reading Speed and Reading Accuracy After Exclusion of Misclassified Participants, see Supplementary figure s8.

Table 1. Descriptive statistics and summary of paired samples t-tests for reading abilities and history of reading problems of dyslexic and typical reader groups. SD = standard deviation.

	Dyslexic		Typical				
	Mean	SD	Mean	SD	t(33)	P	D_{av}
ARHQ-Ice	0.70	0.12	0.32	0.13	-13.18	< .001	-3.04
IS-FORM Common							
Word accuracy (%)	93.66	4.74	96.81	3.56	3.39	.002	0.76
Words/minute	64.49	17.62	100.37	20.43	6.89	< .001	1.88
IS-FORM Uncommon							
Word accuracy (%)	81.35	10.31	94.07	7.09	6.12	< .001	1.46
Words/minute	41.06	12.57	71.76	15.76	9.14	< .001	2.17
IS-PSEUDO							
Pseudoword accuracy (%)	62.77	19.09	86.29	12.74	6.09	< .001	1.48
Pseudowords/minute	31.40	16.09	48.96	13.10	4.55	< .001	1.20

Verification of Group Classification

A binary logistic regression was run with ARHQ-Ice scores, reading speed, and reading accuracy as predictors and group as a dependent variable. The confusion matrix showed that group membership was correctly predicted in 95.6 percent of cases. Three participants were misclassified, one from the dyslexic reader group (classified as a typical reader) and two from the typical reader group (classified as dyslexic readers). They were excluded from the analyses along with their matched participants. After exclusion, 62 participants remained, where thirty-one of them reported a previous diagnosis of dyslexia (19 women; mean age: 36.9 years, range 18-62) and another 31 were self-reported typical readers (19 women; mean age: 36.5, range 18-67). In each group, 5 people had completed the first level of schooling (high school), 14 the second level (gymnasium), 7 the third level (undergraduate degree), and 5 had completed the fourth level (graduate degree). Unless otherwise noted, all further analyses are limited to the remaining 62 participants.

Other Disorders

All participants reported normal or corrected-to-normal vision, two dyslexic readers and two typical readers reported hearing impairments, 10 dyslexic readers reported dyscalculia (consistent with previously described co-occurrences of dyslexia and dyscalculia, e.g. Wilson et al., 2015), one dyslexic reader reported an autism spectrum disorder and one reported language problems other than dyslexia. A previous diagnosis of ADHD was reported by 10 dyslexic readers and three typical readers, which is consistent with the well-known connection between dyslexia and ADHD (see e.g. Germanò et al., 2010). Also, in the screening questionnaires for ADHD (Behavioral Evaluation Questionnaire for Adults I and II), 13 dyslexic readers and eight typical readers scored over suggested screening scores on the childhood ADHD measure, and 5 dyslexic readers and 1 typical reader on the current ADHD measure. Dyslexic participants had significantly higher scores than typical readers for both childhood (dyslexic readers: M = 22.52, SD = 15.65; typical readers: *M* = 14.42, *SD* = 13.55; paired samples t-test, $t(30) = -2.46$, $p = .02$, $D_{av} = -0.55$, 95% *CI* [-14.81,-1.38]), and current ADHD symptoms (dyslexic readers: *M* = 16.68, *SD* = 12.38; typical readers: *M* = 11.16, *SD* = 8.97; paired samples t-test, $t(30) = -2.10$, $p = .04$, $D_{av} = -0.52$, 95% *CI* [-10.88, -0.16]).

Object Recognition Task: Overall Group Differences and Correlations

The reaction times of the two groups were comparable (paired samples t-tests, all four subtasks *p*s > 0.58, $D_{avs} > 0.14$), but dyslexic readers were less accurate than matched controls at recognizing houses both featurally and configurally. The dyslexic group was significantly less accurate at featural processing of houses (*M* = 75.66%, *SD* = 5.61%) than the typical group (*M* = 78.70%, *SD* = 5.60%; $t(30) = 2.13$, $p = .04$, $D_{av} = .54$, 95% *CI* [0.13, 5.96]) as well as at configural processing of houses (dyslexic readers *M* = 80.70, *SD* = 7.47; typical readers *M* = 85.13%, *SD* = 6.75%; $t(30) = 2.53$, $p = .02$, $D_{av} = 0.62$, 95% *CI* [0.86, 8.01]). Notably, however, neither featural nor configural processing of faces differed significantly between the groups (see figure 4; featural faces:

dyslexic readers *M* = 81.23%, *SD* = 6.61%; typical readers *M* = 81.97%, *SD* = 6.48%; *t*(30) = 0.50, *p* = .62, D_{av} = 0.11, 95% *CI* [-2.27, 3.75]; configural faces: dyslexic readers *M* = 76.33%, *SD* = 8.63%; typical readers *M* = 77.70%, *SD* = 7.36%; t(30) = 0.78, *p* = .44, D_{av} = 0.17, 95% *CI* [-2.24, 5.00]; for group differences additionally broken up by difficulty levels, see Supplementary figure s1; for group differences after excluding participants with potential comorbidities, see Supplementary Information).

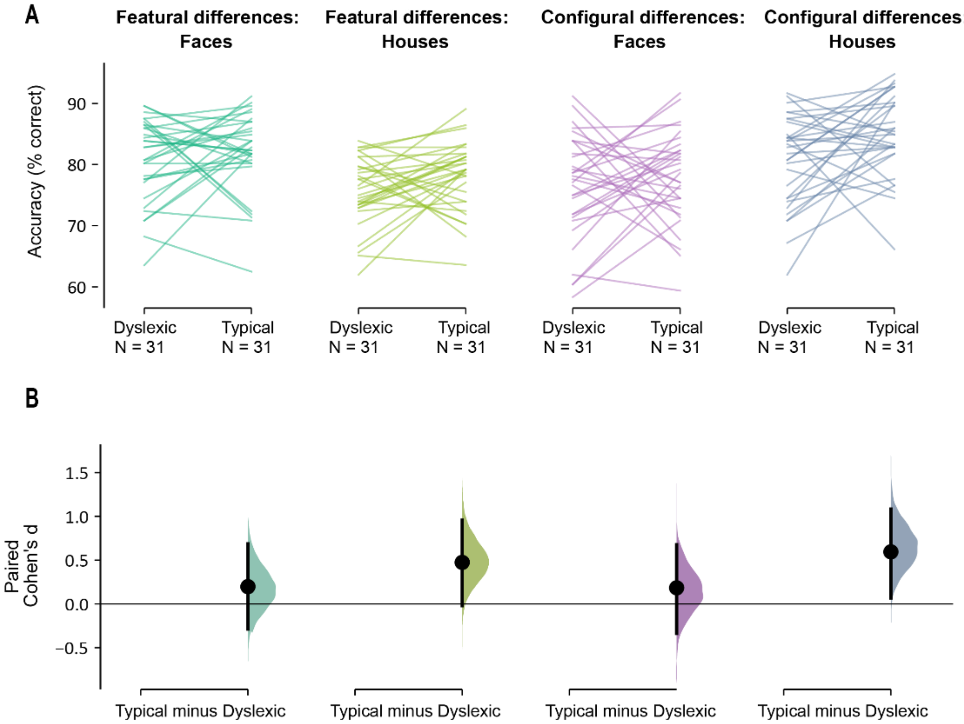


Figure 4. Cumming estimation plot for paired Dav for featural and configural processing of faces and houses. A. The raw data is plotted for each paired set of observers connected by a line; B. Each paired mean difference is plotted as a bootstrap sampling distribution (5000 bootstrap samples were taken). Mean differences are depicted as dots; 95% confidence intervals are indicated by the ends of the vertical error bars, and they are bias-corrected and accelerated.

We performed a 2 x 2 x 2 repeated measures ANOVA (table 2) with group (dyslexic and typical readers), stimulus (faces and houses) and process (featural and configural) as factors and accuracy as the dependent measure. The main effect of group was not significant according to a two-sided test (*p* = .07). No main effects of stimulus and process were found. The interaction of stimulus and process was significant (*p* = <.001) indicating that featural processing of houses was less accurate than featural processing of faces, while configural processing of houses was more accurate than configural processing of faces. There were no other significant interactions (between group and process, group and stimulus, or group, process, and stimulus).

Table 2. Summary of 2 x 2 x 2 repeated measures ANOVA. ^η_G^2 indicates generalized eta-squared.

	F	p	$\hat{\eta}_G^2$
Group	3.48	.07	.03
Stimulus	0.61	.44	.002
Process	0.64	.43	.001
Group × Stimulus	2.65	.11	.01
Group × Process	1.51	.23	.002
Stimulus × Process	164.20	< .001	.12
Group × Stimulus × Process	0.50	.49	<.001

The lack of main effects and interactions with group was surprising, although the non-significant main effect of group should be interpreted in the context that the test is conservative given that our hypothesis was clearly one-sided (i.e., dyslexic readers were predicted to do worse than typical readers). One possible reason is that visual recognition problems are modulated by educational level, as we have previously seen that group effects in visual recognition were solely driven by dyslexic readers without higher education (Sigurdardottir, Hjartarson, et al., 2019, their figure 3). This was supported by a significant three-way interaction between educational level, group, and stimulus, where the group difference was larger for houses compared to faces, but only for participants with lower educational levels. For further information, see "Repeated measures ANOVA with covariates" in the supplementary materials. We return to the group x process null result in the subchapter on representational similarity analysis (RSA).

Reading History Problems, Reading Speed, and Reading Accuracy

We also assessed the association between performance on the visual tasks included in the object recognition task (accuracy for the four subtasks, as well as total accuracy for faces, houses, featural, and configural processing) and the three measures of reading performance (ARHQ-Ice, reading speed, reading accuracy). The results are summarized in figure 5. For additional analyses excluding participants with missing values for reading speed and accuracy for the IS-FORM/IS-PSEUDO reading tasks, and broken down by group membership, see Supplementary Information, Correlations of Object Recognition, Reading Speed And Reading Accuracy After Excluding Participants With Missing Reading Scores .

Featural face accuracy was not significantly correlated with any measures of reading performance, while configural face accuracy as well as total face accuracy were significantly associated with greater reading speed and accuracy. Featural house, configural house, and total house accuracy were significantly correlated with all measures of reading performance, where lower accuracy was associated with a greater history of reading problems, slower reading, and less accurate reading. Featural and configural processing (irrespective of stimulus type) were correlated with all measures of reading.

For a better understanding of the specificity of the relationship between performance on visual tasks and reading problems, we calculated partial correlation coefficients (figure 5). Total house accuracy was significantly correlated with all reading measures when total face accuracy was controlled for, showing that lower accuracy for houses was particularly related to reading difficulties. In addition, configural face accuracy was significantly correlated with reading accuracy when featural face accuracy was kept constant. Other partial correlations were not significant. For correlations of object recognition, reading speed, and reading accuracy divided by group, see Supplementary Information.

Figure 5. Correlations between reading measures and accuracy on the object recognition task. The three leftmost panels show zero order correlations (Pearson's r) and the three rightmost panels show partial correlation coefficients. Each box on the right-hand side indicates mutually partialled-out variables. The upper row of the top-most box shows correlations with configural processing accuracy when featural accuracy is partialled out, while the lower row of the same box shows correlations with featural processing accuracy when configural accuracy is partialled out. The upper row of the second box from the top shows correlations with house accuracy when face accuracy is partialled out, while the lower row of the same box shows correlations with face accuracy when house accuracy is partialled out. The upper row of the third box from the top shows correlations with configural house accuracy when featural house accuracy is partialled out, while the lower row of the same box shows correlations with featural house accuracy when configural house accuracy is partialled out. The upper row of the bottom box shows correlations with configural face accuracy when featural face accuracy is partialled out, while the lower row of the same box shows correlations with featural face accuracy when configural face accuracy is partialled out. Asterisks indicate significance levels: *p < .05, **p < .01, *** p < .001.

Representational Similarity Analysis (RSA)

While univariate methods are useful for comparing group averages, they cannot detect all informative data patterns. For example, how well featural performance predicts configural performance (and vice versa) can inform us on how differentiable these processes are. By relying solely on group averages, such relationships will be overlooked. In our case, dyslexic and typical readers could have comparable accuracy but the relations between conditions might still differ between the two groups. As an example, in general we would expect performance on featural trials to be more correlated with other featural trials than configural trials, and vice versa. However, this expected pattern might not be apparent in dyslexic readers if featural and configural trials are not actually differently processed. Representational similarity analysis (RSA), which originates in systems neuroscience, can provide a fuller description of the structure of information representation in each group. We use RSA to correlate individual responses within each group and evaluate the similarity of these correlation matrices (reference models) with predicted data patterns (conceptual models; Kriegeskorte et al., 2008).

To illustrate, assume that we have two conditions "A" and "B" (see figure 6). We will overlook the relationship between "A" and "B" if we just compare the total accuracy of "A" or "B" between two groups of dyslexic and typical readers. Two groups might be equally good at condition "A" or "B" but employ different strategies to achieve that. If "A" and "B" tap into different skills or require different strategies, the accuracy of a trial with task "A" should be more correlated with the accuracy of another trial of task "A" than a trial with task "B". In this case, the left pattern of figure 6 will be our conceptual model that assumes "A" and "B" are independent skills or require different strategies. "If "A" and "B" are not independent but instead tap into the same skills, then a trial from "A" should be no more correlated with another trial with "A" than it is with a trial with "B", and we would see the right pattern in figure 6. RSA measures how well such patterns or conceptual models correspond to the real data.

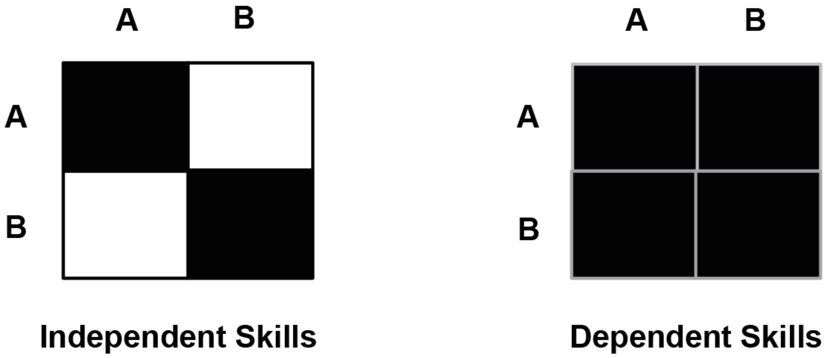


Figure 6. illustrative conceptual models. Darker colors indicate higher correlation.

In the first step of the RSA, we developed our reference models, one for dyslexic readers and another for typical readers. The reference models are correlation matrices of the accuracy of featural and configural processing of faces and houses with the four different difficulty levels (figure 7, panel A). We calculated the mean accuracy for each participant on all 16 combinations of trial types, including two different stimulus types (faces or houses), two different process types (featural or configural), and four different difficulty levels (0, 1, 2, 3). Each reference model cell represents the correlation across participants of two trial types, such as the correlation between people's accuracy for featural faces with the difficulty level of 1 and their accuracy with configural houses with the difficulty level of 2.

We then created three conceptual models based on possible predicted patterns for stimuli, processes, and difficulty levels. For the stimulus matrix, the values of the conceptual model were "1" when the stimuli on the compared trial types came from the same category (both faces or both houses) but were otherwise set to "0" (one face and the other house). A similar pattern would be found in a reference model if people's performance on a stimulus category was better predicted by their performance on other trials from that category than trials from a different category. The process matrix, the second conceptual model, followed the same logic, but for processes (featural versus configural): the provided values were "1" when the processes matched (both featural or both configural), but "0" otherwise (one featural and the other configural). A reference model would be expected to follow this pattern if the putative processes were separable, so a person's performance with one type of process would better predict another trial where that same process was used, as opposed to a trial with the other process. A scale (1, 0.75, 0.25, 0) was used for the conceptual model of difficulty levels, the difficulty level matrix, where greater numbers indicated greater similarity in difficulty; "1" therefore stood for identical difficulty levels and "0" stood for the most different levels (figure 7, panel B). It should be noted that we did not necessarily expect to see this pattern in the reference models, as it would require performance on trials of a specific difficulty level to best predict performance at that same difficulty level. Different difficulty levels might however not depend on distinct mechanisms; if a participant performs well on easy trials compared to others, then he or she might also perform comparatively well on more demanding trials. This conceptual model was therefore included mainly for the sake of completeness.

To reduce the possibility of spuriously low p-values, each model's diagonal (which has correlational coefficients of 1 by definition) and one off-diagonal triangle (a mirror version of the other off-diagonal triangle) was omitted and then converted into a vector before comparing reference models with the three conceptual models (stimuli, processes, and difficulty levels). Two multiple regression analyses were performed, one with the typical reader reference vector and another one with the dyslexic reader reference vector as dependent variables. The three conceptual vectors were used as independent variables to estimate to what degree stimuli, processes, and difficulty levels contributed to the reference vectors for dyslexic and typical readers.

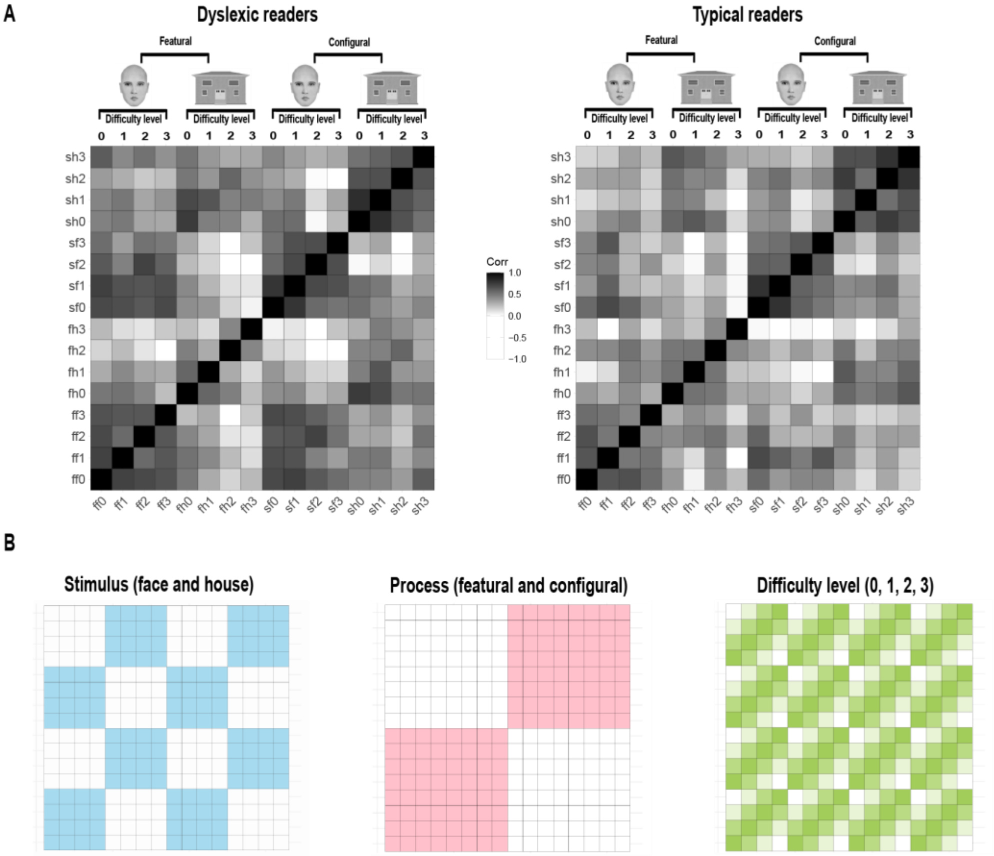


Figure 7. A. Reference models for Representational Similarity Analysis (RSA): Correlation matrices of dyslexic and typical readers; B. Conceptual models. Left: Stimulus matrix; Middle: Process matrix; Right: Difficulty level matrix. Abbreviations: Corr: correlation (Pearson's r); ff: featural faces; fh: featural houses; sf: second-order configural faces; sh: second-order configural houses; 0-3: difficulty levels.

As summarized in table 3, a significant regression equation was found for typical readers ($p < .001$). Both stimulus ($t(116) = 8.13$, $p < .001$) and process ($t(116) = 4.40$, $p < .001$), but not difficulty levels ($t(116) = 0.86$, $p = .39$), significantly predicted the typical reader reference model. The regression equation was also significant for dyslexic readers ($p < .001$), but the only significant independent predictor for the dyslexic reader reference model was stimulus ($t(116) = 9.29$, $p < .001$), while process ($t(116) = 0.65$, $p = .52$) and difficulty levels ($t(116) = 1.32$, $p = .19$) were not significant. Stimulus type (face vs. house) dominated the patterns of both typical and dyslexic readers and both groups performed differently for different stimulus classes, but how well the data pattern represented processes (featural vs. configural) varied considerably between the two groups. Performance was more consistent for typical readers when process was constant. Dyslexic readers did not perform differently when processes were expected to differ, and seemingly relied on a single process that was indistinguishable for supposed featural and configural trials. Group differences in the separability of processes could not be attributed to disorders other than dyslexia (see Supplementary Information, Representational Similarity Analysis (RSA): Group Differences in The Separability Of Processes After Excluding Participants With Other Disorders).

Table 3. Summary of representational similarity analyses (RSA). Multiple regressions with dyslexia or typical vectors as dependent variables, and stimulus, process and difficulty level as predictors.

	Dyslexic				Typical			
Predictor	**b**	**95% CI**	**t (116)**	**P**	**b**	**95% CI**	**t (116)**	**p**
Intercept	0.23	[0.16, 0.31]	6.27	< .001	0.17	[0.11, 0.24]	5.05	< .001
Stimulus	0.28	[0.22, 0.34]	9.29	< .001	0.23	[0.17, 0.28]	8.13	< .001
Process	0.02	[-0.04, 0.08]	0.65	.52	0.12	[0.07, 0.18]	4.40	< .001
Difficulty level	0.06	[-0.03, 0.14]	1.32	.19	0.03	[-0.05, 0.11]	0.86	.39

We redid the RSA separately for faces and houses to explore whether process differences between dyslexic and typical readers could be identified for both stimulus categories. Process was not a significant predictor of the reference model for dyslexic readers, and this was true for both faces ($b = 0.05$, $t(25) = 1.41$, $p = .17$) and houses ($b = 0.05$, $t(25) = 0.85$, $p = .40$). Importantly, for typical readers, process for faces was a significant predictor ($b = 0.15$, $t(25) = 2.87$, $p = .01$), and process for houses was close to significance ($b = 0.14$, $t(25) = 1.96$, $p = .06$). Therefore, dyslexic readers apparently rely on only one process, and do not distinguish between supposedly featural vs. configural faces, or featural vs. configural houses.

As the results of Sigurdardottir, Hjartarson, et al. (2019) and our univariate analyses (see Supplementary Information, Repeated measures ANOVA with covariates) showed that the visual problems of dyslexic readers were modulated by educational level, we investigated whether group differences in the separability of processes were also restricted to lower educational levels and found out that it was not. For this, we developed a novel method for doing individual statistics on RSA analyses. For details, see Supplementary Information, Representational Similarity Analysis (RSA): Possible Modulation by Educational Level .

To assess the likelihood that the difference in process between the groups was due to chance, we performed a two-sided permutation test where the group labels (dyslexic and typical readers) were randomized. Next, RSA was calculated based on the permuted labels, i.e. two randomized ("fake") groups were created, and their reference models were calculated. We then computed the difference between the unstandardized *b*s of process, stimulus, and difficulty level for the randomized groups, repeating these steps 10,000 times and comparing the distribution of these random group differences to the *b* differences for the original unpermuted RSA. There was no significant difference between permuted and unpermuted RSAs for stimulus (p = .72) and difficulty level (p = .79), but the difference for process was significant (p = .02) suggesting that the observed process difference in dyslexic and typical readers is real.

Discussion

Our first major goal was to examine whether dyslexic readers have problems with recognizing non-word visual objects, and whether any such deficits are domain-specific, restricted to the processing of written words, or domain-general, affecting the recognition of other visual stimuli, such as faces and houses. Visual problems in dyslexia generalized to other visual domains, and seemingly more to some visual categories than others, as our participants with dyslexia had difficulty with house recognition but not face recognition. Furthermore, lower accuracy for houses was related to reading difficulties even when accuracy for faces was kept constant. A second major goal was to investigate whether visual problems in dyslexia are specific to either featural or configural processing. Initially, problems in dyslexic readers did not seem to be process specific as there was no detectable group difference in mean accuracy for featural vs. configural processing. But representational similarity analyses (RSA) revealed differences in the recognition processes used by dyslexic and typical readers that were not detected by traditional univariate analyses. Featural and configural processes were clearly separable in typical readers, while dyslexic readers appeared to rely on a single process to identify visual objects. This effect for processes was general, not restricted to either faces or houses.

How can our findings be integrated with previous studies that have produced contradictory results? As in the current study, Sigurdardottir et al. (2015) found problems in non-face object recognition of dyslexic readers and in Brachacki et al. (1995) dyslexic readers recognized traffic signs less accurately than typical readers. Huestegge et al. (2014) showed that dyslexic readers have problems in representing highly detailed visual objects in long-term memory. Conversely, Gabay et al. (2017) and Sigurdardottir et al. (2018) failed to find problems with object perception in dyslexic readers. This discrepancy could be rooted in methodological differences. The tasks were perceptual in Gabay et al. (2017) where participants estimated whether two simultaneously presented pictures of cars were identical. Similarly, participants in Sigurdardottir et al. (2018) matched simultaneously presented novel objects ("YUFOs") with which observers had no visual experience. Conversely, the current tasks and in Sigurdardottir et al. (2015) involved a memory component where participants had to indicate which objects they had seen previously. In our current study, we nonetheless found intact memory for faces so dyslexic readers are not likely to have a general problem with poor memory. Additionally, Sigurdardottir et al. (2015) assessed color memory which appeared to be intact in dyslexic readers, and visual problems of dyslexic readers in Sigurdardottir, Hjartarson, et al. (2019) were not connected to verbal working memory problems.

Another factor that might explain the apparent contradictions is education level. Dyslexic readers with lower educational levels are likely to have more severe dyslexia. The difference in house recognition accuracy between dyslexic and typical readers was largely driven by those with lower educational levels, while dyslexic readers with higher educational levels displayed performance comparable to matched typical readers. Earlier research has also suggested that visual recognition problems of dyslexic readers differ by education levels (Sigurdardottir, Hjartarson, et al., 2019), although here we found this for house recognition while the earlier study reported this for face recognition. This may explain why Gabay et al. (2017) did not find a problem with non-face object recognition, since their participants were university students.

Nevertheless, while dyslexic readers' problems were not restricted to recognizing words, they had no detectable difficulty with face recognition, and dyslexia in this study was only associated with non-face object recognition problems. This is compatible with the claim that domain-specific processes are involved in face processing (e.g. Yovel & Kanwisher, 2004), and that recognition of faces and words can be selectively affected by brain injury or developmental disorders (Robotham & Starrfelt, 2017). Intact face recognition in dyslexia stands in contrast to accounts suggesting a mutual dependency between words and faces (Behrmann & Plaut, 2019; Dehaene et al., 2010; Plaut & Behrmann, 2011). Our result is also somewhat at odds with results from the Back of the Brain project (Rice et al., 2020) where the authors claim that the general organizational principle for patients with posterior cerebral artery stroke was that of associations between word and face processing; a minority of patients did nonetheless show disproportionate deficits for word recognition. We should note that according to some accounts (Bishop, 1997; D'Souza & Karmiloff-Smith, 2011), findings of developmental deficits (e.g., dyslexia) are not necessarily related to normal cognition, as they postulate that brains of people with developmental deficits develop differently.

Recognizing faces and words is generally done daily, while recognizing houses is arguably a task in which people engage less frequently. The lack of a face recognition problem in our univariate analysis is not consistent with a visual expertise problem in dyslexia (Lieder et al., 2019; Sigurdardottir et al., 2017; 2018). Such a problem is expected to manifest as greater recognition difficulties for visual categories with which people have the most experience, such as faces and words. We nonetheless should not claim based on these data alone that visual recognition problems in dyslexic readers are completely unrelated to experience; houses are, after all, familiar objects. Neither should we make strong claims that people with dyslexia have problems with all non-face object recognition as houses might not be representative of all object categories. According to Richler et al. (2017), the assumption that non-face object processing has a common mechanism that varies little from one non-face category to another can be questioned. Several studies have reported dissociations for brain areas involved in recognizing animals vs. tools (e.g., Chao et al., 2002), large vs. small objects (e.g., Konkle & Oliva, 2012) or curvilinear vs. rectilinear objects (e.g., Nasr et al., 2014; Yue et al., 2014). Also, according to recent research, the mean pairwise correlation (r = 0.33–0.34) in performance across object recognition tests (e.g., butterflies, cars, planes, shoes, dinosaurs; McGugin et al., 2012; Van Gulick et al., 2016) was no higher than the usual correlation between face and non-face object recognition tests (e.g., r = 0.37 in Dennett et al., 2012). We chose our stimuli as they are well-controlled and span a wide range of difficulty levels (Collins et al., 2012) which can therefore capture a wide range of individual differences. To generalize our results, it might nonetheless be beneficial to include more visual categories in future investigations and manipulate featural and configural information in words.

Several studies have found impairments in face processing of dyslexic readers (Collins et al., 2017; Gabay et al., 2017; Sigurdardottir et al., 2018; Sigurdardottir, Hjartarson, et al., 2019; Sigurdardottir et al., 2015) while others have not (Brachacki et al., 1994; Holmes & McKeever, 1979; Rüsseler et al., 2003; Smith-Spark & Moore, 2009). We should note that configural face processing accuracy was positively correlated with reading accuracy when featural face processing accuracy was kept constant, although it was not associated with dyslexia (see partial correlations in figure 6). This is seemingly inconsistent with Ventura et al. (2013) who showed that in comparison to illiterates, literates process faces less holistically, but more consistent with Cao et al. (2019) who showed that illiterates were less sensitive to changes in the configural processing of faces and houses and concluded that later experience in reading can reshape configural processing. Cao et al. (2019) argued that the paradigm might be a key factor to explain the apparent discrepancies between these results, as they measured second-order configural processing, but Ventura et al. (2013) used the complete composite face paradigm, which represents the failure of selective attention.Therefore, inadequate experience with reading could possibly lead to both reduced reading accuracy and limited ability for second-order configural face processing.

As reviewed by Ventura (2014), visual word representations of expert readers may overtake cortical space that otherwise would have been dedicated to face processing. Reading acquisition could therefore trigger right hemispheric lateralization for faces, and additionally have detrimental effects on face processing abilities as suggested by Dehaene et al. (2010). Our data show that the accuracy of configural face processing is positively correlated with reading performance. This is more in alignment with the possibility that literacy enhances representational similarity between text and faces and reorganizes cortical function without inducing direct cortical competition with other visual categories (Hervais-Adelman et al., 2019).

One possible reason for conflicting results on visual processing in dyslexia is that some tasks involve a visual processing mechanism deficient in dyslexia, while dyslexic readers can in other cases rely on a separate intact visual mechanism. However, in the current study, the problem with houses was found for both featural and configural manipulations, and we found neither featural nor configural processing difficulties for face recognition in the univariate analysis of accuracy. This may suggest that dyslexia is not process specific and is, at a first glance, inconsistent with studies showing a featural processing deficiency (Sigurdardottir, Arnardottir, et al., 2019) in dyslexic readers. But there are several alternative explanations for this result. According to Rakover (2002), configural changes such as those used in the current study may inadvertently involve featural changes, and vice versa. For example, manipulating the space between the eyes can be interpreted as a featural change in the nasal bridge. Furthermore, we modified second-order relations, one type of configural information, but other types of configural processing may be unrelated to it (Maurer et al., 2002). Rezlescu et al. (2017) argue against the use of a single term for configural processing, because different tasks reflect distinct perceptual mechanisms. The same may even be true for featural processing. For example, Sigurdardottir, Arnardottir, et al. (2019) used a different way of manipulating featural vs. configural processing; while we changed the distances between features for configural processing, they manipulated the form of the skull, muscles, and fat structure. For featural manipulation, internal features from one face were transferred to another face with a different global form (for details, see Van Belle et al., 2009). These different manipulations might lead to similar results, but the processes for performing these tasks may differ. Similar to the different manipulations that supposedly all measure "configural processing", different featural manipulations might not assess the same underlying mechanisms. Further research is required to investigate similarities and differences in these methods. But to summarize, the use of particular stimulus manipulations does not guarantee "featural" and "configural" processing.

Differences in task demands may also contribute to literature inconsistencies. For example, in our face task the sample and the match were identical images, and participants were required to hold the sample face in memory for 300 ms (visual short-term or working memory), while previous studies reporting group differences in face processing used visual tasks that were based on long-term memory and/or non-identical samples and matches (e.g. Gabay et al., 2017; Sigurdardottir, Arnardottir, et al., 2019; Sigurdardottir et al., 2018; Sigurdardottir et al., 2015). Even though we included size differences between sample and choice images specifically to discourage low-level strategies, the current task may still have enabled template matching or other methods that do not tax invariant high-level visual processes. Left hemisphere regions hypoactive in dyslexic readers (Richlan et al., 2011) might be particularly relevant for invariant object recognition – hypothesized to involve a feature- or part-based strategy – so weaknesses in such a system are likely less noticeable when a task can be solved with a holistic processing strategy, by matching to specific exemplars (Marsolek, 1999; but see Curby et al., 2004).

Also, as in our task participants were first presented with a sample that was either a face or a house, they were aware that the task would involve a particular object category, but the optimal processing strategy – featural vs. configural – was unknown until the match and foil appeared. Participants could therefore have changed strategy primarily based on stimulus category and not process. This could be why the process conceptual model pattern is not very apparent even for typical readers, as depicted in figure 7. While the univariate analysis did not reveal process differences, it was still clear from the more sensitive RSA analysis that the typical readers indeed solved the featural and configural tasks using different strategies, while the dyslexic readers did not.

Importantly, univariate analyses of accuracy miss that two people with equivalent performance can solve a task using different strategies. Even though no overall group differences were found for featural vs. configural accuracy, this does not mean that the tasks were solved using similar representations or the same underlying mechanisms. Our RSA analyses revealed differences in the visual recognition processes used by the two groups. To clarify, if two different processes (one featural and another configural) or stimuli (one face and another one house) were in fact supported by separable mechanisms, we expected performance for one type of process or stimulus to better predict a trial with an equivalent process or stimulus than a trial with a different process or stimulus. The RSA results showed that typical readers performed differently depending on both stimulus category (faces vs. houses) and processing type (featural vs. configural), while dyslexic readers performed differently based on stimulus category, but not process. The RSA also showed that process differences between dyslexic and typical readers are not restricted to only one stimulus category as the pattern was observed for both faces and houses.

These group differences in process might be experience-dependent. If dyslexic readers have problems with acquiring expertise (Brachacki et al., 1995; Lieder et al., 2019; Sigurdardottir et al., 2017), then the underlying processing mechanisms could differ from typical readers with acquired expertise, as experience with a stimulus category can affect how it is processed (Bukach et al., 2006; Gauthier & Bukach, 2007; Gauthier & Tarr, 2002; Hsiao & Cottrell, 2009; Zhang & Cottrell, 2005). Cao et al. (2019) also suggest that literacy acquisition can reshape configural processing in general. The lack of a process by education interaction in our RSA results however makes it less likely that processing differences between dyslexic and typical readers are related to differences in reading experience as dyslexic readers with different educational backgrounds likely have considerably different reading experience. Another possibility is that some cases of dyslexia are a product of fewer available intact visual processes, as flexibly switching between different processes can be important for recognizing objects. Exactly pinning down these visual processes is a topic worthy of further study.

Developmental dyslexia is not the only developmental disorder where featural or configural processing could differ. For example, a meta-analysis revealed that people with autism are slower at global or configural perception than controls (Van der Hallen et al., 2015). Another study suggested that children with autism prefer to report the local properties of a stimulus (Koldewyn et al., 2013). When asked to report the global properties, their performance was comparable to that of the control group, indicating that they have a disinclination in using global processing rather than a disability. Kalanthroff et al. (2013) reported no differences in interference between irrelevant global stimuli and irrelevant local stimuli, implying that there is no global-to-local interference in adults with ADHD. Song and Hakoda (2012, 2015) argued that people with ADHD have local-to-global interference rather than global-to-local interference. As indicated in the supplementary material, omitting people with autism or ADHD, as well as their matched participants, had only a minor impact on our Representational Similarity Analysis (RSA); therefore, they are not the primary reason for the inseparability of featural and configural processing in dyslexia found here. There is evidence that face recognition deficits in developmental prosopagnosia might be linked to a weakness in holistic processing of faces (Avidan et al., 2011; DeGutis et al., 2014; Palermo et al., 2011; Towler et al., 2018). In future studies we aim to perform RSA on featural and configural processing in developmental prosopagnosia comparing this with dyslexic readers, especially given that RSA can uncover differences that might otherwise go undetected in more traditional univariate analyses. More generally, our RSA results show that relying solely on mean accuracy to look for group differences can be deceptive as group means could be equal while data patterns could differ, hinting at different representations and different underlying mechanisms. Therefore, RSA implementations can be used effectively to explore variations in the visual representations of dyslexic and typical readers as well as other comparisons of group differences, such as other developmental disorders.

Conclusions

We draw two main conclusions here. Firstly, our results suggest that some dyslexic readers show object recognition impairments. This argues that problems in dyslexia are not restricted to reading which suggests that dyslexia— and visual word processing more generally — is not domain-specific. Such visual processing deficits cannot easily be attributed to phonological problems, consistent with the idea that dyslexia is a heterogeneous disorder. One alternative could be a high-level visual deficit in some dyslexic readers which could manifest as difficulties in reading. The direction of causality nonetheless needs further study.

Secondly, our RSA results suggest that dyslexic readers rely on a single process to identify visual objects. This demonstrates the effectiveness of representational similarity analysis (RSA) in behavioral studies. While univariate analyses failed to uncover process differences between dyslexic and typical readers, RSA revealed that dyslexic readers depend on only a single visual process regardless of whether features or configurations are task-relevant. This process effect was general, occurring for both faces and houses. Our results suggest that dyslexic readers' general failure to use different processes may be responsible for their reading problems and that for efficient reading, both featural and configural processing are required.

Acknowledgements

We would like to thank Jane E. Joseph for providing us with visual stimuli which were a part of a larger set developed by Collins et al. (2012). We are grateful to Guðrún Rakel Eiríksdóttir for her help with task coding, Berglind Anna Víðisdóttir and Sigurður Þór Þórðarson for their help with recruitment/running participants, and Randi Starrfelt for her feedback on this project. This work was supported by The Icelandic Research Fund (Grant No. 174013-051) and the University of Iceland Research Fund. Part of this paper was presented as poster at The European Conference on Visual Perception (ECVP) 2019, Leuven, Belgium.

Open Practices Statement

A preprint of this paper is available at https://psyarxiv.com/u6vqb. Data can be made available to other researchers upon request provided that the National Bioethics Committee of Iceland grants them permission for such access and provided that such access adheres to all Icelandic laws regarding data privacy and protection.

References

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. Neuropsychologia, 49(9), 2541-2552.

Behrmann, M., & Plaut, D. C. (2013). Distributed circuits, not circumscribed centers, mediate visual recognition. *Trends in cognitive sciences, 17*(5), 210-219.

Behrmann, M., & Plaut, D. C. (2019). Hemispheric Organization for Visual Object Recognition: A Theoretical Account and Empirical Evidence. *Perception*, 0301006619899049.

Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., & Buchanan, L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage, 33*(2), 739-748.

Bishop, D. V. (1997). Cognitive neuropsychology and developmental disorders: Uncomfortable bedfellows. The Quarterly Journal of Experimental Psychology Section A, 50(4), 899-923.

Bjornsdottir, G., Halldorsson, J. G., Steinberg, S., Hansdottir, I., Kristjansson, K., Stefansson, H., & Stefansson, K. (2014). The adult reading history questionnaire (ARHQ) in Icelandic: psychometric properties and factor structure. *Journal of Learning Disabilities, 47*(6), 532-542.

Brachacki, G. W., Fawcett, A. J., & Nicolson, R. I. (1994). Adults with dyslexia have a deficit in voice recognition. *Perceptual and motor skills, 78*(1), 304-306.

Brachacki, G. W., Nicolson, R. I., & Fawcett, A. J. (1995). Impaired recognition of traffic signs in adults with dyslexia. *Journal of Learning Disabilities, 28*(5), 297-301.

Brady, N., Darmody, K., Newell, F., & Cooney, S. M. (2020). Holistic processing of words and faces in dyslexia.

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: the power of an expertise framework. *Trends in cognitive sciences, 10*(4), 159-166.

Cao, X., Yang, Q., Zhong, P., & Chen, C. (2019). The characteristics of face configural effect in illiterates and literates. *Acta psychologica, 201*, 102951.

Chao, L. L., Weisberg, J., & Martin, A. (2002). Experience-dependent modulation of category-related cortical activity. Cerebral Cortex, 12(5), 545-551.

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain, 125*(5), 1054-1069.

Collins, E., Dundas, E., Gabay, Y., Plaut, D. C., & Behrmann, M. (2017). Hemispheric organization in disorders of development. *Visual cognition, 25*(4-6), 416-429.

Collins, Zhu, X., Bhatt, R. S., Clark, J. D., & Joseph, J. E. (2012). Process and domain specificity in regions engaged for face processing: an fMRI study of perceptual differentiation. *Journal of cognitive neuroscience, 24*(12), 2428-2444.

Conway, A., Brady, N., & Misra, K. (2017). Holistic word processing in dyslexia. *PloS one, 12*(11), e0187326.

Cox, D. D. (2014). Do we understand high-level vision? *Current opinion in neurobiology, 25*, 187-193.

Curby, K. M., Hayward, W. G., & Gauthier, I. (2004). Laterality effects in the recognition of depth-rotated novel objects. *Cognitive, Affective, & Behavioral Neuroscience, 4*(1), 100-111.

DeGutis, J., Cohan, S., & Nakayama, K. (2014). Holistic face training enhances face processing in developmental prosopagnosia. Brain, 137(6), 1781-1798.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384-398.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in cognitive sciences, 15*(6), 254-262.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *science, 330*(6009), 1359-1364.

De Martino, S., Espesser, R., Rey, V., & Habib, M. (2001). The "temporal processing deficit" hypothesis in dyslexia: New experimental evidence. *Brain and cognition, 46*(1-2), 104-108.

Dennett, H. W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., & Duchaine, B. (2012). The Cambridge Car Memory Test: A task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. Behavior research methods, 44(2), 587-605.

D'Souza, D., & Karmiloff-Smith, A. (2011). When modularization fails to occur: A developmental perspective. Cognitive neuropsychology, 28(3-4), 276-287.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is" special" about face perception? *Psychological review, 105*(3), 482.

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic bulletin & review, 2*(4), 460-493.

Franceschini, S., Bertoni, S., Gianesini, T., Gori, S., & Facoetti, A. (2017). A different vision of dyslexia: Local precedence on global perception. Scientific Reports, 7(1), 1-10.

Franceschini, S., Bertoni, S., Puccio, G., Mancarella, M., Gori, S., & Facoetti, A. (2021). Local perception impairs the lexical reading route. Psychological research, 85(4), 1748-1756.

Gabay, Y., Dundas, E., Plaut, D., & Behrmann, M. (2017). Atypical perceptual processing of faces in developmental dyslexia. *Brain and language, 173*, 41-51.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition, 103*(2), 322-330.

Gauthier, I., McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Van Gulick, A. E. (2014). Experience moderates overlap between object and face recognition, suggesting a common ability. *Journal of Vision, 14*(8), 7-7.

Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. *Current opinion in neurobiology, 11*(2), 219-224.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature neuroscience, 3*(2), 191-197.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance, 28*(2), 431.

Germanò, E., Gagliano, A., & Curatolo, P. (2010). Comorbidity of ADHD and dyslexia. *Developmental neuropsychology, 35*(5), 475-493.

Giofrè, D., Toffalini, E., Provazza, S., Calcagnì, A., Altoè, G., & Roberts, D. J. (2019). Are children with developmental dyslexia all the same? A cluster analysis with more than 300 cases. *Dyslexia, 25*(3), 284-295.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in cognitive sciences, 15*(1), 3-10.

Goswami, U. (2015). Sensory theories of developmental dyslexia: three challenges for research. *Nature Reviews Neuroscience, 16*(1), 43-54.

Grainger, J., Lété, B., Bertand, D., Dufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition, 123*(2), 280-292.

Grainger, J., & Ziegler, J. (2011). A dual-route approach to orthographic processing. *Frontiers in psychology, 2*, 54.

Griffin, J. W., Bauer, R., & Scherf, K. S. (2021). A quantitative meta-analysis of face recognition deficits in autism: 40 years of research. Psychological bulletin, 147(3), 268

Hervais-Adelman, A., Kumar, U., Mishra, R. K., Tripathi, V. N., Guleria, A., Singh, J. P., Eisner, F., & Huettig, F. (2019). Learning to read recycles visual cortical networks without destruction. Science advances, 5(9), eaax0262.

Hills, C. S., Pancaroglu, R., Duchaine, B., & Barton, J. J. (2015). Word and text processing in acquired prosopagnosia. *Annals of neurology, 78*(2), 258-271.

Hinshelwood, J. (1896). A CASE OF DYSLEXIA: A PECULIAR FORM OF WORD-BLINDNESS. 1. *The Lancet, 148*(3821), 1451-1454.

Holmes, D. R., & McKeever, W. F. (1979). Material specific serial memory deficit in adolescent dyslexics. *Cortex, 15*(1), 51-62.

Hsiao, J. H., & Cottrell, G. W. (2009). Not all visual expertise is holistic, but it may be leftist: The case of Chinese character recognition. *Psychological science, 20*(4), 455-463.

Huestegge, L., Rohrßen, J., van Ermingen-Marbach, M., Pape-Neumann, J., & Heim, S. (2014). Devil in the details? Developmental dyslexia and visual long-term memory for details. *Frontiers in psychology, 5*, 686.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature neuroscience, 3*(8), 759.

Kalanthroff, E., Naparstek, S., & Henik, A. (2013). Spatial processing in adults with attention deficit hyperactivity disorder. Neuropsychology, 27(5), 546.

Koldewyn, K., Jiang, Y. V., Weigelt, S., & Kanwisher, N. (2013). Global/local processing in autism: Not a disability, but a disinclination. Journal of autism and developmental disorders, 43(10), 2329-2340.

Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipitotemporal cortex. Neuron, 74(6), 1114-1124.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience, 2*, 4.

Lefly, D. L., & Pennington, B. F. (2000). Reliability and validity of the adult reading history questionnaire. *Journal of Learning Disabilities, 33*(3), 286-296.

Lieder, I., Adam, V., Frenkel, O., Jaffe-Dax, S., Sahani, M., & Ahissar, M. (2019). Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia. *Nature neuroscience, 22*(2), 256-264.

Lobmaier, J. S., Bölte, J., Mast, F. W., & Dobel, C. (2010). Configural and featural processing in humans with congenital prosopagnosia. *Advances in Cognitive Psychology, 6*, 23.

Magnússon, P., Smári, J., Sigurðardóttir, D., Baldursson, G., Sigmundsson, J., Kristjánsson, K., Sigurðardóttir, S., Hreiðarsson, S., Sigurbjörnsdóttir, S., & Guðmundsson, Ó. Ó. (2006). Validity of self-report and informant rating scales of adult ADHD symptoms in comparison with a semistructured diagnostic interview. *Journal of Attention disorders, 9*(3), 494-503.

Marsolek, C. J. (1999). Dissociable neural subsystems underlie abstract and specific object recognition. *Psychological science, 10*(2), 111-118.

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision, 5*(1), 6-6.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in cognitive sciences, 6*(6), 255-260.

McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Gauthier, I. (2012). The Vanderbilt Expertise Test reveals domain-general and domain-specific sex effects in object recognition. Vision research, 69, 10-22.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in cognitive sciences, 11*(1), 8-15.

McKone, E., & Robbins, R. (2007). The evidence rejects the expertise hypothesis: Reply to Gauthier & Bukach. *Cognition, 103*(2), 331-336. Morgan, W. P. (1896). A case of congenital word blindness. *British medical journal, 2*(1871), 1378.

Moeller, M. P., Tomblin, J. B., Yoshinaga-Itano, C., Connor, C. M., & Jerger, S. (2007). Current state of knowledge: Language and literacy of children with hearing impairment. Ear and hearing, 28(6), 740-753.

Nasr, S., Echavarria, C. E., & Tootell, R. B. (2014). Thinking outside the box: rectilinear shapes selectively activate scene-selective cortex. Journal of neuroscience, 34(20), 6721-6735.

Norton, E. S., Beach, S. D., & Gabrieli, J. D. (2015). Neurobiology of dyslexia. *Current opinion in neurobiology, 30*, 73-78.

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. Neuropsychologia, 49(5), 1226-1235.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods, 51*(1), 195-203.

Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature, 423*(6941), 752.

Plaut, D. C., & Behrmann, M. (2011). Complementary neural representations for faces and words: A computational exploration. *Cognitive neuropsychology, 28*(3-4), 251-275.

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage, 19*(3), 473-481.

Rakover, S. S. (2002). Featural vs. configurational information in faces: A conceptual and empirical analysis. *British journal of psychology, 93*(1), 1-30.

Reid, A. A. (2018). Neuroimaging reveals heterogeneous neural correlates of reading deficit in individuals with dyslexia consistent with a multiple deficit model. In *Neuroimaging-Structure, Function and Mind*. IntechOpen.

Rezlescu, C., Susilo, T., Wilmer, J. B., & Caramazza, A. (2017). The inversion, part-whole, and composite effects reflect distinct perceptual mechanisms with varied relationships to face recognition. *Journal of Experimental Psychology: Human Perception and Performance, 43*(12), 1961.

Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of cognitive neuroscience, 16*(2), 189-203.

Rice, G. E., Kerry, S. J., Robotham, R. J., Leff, A., Ralph, M. A. L., & Starrfelt, R. (2020). Revealing the graded spectrum of visual perceptual function following posterior cerebral artery stroke.

Richlan, F., Kronbichler, M., & Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *Neuroimage, 56*(3), 1735-1742.

Richler, J. J., Floyd, R. J., & Gauthier, I. (2014). The Vanderbilt Holistic Face Processing Test: A short and reliable measure of holistic face processing. *Journal of Vision, 14*(11), 10-10.

Richler, J. J., Wilmer, J. B., & Gauthier, I. (2017). General object recognition is specific: Evidence from novel and familiar objects. Cognition, 166, 42-55.

Roberts, D. J., Ralph, M. A. L., Kim, E., Tainturier, M.-J., Beeson, P. M., Rapcsak, S. Z., & Woollams, A. M. (2015). Processing deficits for familiar and novel faces in patients with left posterior fusiform lesions. *Cortex, 72*, 79-96.

Robotham, R. J., & Starrfelt, R. (2017). Face and word recognition can be selectively affected by brain injury or developmental disorders. *Frontiers in psychology, 8*, 1547.

Rossion, B., Dricot, L., Devolder, A., Bodart, J.-M., Crommelinck, M., Gelder, B. d., & Zoontjes, R. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of cognitive neuroscience, 12*(5), 793-802.

Rüsseler, J., Johannes, S., & Münte, T. F. (2003). Recognition memory for unfamiliar faces does not differ for adult normal and dyslexic readers: an event-related brain potential study. *Clinical neurophysiology, 114*(7), 1285-1291.

Shaywitz, S. E. (1998). Dyslexia. *New England Journal of Medicine, 338*(5), 307-312.

Sigurdardottir, H. M., Arnardottir, A., Halldorsdottir, E. T., Omarsdottir, H. R., & Valgeirsdottir, A. S. (2019). Faces and words are both associated and dissociated: Evidence from visual problems in dyslexia.

Sigurdardottir, H. M., Danielsdottir, H. B., Gudmundsdottir, M., Hjartarson, K. H., Thorarinsdottir, E. A., & Kristjánsson, Á. (2017). Problems with visual statistical learning in developmental dyslexia. *Scientific Reports, 7*(1), 1-12.

Sigurdardottir, H. M., Fridriksdottir, L. E., Gudjonsdottir, S., & Kristjánsson, Á. (2018). Specific problems in visual cognition of dyslexic readers: Face discrimination deficits predict dyslexia over and above discrimination of scrambled faces and novel objects. *Cognition, 175*, 157-168.

Sigurdardottir, H. M., Hjartarson, K. H., Gudmundsson, G. L., & Kristjánsson, Á. (2019). Own-race and other-race face recognition problems without visual expertise problems in dyslexic readers. *Vision research, 158*, 146-156.

Sigurdardottir, H. M., Ívarsson, E., Kristinsdóttir, K., & Kristjánsson, Á. (2015). Impaired recognition of faces and objects in dyslexia: Evidence for ventral stream dysfunction? *Neuropsychology, 29*(5), 739.

Smith-Spark, J. H., & Moore, V. (2009). The representation and processing of familiar faces in dyslexia: Differences in age of acquisition effects. *Dyslexia, 15*(2), 129-146.

Song, Y., & Hakoda, Y. (2012). The interference of local over global information processing in children with attention deficit hyperactivity disorder of the inattentive type. Brain and Development, 34(4), 308-317.

Song, Y., & Hakoda, Y. (2015). Lack of global precedence and global-to-local interference without local processing deficit: A robust finding in children with attention-deficit/hyperactivity disorder under different visual angles of the Navon task. Neuropsychology, 29(6), 888.

Starrfelt, R., & Gerlach, C. (2007). The visual what for area: words and pictures in the left fusiform gyrus. *Neuroimage, 35*(1), 334-342.

Towler, J., Fisher, K., & Eimer, M. (2018). Holistic face perception is impaired in developmental prosopagnosia. Cortex, 108, 112-126.

Tso, R. V.-y., Au, T. K.-f., & Hsiao, J. H.-w. (2014). Perceptual expertise: can sensorimotor experience change holistic processing and left-side bias? *Psychological science, 25*(9), 1757-1767.

Tso, R. V. Y., Chan, R. T. C., & Hsiao, J. H.-w. (2020). Holistic but with reduced right-hemisphere involvement: The case of dyslexia in Chinese character recognition. *Psychonomic bulletin & review*, 1-10.

Valdois, S., Bosse, M. L., & Tainturier, M. J. (2004). The cognitive deficits responsible for developmental dyslexia: Review of evidence for a selective visual attentional disorder. *Dyslexia, 10*(4), 339-363.

Van der Hallen, R., Evers, K., Brewaeys, K., Van den Noortgate, W., & Wagemans, J. (2015). Global processing takes time: A meta-analysis on local–global visual processing in ASD. Psychological bulletin, 141(3), 549.

Van Gulick, A. E., McGugin, R. W., & Gauthier, I. (2016). Measuring nonvisual knowledge about object categories: The semantic Vanderbilt expertise test. Behavior research methods, 48(3), 1178-1196.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of child psychology and psychiatry, 45*(1), 2-40.

Ventura, P. (2014). Let's face it: reading acquisition, face and word processing. Frontiers in psychology, 5, 787.

Ventura, P., Domingues, M., Ferreira, I., Madeira, M., Martins, A., Neto, M., & Pereira, M. (2019). Holistic Word Processing Is Involved in Fast Parallel Reading. *Journal of Expertise/March, 2*(1).

Ventura, P., Fernandes, T., Cohen, L., Morais, J., Kolinsky, R., & Dehaene, S. (2013). Literacy acquisition reduces the influence of automatic holistic processing of faces and houses. Neuroscience letters, 554, 105-109.

Vogel, A. C., Petersen, S. E., & Schlaggar, B. L. (2014). The VWFA: it's not just for words anymore. *Frontiers in human neuroscience, 8*, 88.

Wilson, A. J., Andrewes, S. G., Struthers, H., Rowe, V. M., Bogdanovic, R., & Waldie, K. E. (2015). Dyscalculia and dyslexia in adults: cognitive bases of comorbidity. *Learning and Individual Differences, 37*, 118-132.

Wong, A. C.-N., Bukach, C. M., Yuen, C., Yang, L., Leung, S., & Greenspon, E. (2011). Holistic processing of words modulated by reading experience. *PloS one, 6*(6).

Wong, A. C.-N., & Gauthier, I. (2007). An analysis of letter expertise in a levels-of-categorization framework. *Visual cognition, 15*(7), 854-879.

Wong, A. C.-N., Wong, Y. K., Lui, K. F., Ng, T. Y., & Ngan, V. S. (2019). Sensitivity to configural information and expertise in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 45*(1), 82.

Yovel, G., & Kanwisher, N. (2004). Face perception: domain specific, not process specific. *Neuron, 44*(5), 889-898.

Yue, X., Pourladian, I. S., Tootell, R. B., & Ungerleider, L. G. (2014). Curvature-processing network in macaque visual cortex. Proceedings of the National Academy of Sciences, 111(33), E3467-E3475.

Zhang, L., & Cottrell, G. W. (2005). Holistic processing develops because it is good. Proceedings of the cognitive science society,

Zhou, G., Cheng, Z., Zhang, X., & Wong, A. C.-N. (2012). Smaller holistic processing of faces associated with face drawing experience. *Psychonomic bulletin & review, 19*(2), 157-162.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological science, 21*(4), 551-559.

Supplementary Material

Object Recognition: Group Differences Broken Down by Difficulty Level

In the object recognition task, dyslexic readers tended to be less accurate than typical readers in featural and configural processing of houses. Figure 8 shows this broken down by difficulty level.

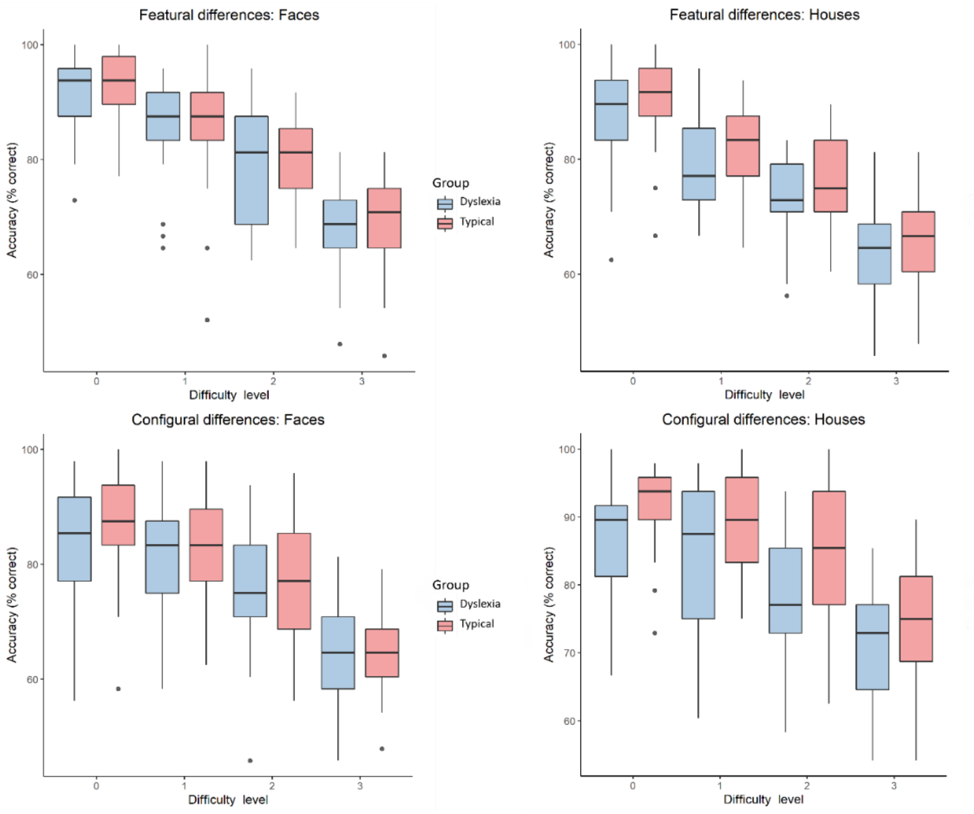


Figure 8. Featural and configural processing of faces and houses by difficulty levels in dyslexic and typical readers. Boxes display the accuracy range. Horizontal lines show median accuracy levels. Box limits indicate the 25th and 75th percentiles. Whiskers extending outward from the box denote variation in the upper and lower 25% of scores. Dots indicate extreme scores.

Object Recognition: Group Differences After Excluding Participants with Potential Comorbidities

To ensure that the results of the object recognition task cannot be attributed to ADHD, we excluded individuals with possible ADHD (who reported an ADHD diagnosis or where scores on Behavioral Evaluation Questionnaire for Adults I or II reached suggested cutoff points for either childhood ADHD symptoms or current ADHD symptoms) and their paired participants, and redid the univariate t-tests of the object recognition task. We redid the same analysis multiple times, each time excluding people who reported dyscalculia, autism, hearing problems, and language problems and their paired participants. Excluding people with any of these disorders only minimally affected the object recognition task outcome.

Repeated measures ANOVA with covariates

In a mixed ANOVA, we coded education as high (undergraduate or graduate degree) and low (no undergraduate or graduate degree), entering this as a factor in addition to stimulus, process, and group. People with low educational level performed less accurately than people with high educational levels (table 4, $p = .04$). As before, the interaction of stimulus and process was significant ($p = <.001$). Interestingly, there was also a significant interaction between education level, group, and stimulus ($p = .02$, figure 9). House recognition of low-education dyslexic readers was poorer than house recognition of typical readers with the same educational level, while house recognition of dyslexic and typical high-education readers was similar. Mean face accuracy was similar in all subgroups.

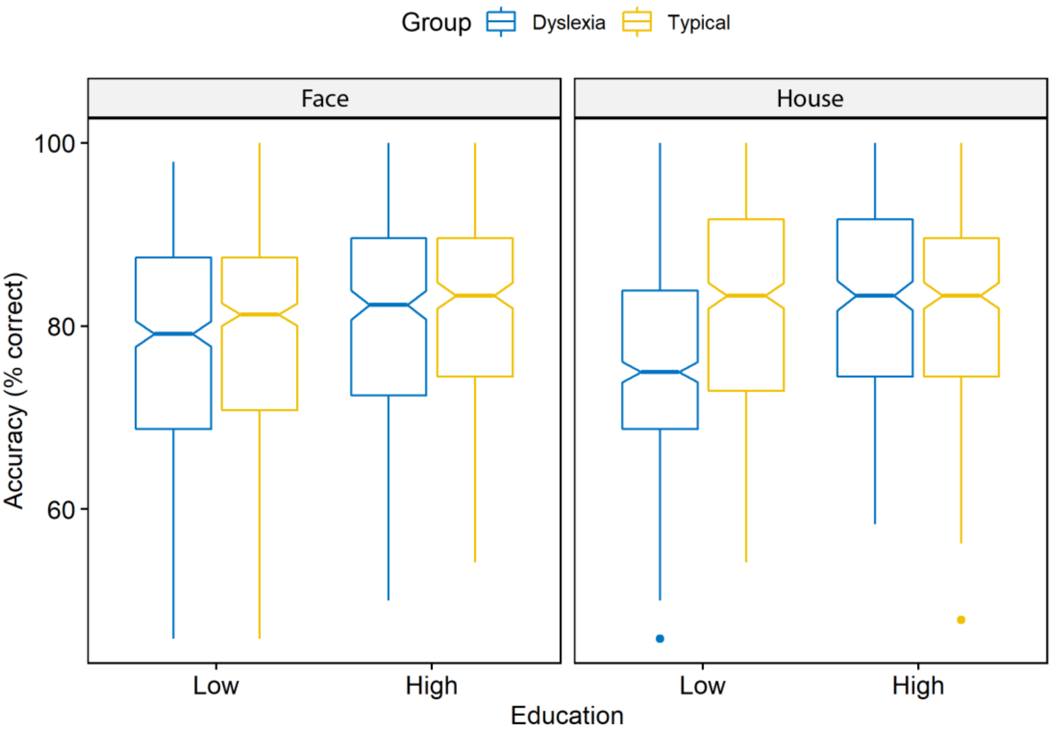


Figure 9. The interaction plot of education level (low and high education), group (dyslexic and typical readers), and stimulus (face or house). Boxes display the range of percentage correct. Horizontal lines show median percent correct. Horizontal lines show median percent correct. Box limits indicate the 25th and 75th percentiles. Whiskers extending outward from the box denote variation in the upper and lower 25% of scores. Dots indicate extreme scores.

To control for a potential role of speed-accuracy tradeoffs, ADHD, and other disorders, we created a new variable "other disorders", where participants who reported dyscalculia, hearing impairments, language problems and autism spectrum disorder were coded with "1" and others with "0". Next, we entered the average z-score of reaction times for the four types of trials (featural faces, featural houses, configural faces and configural houses; all four reaction times were correlated, all *r*s > 0.74), three measures of ADHD (reported ADHD diagnosis coded as "0" for no and "1" for yes, childhood ADHD symptoms, and current ADHD symptoms), and "other disorders" as covariates. The results were like the 2 x 2 x 2 repeated measures ANOVA without other disorders and reaction time except that the main effect of education was no longer significant, but the interaction between education level, stimulus, and group remained significant ($p = .02$), confirming that the result was unlikely to be related to comorbidities or speed-accuracy tradeoffs. In all analyses above, however, no group differences were found for featural vs. configural processing (table 4).

Table 4. Results of 4-way (educational level x group x stimulus x process) repeated measures ANOVAs with and without covariates (ADHD measures, other disorders (reported dyscalculia, hearing impairments, language problems, or autism spectrum disorders) and the average z-score of reaction times for the four types of trials). ^η_G^2 indicates generalized eta-squared.

Effect	With no covariates			With covariates		
	F	p	$\hat{\eta}_G^2$	F	p	$\hat{\eta}_G^2$
Education	4.56	.04	.05	0.06	.81	.001
Group	3.62	.07	.04	0.04	.83	<.001
Stimulus	0.59	.45	.002	0.59	.45	.002
Process	0.62	.44	.001	0.62	.44	.001
Education × Group	2.27	.14	.02	3.75	.06	.04
Education × Stimulus	<.001	.98	<.001	<.001	.98	<.001
Education × Process	0.24	.63	<.001	0.24	.63	<.001
Group × Stimulus	3.07	.09	.01	3.07	.09	.01
Group × Process	1.51	.23	.002	1.51	.23	.002
Stimulus × Process	160.99	<.001	.13	160.99	<.001	.12
Education × Group × Stimulus	5.76	.02	.02	5.76	.02	.02
Education × Group × Process	0.92	.35	.001	0.92	.35	.001
Education× Stimulus × Process	0.41	.53	<.001	0.41	.53	<.001
Group × Stimulus × Process	0.48	.49	.001	0.48	.49	.001
Education× Group × Stimulus × Process	0.06	.80	<.001	0.06	.80	<.001

Representational Similarity Analysis (RSA): Group Differences in The Separability Of Processes After Excluding Participants With Other Disorders

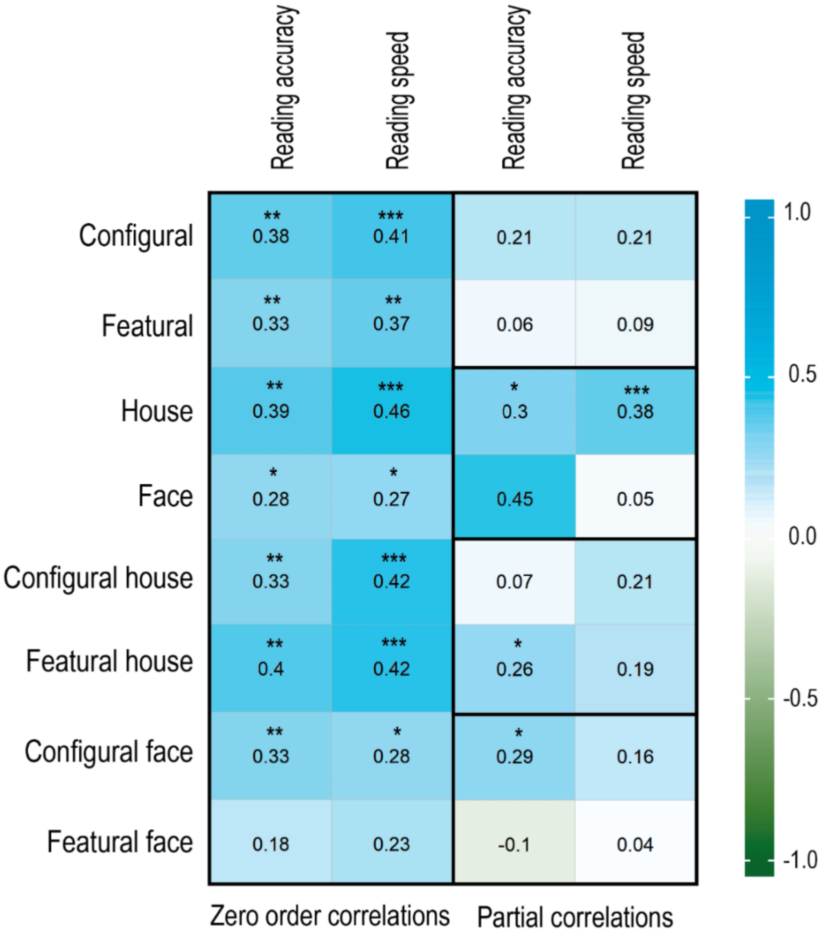
To ensure that group differences in the separability of processes could not be attributed to other disorders, we removed people with possible ADHD (who reported an ADHD diagnosis or whose screening on the Behavioral Evaluation Questionnaire for Adults I and II – see Results: Other Disorders section – reached ADHD cutoff points for either childhood ADHD symptoms or current ADHD symptoms), and their paired participants and redid the Representational Similarity Analysis (RSA, N = 12 remaining in each group). The outcome was comparable, the only exception being the significance of difficulty levels for dyslexic readers ($b = 0.08$, $t(116)= 2.41$, $p = .02$). We did the same for reported dyscalculia, autism, language problems, and hearing problems. The processes effect was minimally influenced by the absence of individuals with any of these disorders. The main RSA result is therefore unlikely to be due to comorbidity with ADHD or other disorders.

Representational Similarity Analysis (RSA): Possible Modulation by Educational Level

RSA was conducted across groups, with and without each individual participant, and the differences between unstandardized *b*s for process with and without the inclusion of that same participant were calculated to estimate his or her contribution to the total correlation coefficients (unstandardized *bs*). Consider participant x as an example; RSA was run across groups, once with participant x in the group and once after excluding participant x. Then the differences between the unstandardized *b* with and without participant x were computed for process (0.072 - 0.067 = 0.005). This was replicated for all participants so that all participants had an individual process value as an estimate for their contribution. Then a 2 x 2 mixed ANOVA was performed with these individual process values (featural and configural) as the dependent variable, educational level (low and high education) as a between-subjects factor and group (dyslexic and typical readers) as a paired (repeated measures) factor. The outcome revealed a significant main effect of group ($\hat{\eta}^2_G = .09$,$F(29) = 6.83$, $p = .01$). The main effect of education was, however, not significant ($\hat{\eta}^2_G = .00$, $F(29) = .09$, $p = .75$) and there was no significant interaction between group and education ($\hat{\eta}^2_G = .01$, $F(29) = .86$, $p = .36$). Educational level therefore does not contribute to group differences in the separability of processes.

Correlations of Object Recognition, Reading Speed And Reading Accuracy After Excluding Participants With Missing Reading Scores

Three participants (two typical readers and one dyslexic reader) who had missing values for reading speed and accuracy for the IS-FORM/IS-PSEUDO reading tasks were removed and the association between performance on the object recognition task (accuracy for the four subtasks, as well as total accuracy of faces, houses, featural and configural processing) and the two measures of reading performance (reading speed, reading accuracy) were assessed again. Excluding participants with missing values for reading speed and accuracy for the IS-FORM/IS-PSEUDO reading tests only minimally affected the results. The results are summarized in supplementary figure 10.



	Zero order correlations		Partial correlations	
	Reading accuracy	Reading speed	Reading accuracy	Reading speed
Configural	** 0.38	*** 0.41	0.21	0.21
Featural	** 0.33	** 0.37	0.06	0.09
House	** 0.39	*** 0.46	* 0.3	*** 0.38
Face	* 0.28	* 0.27	0.45	0.05
Configural house	** 0.33	*** 0.42	0.07	0.21
Featural house	** 0.4	*** 0.42	* 0.26	0.19
Configural face	** 0.33	* 0.28	* 0.29	0.16
Featural face	0.18	0.23	-0.1	0.04

Figure 10. Correlations between reading measures and accuracy on the object recognition task. The three leftmost panels show zero order correlations (Pearson's r) and the three rightmost panels show partial correlation coefficients. Each box on the right-hand side indicates mutually partialled-out variables. The upper row of the top-most box shows correlations with configural processing accuracy when featural accuracy is partialled out, while the lower row of the same box shows correlations with featural processing accuracy when configural accuracy is partialled out. The upper row of the second box from the top shows correlations with house accuracy when face accuracy is partialled out, while the lower row of the same box shows correlations with face accuracy when house accuracy is partialled out. The upper row of the third box from the top shows correlations with configural house accuracy when featural house accuracy is partialled out, while the lower row of the same box shows correlations with featural house accuracy when configural house accuracy is partialled out. The upper row of the bottom box shows correlations with configural face accuracy when featural face accuracy is partialled out, while the lower row of the same box shows correlations with featural face accuracy when configural face accuracy is partialled out. Asterisks indicate significance levels: *p < .05, **p < .01, *** p < .001.

Correlations of Object Recognition, Reading Speed and Reading Accuracy Divided by Group

Correlational analysis was originally done across groups. Dyslexic and typical readers however almost surely not only differ in terms of their reading experience, but also in other ways. Therefore, as suggested by a reviewer, we performed an exploratory analysis to investigate whether there is a difference between dyslexic and typical readers in correlations of object recognition with ARHQ, reading speed, and reading accuracy (figure 6). The results for dyslexic and typical readers are summarized in figure 11 and 12, respectively. Our result showed that for the zero-order correlations for dyslexic readers, the ARHQ was negatively associated with configural processing and configural house processing, reading speed was positively correlated with house processing, and reading accuracy was positively associated with featural house processing. Partial correlations of dyslexic readers were not significant. As for the zero-order correlations for typical readers, the speed and accuracy of reading was positively associated with overall face processing, configural face processing, and configural processing, and reading speed was positively correlated with featural processing. For the partial correlation analyses of typical readers, face and configural face were positively correlated with reading accuracy.

We should note that the lack of significance does not mean that these results differ significantly from the main analysis. For example, featural processing is significantly correlated with ARHQ in the main analysis but not for either group separately. However, their actual r values are close-to the same in all cases, but in the subgroups, we have effectively decreased our sample size by half, thus diminishing power, i.e., the correlations need to be bigger to be significant in a smaller sample. When we divided by group, we are truncating the range of all dependent variables, as there is less variability in ARHQ, reading speed, and reading accuracy within each group than between them. This is a-priori expected to diminish all correlations. Also, just because we find a significant correlation in one group and not the other, that does not necessarily mean that the two correlations are different from each other (they might be very similar, but one could be barely below significance and the other barely above). As a further exploratory analysis, we followed this up with a formal check of a group interaction. We compared linear models with and without an interaction effect for each category (e.g., face, house, featural, configural, featural face, featural house, configural face, configural house) by Hierarchical Linear Regression. The result showed that adding an interaction term never significantly improved any regression models (all Fs < 3.15, all ps > .08). Although we cannot make a strong claim based on our findings, it is worth investigating if the association between reading and face processing develops quite differently for people with and without reading problems.

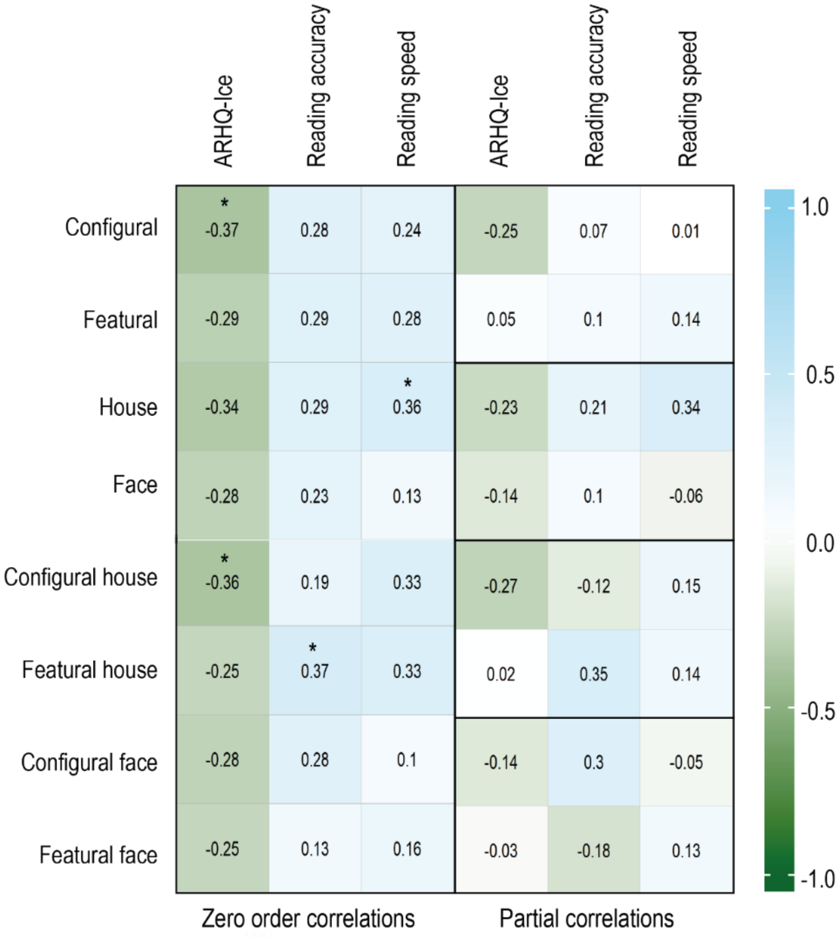
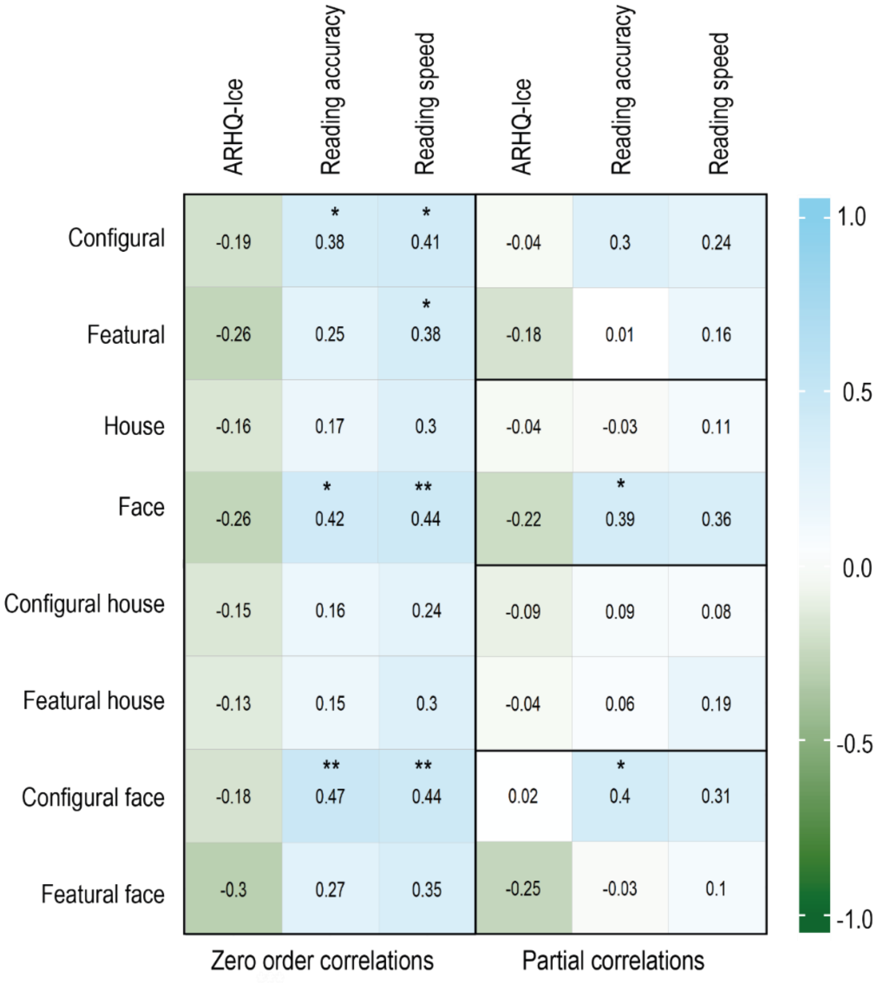


Figure 11. Correlations between reading measures and accuracy on the object recognition task in dyslexic readers. The three leftmost panels show zero order correlations (Pearson's r) and the three rightmost panels show partial correlation coefficients. Each box on the right-hand side indicates mutually partialled-out variables. The upper row of the top-most box shows correlations with configural processing accuracy when featural accuracy is partialled out, while the lower row of the same box shows correlations with featural processing accuracy when configural accuracy is partialled out. The upper row of the second box from the top shows correlations with house accuracy when face accuracy is partialled out, while the lower row of the same box shows correlations with face accuracy when house accuracy is partialled out. The upper row of the third box from the top shows correlations with configural house accuracy when featural house accuracy is partialled out, while the lower row of the same box shows correlations with featural house accuracy when configural house accuracy is partialled out. The upper row of the bottom box shows correlations with configural face accuracy when featural face accuracy is partialled out, while the lower row of the same box shows correlations with featural face accuracy when configural face accuracy is partialled out. Asterisks indicate significance levels: *p < .05, **p < .01, *** p < .001.



	Zero order correlations			Partial correlations		
	ARHQ-Ice	Reading accuracy	Reading speed	ARHQ-Ice	Reading accuracy	Reading speed
Configural	-0.19	* 0.38	* 0.41	-0.04	0.3	0.24
Featural	-0.26	0.25	* 0.38	-0.18	0.01	0.16
House	-0.16	0.17	0.3	-0.04	-0.03	0.11
Face	-0.26	* 0.42	** 0.44	-0.22	* 0.39	0.36
Configural house	-0.15	0.16	0.24	-0.09	0.09	0.08
Featural house	-0.13	0.15	0.3	-0.04	0.06	0.19
Configural face	-0.18	** 0.47	** 0.44	0.02	* 0.4	0.31
Featural face	-0.3	0.27	0.35	-0.25	-0.03	0.1

Figure 12. Correlations between reading measures and accuracy on the object recognition task in typical readers. The three leftmost panels show zero order correlations (Pearson's r) and the three rightmost panels show partial correlation (Pearson's r) and the three rightmost panels show partial correlation coefficients. Each box on the right-hand side indicates mutually partialled-out variables. The upper row of the top-most box shows correlations with configural processing accuracy when featural accuracy is partialled out, while the lower row of the same box shows correlations with featural processing accuracy when configural accuracy is partialled out. The upper row of the second box from the top shows correlations with house accuracy when face accuracy is partialled out, while the lower row of the same box shows correlations with face accuracy when house accuracy is partialled out. The upper row of the third box from the top shows correlations with configural house accuracy when featural house accuracy is partialled out, while the lower row of the same box shows correlations with featural house accuracy when configural house accuracy is partialled out. The upper row of the bottom box shows correlations with configural face accuracy when featural face accuracy is partialled out, while the lower row of the same box shows correlations with featural face accuracy when configural face accuracy is partialled out. Asterisks indicate significance levels: *p < .05, **p < .01, *** p < .001.

Lexical Decision: Word Length and Frequency

The lexical decision test was developed to measure the word length effect in Icelandic (Sigurdardottir et al., 2019), or the correlation between visual word recognition time and the number of letters in the word (Barton et al., 2014). An elevated word length effect is considered a measure of a letter-by-letter reading strategy which is the hallmark of pure alexia (Barton et al., 2014). The word length effect decreases as children's reading abilities improve (Zoccolotti et al., 2005). Dyslexic readers show an elevated word length effect, probably reflecting a letter-by-letter reading strategy (Juphard et al., 2004; Martens & de Jong, 2006; Ziegler et al., 2003). An increased word length effect has also been found following damage to ventral regions which are associated with visual recognition problems (Barton et al., 2014; Behrmann et al., 1998). Therefore, word length effects may also be correlated with visual recognition issues arising from ventral stream impairment in dyslexia.

In the lexical decision test, the stimuli are 4, 5 or 6 letter real words and pseudowords. Twenty real words (10 high-frequency and 10 low-frequency words) and 20 pseudowords of each length were used. Participants completed two practice trials followed by 120 experimental trials with two breaks. On each trial, a word or pseudoword was displayed (Arial font, black letters, approximately 1° height). Participants indicated with a keypress whether the stimulus shown was a real word or a pseudoword. In one version of the test, Z represented words and M represented pseudowords (marked with yellow stickers), while in the other, the key mapping was reversed (counterbalanced by participant pair). Following an incorrect response, the (pseudo)word turned red for 500 ms with a sound. During each 500 ms inter-trial interval, a fixation point (height approximately 1°) was shown at screen center. Reaction time and accuracy of (pseudo)words with different number of letters were measured.

The word length effect was estimated for each individual participant in the lexical decision task by calculating a response time slope (milliseconds per letter) or the expected increase in response times for each letter added to a word or pseudoword. Dyslexic readers tended to have a larger word length effect than typical readers (dyslexic readers $M = 54.69$, $SD = 63.93$; typical readers $M = 15.60$, $SD = 36.05$; $t(30) = -2.75$, $p = .01$, $D_{av} = 0.78$, 95% CI [-68.08, -10.09]). Additional analysis showed that dyslexic readers tended to be slower (figure 13) and less accurate (figure 14) than typical readers for both high and low frequency words but the group difference was especially large for pseudowords. Pseudowords furthermore showed a greater word length effect for dyslexic readers compared to typical readers. The main reason for calculating this word length effect was to see whether it could statistically account for any visual processing problems on the object recognition task. No correlations between the word length effect slope and the four conditions of the object recognition task were significant (all *ps* > .43, all absolute *rs* > .06).

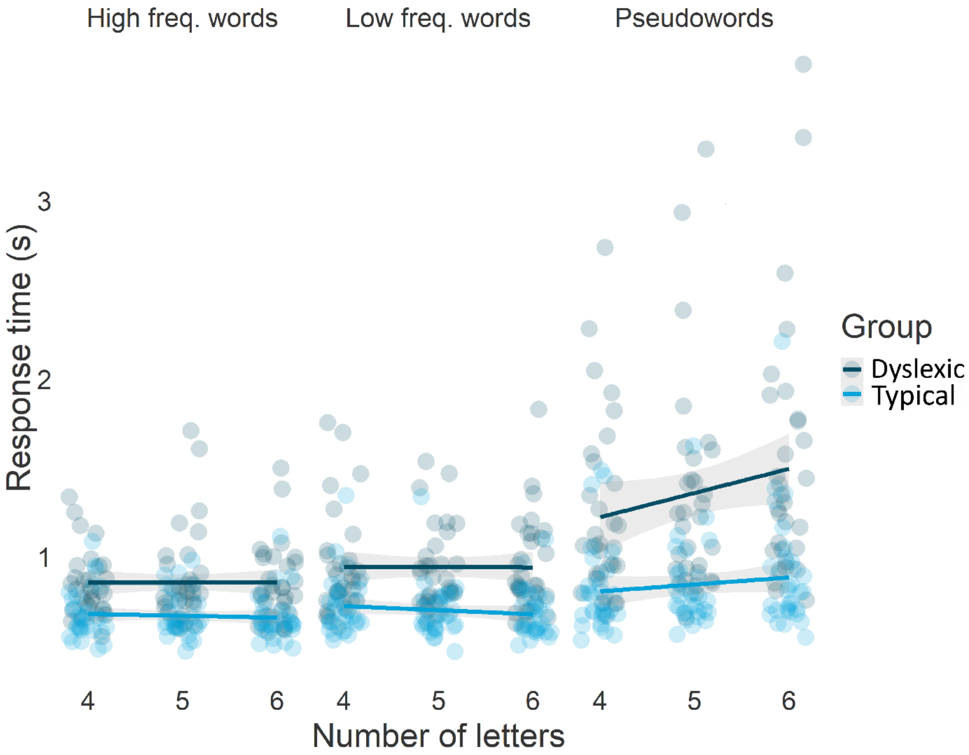


Figure 13. Mean response times (RTs) for lexical decision by the number of letters in a (pseudo)word, frequency of word (high, low, none), and group membership. Only correct trials were included, and trials over three standard deviations from the RT mean of each participant were additionally thrown out. Each dot shows a participant's mean RT for a particular combination of letter number and word frequency. Dots are jittered to minimize overlap. Lines show linear fits for each group. Shaded areas indicate 95% confidence intervals.

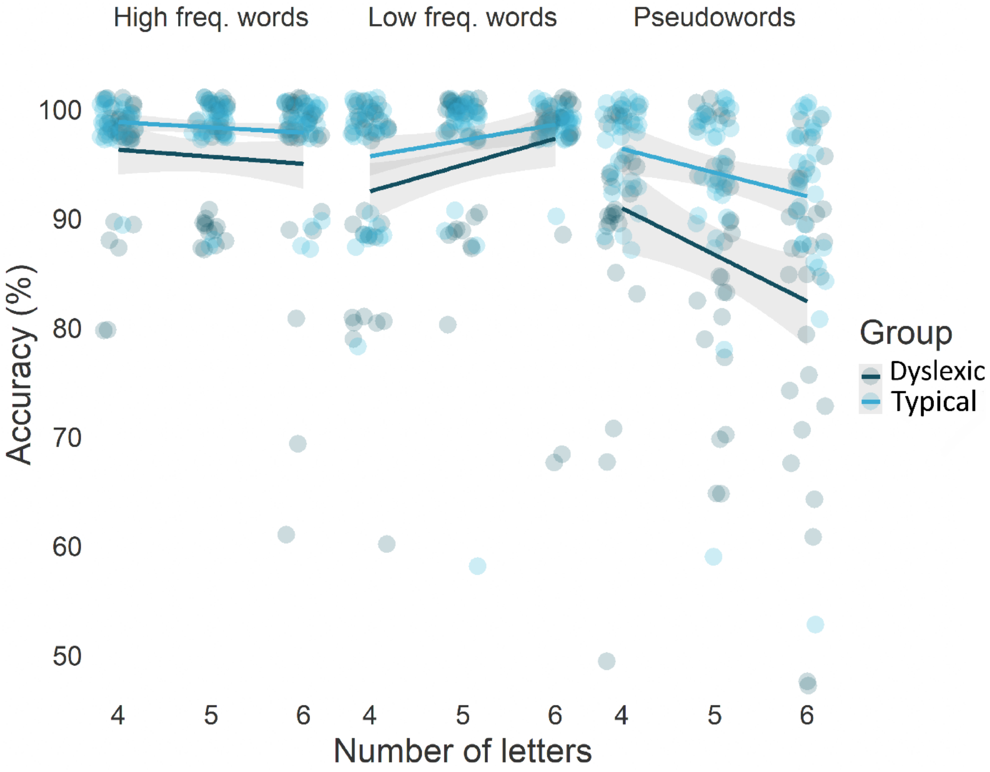


Figure 14. Accuracy (percent correct) for lexical decision by the number of letters in a (pseudo)word, frequency of word (high, low, none), and group membership. All trials were included. Each dot shows a participant's mean accuracy for a particular combination of letter number and word frequency. Dots are jittered to minimize overlap. Lines show linear fits for each group. Shaded areas indicate 95% confidence intervals.

Reading Speed and Reading Accuracy After Exclusion of Misclassified Participants

After excluding misclassified participants, almost all participants in the dyslexia group read less accurately and more slowly (read fewer words or pseudowords per minute) than their matched typical readers, and the group effect size is very large (figure 15)

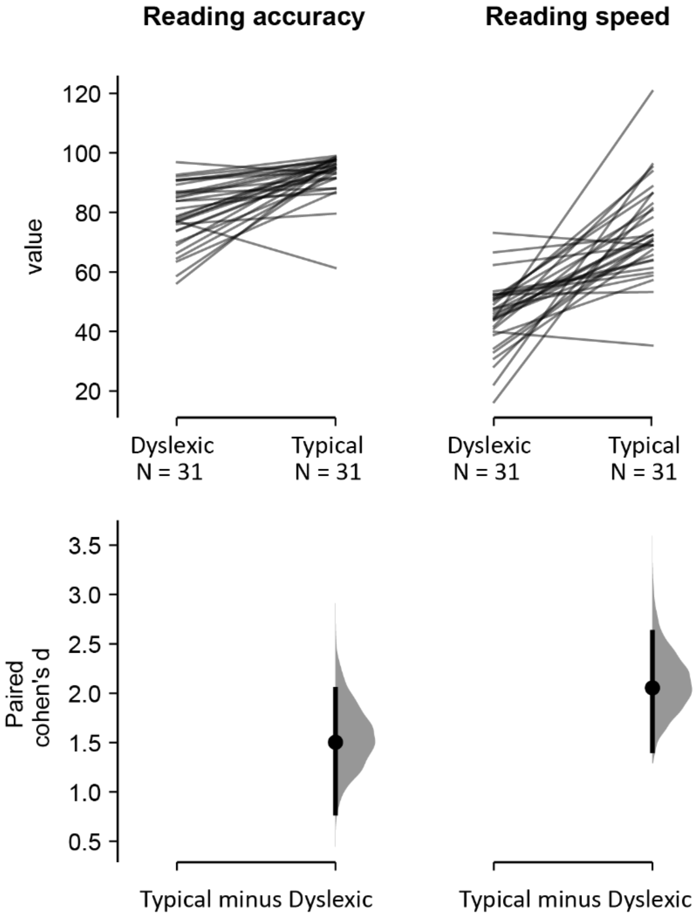


Figure 15. Cumming estimation plot for paired Dav for reading speed and reading accuracy after exclusion of misclassified participants. The raw data is plotted on the upper panel for each paired set of observers connected by a line. In the lower panel, each paire mean difference is plotted as a bootstrap sampling distribution (5000 bootstrap samples were taken). Mean differences are depicted as dots; 95% confidence intervals are indicated by the ends of the vertical error bars, and they are bias-corrected and accelerated.

Supplementary References

Barton, J. J., Hanif, H. M., Eklinder Björnström, L., & Hills, C. (2014). The word-length effect in reading: A review. *Cognitive neuropsychology, 31*(5-6), 378-412.

Behrmann, M., Nelson, J., & Sekuler, E. (1998). Visual complexity in letter-by-letter reading: Pure alexia is not pure. *Neuropsychologia, 36*(11), 1115-1132.

Juphard, A., Carbonnel, S., & Valdois, S. (2004). Length effect in reading and lexical decision: Evidence from skilled readers and a developmental dyslexic participant. *Brain and cognition, 55*(2), 332-340.

Martens, V. E., & de Jong, P. F. (2006). The effect of word length on lexical decision in dyslexic and normal reading children. *Brain and language, 98*(2), 140-149.

Sigurdardottir, H. M., Arnardottir, A., Halldorsdottir, E. T., Omarsdottir, H. R., & Valgeirsdottir, A. S. (2019, July 16). Faces and words are both associated and dissociated: Evidence from visual problems in dyslexia. https://doi.org/10.31234/osf.io/n2kp6

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of experimental child psychology, 86*(3), 169-193.

Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain and language, 93*(3), 369-373.

Paper II

Using representational similarity analysis to reveal category and process specificity in visual object recognition

Bahareh Jozranjbar[1] (baj18@hi.is), Árni Kristjánsson[1], Randi Starrfelt[2],
Christian Gerlach[3], Heida Maria Sigurdardottir[1]

[1] Icelandic Vision Lab
Department of Psychology
University of Iceland

[2] Department of Psychology
University of Copenhagen
Copenhagen, Denmark

[3] Department of Psychology
University of Southern Denmark
Odense, Denmark

Abstract

Cross-condition comparisons on neurodevelopmental conditions are central in neurodiversity research. In the realm of visual perception, the performance of participants with different category-specific disorders such as developmental prosopagnosia (problems with faces) and dyslexia (problems with words) have contributed to understanding of perceptual processes involved in word and face recognition. Alterations in face and word recognition are present in several neurodiverse populations, and improved knowledge about their relationship may increase our understanding of this variability of impairment. The present study investigates organizing principles of visual object processing and their implications for developmental disorders of recognition. Some accounts suggest that distinct mechanisms are responsible for recognizing objects of different categories, while others propose that categories share or even compete for cortical resources. We took an individual differences approach to estimate the relationship between abilities in recognition. Neurotypical participants (N = 97 after outlier exclusion) performed a match-to-sample task with faces, houses, and pseudowords. Either individual features or feature configurations were manipulated. To estimate the separability of visual recognition mechanisms, we used representational similarity analysis (RSA) where correlational matrices for accuracy were compared to predicted data patterns. Recognition abilities separated into face recognition on one hand and house/pseudoword recognition on the other, indicating that face recognition may rely on relatively selective mechanisms in neurotypicals. We also found evidence for a general visual object recognition mechanism, while some combinations of category (faces, houses, words) and processing type (featural, configural) likely rely on additional mechanisms. Developmental conditions may therefore reflect combinations of impaired and intact aspects of specific and general visual object recognition mechanisms, where featural and configural processes for one object category separate from the featural or configural processing of another. More generally, RSA is a promising approach for advancing understanding of neurodiversity, including shared aspects and distinctions between neurodevelopmental conditions of visual recognition.

Keywords:

Face recognition, Object recognition, Word recognition, Visual Recognition, Reading

Introduction

Face and word processing have been at the center of the domain generality/specificity debate, and findings from neurodevelopmental conditions like dyslexia and prosopagnosia have provided important observations that have increased our understanding of functional interdependency and segregation. In the following, we will go through some of this evidence, while keeping in mind that interpreting it in the context of typical development of visual processing may not be straightforward.

Developmental dyslexia is a reading disorder that occurs despite normal intellectual capacity, adequate educational opportunities, and intact sensory abilities (Kristjansson & Sigurdardottir, 2022; Shaywitz, 1998). In comparison, developmental prosopagnosia is characterized by severe face recognition problems (Duchaine & Nakayama, 2006). Both disorders occur in the absence of brain injury (Duchaine & Nakayama, 2006; Shaywitz, 1998). In principle, recognition problems in dyslexia or prosopagnosia could be limited to words and faces, respectively, consistent with potential domain-specificity of these categories (e.g., Dehaene & Cohen, 2011; Kanwisher, 2000; Kleinschmidt & Cohen, 2006; Rhodes et al., 2004; Yovel & Kanwisher, 2004). However, they could also extend to other visual categories in accordance with domain-general accounts (Gauthier et al., 2014; Hills et al., 2015; Sigurdardottir et al., 2015, 2018, 2019). In fact, these two possibilities need not even apply identically to faces and words, and given the variability of neurodevelopmental trajectories, it is likely that some people might show selective impairment with one category of objects, while others will show more general deficits (Gerlach et al., 2022). For example, word recognition might be more domain-general while faces could involve content-specific mechanisms (B. Duchaine & Nakayama, 2005; Kanwisher, 2000; Rhodes et al., 2004; Yovel & Kanwisher, 2004).

Visual object recognition may rely on both dissociable and shared mechanisms for different object categories or for different aspects of objects such as their features versus their configurations of features. According to domain-specific accounts, some object categories like words and faces are processed by largely independent mechanisms (Kanwisher et al., 1997; Petersen et al., 1988). Other accounts suggest that face and word recognition share or even compete for the same cortical resources (Behrmann & Plaut, 2013, 2015; Dehaene et al., 2010; Dundas et al., 2013) depending on the type of visual processing that the task requires (Sigurdardottir et al., 2021). Dehaene et al. (2010) proposed that words and faces compete for limited cortical resources during reading acquisition; hence, better performance with words could come at a cost – i.e., worse performance with faces. Alternatively, training on orthographic stimuli during reading development could fine-tune abilities such as face perception and may increase the amount of shared representations between words and faces (Hervais-Adelman et al., 2019). Some studies support a domain-general object identification ability that is separate from general intelligence and other cognitive and personality variables and reveals connections across various visual tasks and categories (Hendel et al., 2019; Richler et al., 2019). Richler et al. (2019) found that such a general factor, referred to as *o*, could explain an average of 89% of the variance in performance with novel object categories. Hendel et al., (2019) also found that super-recognizers, who have superior face recognition abilities, performed better than a control group on all visual categories (words, objects, and faces). Consequently, what causes superior face processing in super-recognizers may also explain their superior performance with other visual categories and reflect a general factor in the visual domain, referred to by Hendel and colleagues as factor *VG* (Hendel et al., 2019). Similarly, Maratos et al., (2022) found that known stimuli clustered together regardless of whether they were words or objects, interpreting this as evidence for a common pattern recognition mechanism for familiar objects.

Some studies have reported impaired face recognition in dyslexic readers (Collins et al., 2017; Gabay et al., 2017; Sigurdardottir et al., 2015, 2018, 2019), while others have not (Holmes & McKeever, 1979; Jozranjbar et al., 2021; Rüsseler et al., 2003; Smith-Spark & Moore, 2009). Kühn et al., (2020) reported that some individuals with dyslexia have difficulty with face recognition while others do not. This is perhaps not surprising given that dyslexia is a heterogeneous disorder (Kristjansson & Sigurdardottir, 2022). Also, while dyslexic readers have been shown to have visual recognition problems (Brachacki et al., 1995; Huestegge et al., 2014; Jozranjbar et al., 2021; Mayseless & Breznitz, 2011; Sigurdardottir et al., 2015), the visual processing deficits of dyslexic readers may not generalize to all object types (Gabay et al., 2017; Sigurdardottir et al., 2018). The specificity of word recognition problems in dyslexia and how they relate to face and object recognition problems is therefore still unclear.

A review on general visual recognition problems in developmental prosopagnosia revealed an association between impaired face and object recognition (Geskin & Behrmann, 2018). Consistently, Richler et al., (2019) point out that when several categories are compared, face recognition does not stand out as a particularly distinct ability. While developmental prosopagnosia may frequently co-occur with developmental object agnosia, a deficit in visual object recognition, due to similar risk factors, co-occurrence may not be inevitable; some individuals may have a selective difficulty with faces (Gray & Cook, 2018). Most commonly, however, visual object recognition is affected in developmental prosopagnosia, although not to the same degree as face recognition (Gerlach et al., 2016; Behrmann & Geskin, 2018). Developmental prosopagnosia studies have, however, found little evidence for word recognition disadvantages in this group (Burns et al., 2017; Collins et al., 2017; Robotham & Starrfelt, 2017; Rubino et al., 2016; Starrfelt et al., 2018), consistent with separable developmental trajectories for mechanisms supporting visual face and word processing. But as Geskin and Behrmann, (2018) point out, the sample sizes in such studies tend to be small so additional data may be needed (but see Burns & Bukach, 2021; Gerlach & Starrfelt, 2022).

Even though word and face processing disabilities may not be strictly confined to a category, there might not be a single shared general mechanism for all object recognition; this could depend on what aspects of an object's appearance are task-relevant. Various relatively distinct brain areas have been linked with the processing of fundamental visual features such as symmetry (Sasaki et al., 2005), curvature (Yue et al., 2014), and rectangularity (Nasr et al., 2014), as well as the nature of the processing, such as the analysis of features or their configurations (Lobmaier et al., 2010; Rossion et al., 2000). Featural processing refers to the analysis of specific visual features (e.g., letters in words or eyes in faces) while configural processing involves spatial relations between features (Maurer et al., 2002).

One possible reason why visual word, face, and object processing are sometimes found to be associated and sometimes dissociated is that different tasks and stimulus sets might require similar or different types of visual processing. Individual differences in word and face perception such as those reflected in developmental dyslexia and developmental prosopagnosia could in other words be process-specific rather than category-specific. One possibility is that dyslexia involves a primary impairment in featural processing while prosopagnosia involves configural processing deficits, as word recognition is assumed to be mainly feature-based (Pelli et al., 2003) while configural processing is considered the hallmark of face processing (McKone et al., 2007; McKone & Robbins, 2007; Rossion, 2013). However, some claim that configural processing of faces relies on domain-general mechanisms that can been recruited for object perception through accumulated visual expertise (e.g., Bukach et al., 2006; Gauthier & Bukach, 2007; Gauthier & Tarr, 2002; but see McKone et al., 2007). Another possibility is that both disorders involve impaired configural processing; both words and faces are objects of expertise (e.g., Gauthier et al., 2000; Gauthier & Nelson, 2001; Ventura et al., 2019; Wong & Gauthier, 2007), and gaining expertise for a visual category could increase configural processing of objects belonging to that category (Wong, Palmeri, & Gauthier, 2009; Wong, Palmeri, Rogers, et al., 2009). If people with dyslexia and prosopagnosia have a problem with acquiring visual expertise (Barton et al., 2019; Lieder et al., 2019; Sigurdardottir et al., 2017), their recognition problems may reflect impairments in configural processing.

Developmental prosopagnosia has often been linked to problems with configural face processing (Avidan et al., 2011; DeGutis et al., 2014; Palermo et al., 2011; Towler et al., 2018). But while several studies have reported impaired configural processing in developmental prosopagnosia (Avidan et al., 2011; Liu & Behrmann, 2014; Palermo et al., 2011), other studies have not (Biotti et al., 2017; Le Grand et al., 2006; Susilo et al., 2010; Ulrich et al., 2017). Additionally, some people with developmental prosopagnosia have been found to perform worse than control groups on both featural and configural face processing (Gerlach et al., 2016; Le Grand et al., 2006; Yovel & Duchaine, 2006). Dyslexic readers have shown configural processing comparable to controls for face recognition (Brady et al., 2021; Sigurdardottir et al., 2015, 2019) or even stronger for word recognition (Brady et al., 2020; R. V. Tso et al., 2021; R. V. Y. Tso et al., 2020), suggesting that dyslexic readers could be primarily deficient at featural processing. Sigurdardottir et al., (2021) reported impaired featural processing in dyslexic readers, while Jozranjbar et al. (2021) reported no difference in featural and configural processing accuracy for dyslexic and typical readers. Diverging results may in some cases be due to different manipulations of configural processing (e.g, difference between second-order configural processing and holistic processing; Maurer et al., 2002; Roberts et al., 2015). However, just as all roads might lead to Rome, there can be many possible ways of reaching equivalent performance on a given task. Representational similarity analyses (RSA; see below) in Jozranjbar et al., (2021) revealed that dyslexic readers – unlike typical readers – rely on a single visual processing mechanism for face and house recognition regardless of whether features or configurations are task-relevant. Importantly, the RSA analyses revealed patterns of associations and dissociations that group comparisons of overall accuracy failed to uncover.

While studies on developmental disorders are clearly important, drawing conclusions about typical development or cognitive architecture based on studies of neurodevelopmental disorders is not straightforward. Some accounts (Bishop, 1997; D'Souza & Karmiloff-Smith, 2011) suggest that developmental deficits (e.g., dyslexia and prosopagnosia) partly reflect that the brains of individuals with such deficits develop differently and that investigating developmental disorders to gain insights into normal cognition ignores the dynamic and complex interactions of numerous emerging systems that characterize the developing brain and its functional plasticity (Moses & Stiles, 2002). These accounts emphasize that the infant's brain is initially highly interconnected, and neural networks only become more specialized or modularized throughout development. Even when behavioral scores of people with developmental disorders fall within the normal range, they may be supported by different cognitive and neurological processes in atypical development. Neurodevelopmental disorders may in other words not reflect the continuous process of relative modularization and modularity might be the end state of development, but not its starting point (D'Souza & Karmiloff-Smith, 2011). Thus, face processing of people with developmental prosopagnosia may start out as abnormal and then continue to develop abnormally. Starrfelt and Robotham (2018) suggest that even if there is no distinction between face and object recognition in developmental prosopagnosia, this does not necessarily have implications for the cognitive architecture of the face processing system in typically developed adults, which may still possess module-like perceptual systems for face processing. In addition, findings on developmental dyslexia should be interpreted with caution due to evidence of temporal, motor, attentional, auditory, and visual dysfunction (De Martino et al., 2001; Farmer & Klein, 1995; Giofrè et al., 2019; Goswami, 2011; Kristjánsson & Sigurdardottir, 2022; Norton et al., 2015; Reid, 2018; Valdois et al., 2004; Ziegler et al., 2010), indicating that developmental dyslexia is a heterogeneous deficit. For these reasons, investigating individual differences in face, word, and object recognition in neurotypical populations may provide insights that complement studies of neurodevelopmental conditions.

Current study

The primary aim of the present study is to answer the following general question: *What are the organizing principles of visual object processing?* More specifically: Do different visual categories (here: faces, houses, words) and visual manipulations (here: featural, configural) rely on unique or joint mechanisms? Do some rely on a specific mechanism while others utilize a common, domain-general mechanism such as *o* or *VG*? Is better performance for one category even associated with worse performance for another, indicating a tradeoff? Is the same type of featural vs. configural process applied across visual categories or are such processes category-specific?

We addressed these possibilities by assessing individual differences in visual recognition abilities of a large sample (N = 101) of people who did not complain of problems with word or face recognition. We specifically measured people's use of featural vs. configural information of unfamiliar faces, unfamiliar houses, and unfamiliar words (pseudowords) in a visual delayed match-to-sample recognition task.

To foreshadow our results, our research indicates that there may be a general mechanism for recognizing visual objects, but some categories and manipulations likely require additional mechanisms. The results may help us better understand the similarities and differences between neurodevelopmental conditions where visual recognition is affected, including developmental prosopagnosia, developmental dyslexia, and developmental object agnosia, and could extend our understanding of other populations that may have strengths and weaknesses in visual recognition, including super recognizers, autistic people, and people with Williams syndrome.

Method

The study protocol was reviewed by the Internal Review Board of Icelandic Universities (Siðanefnd háskólanna um vísindarannsóknir; ID SHV2021-001). Informed consent was obtained from every participant in the study, in each session. The study was preregistered on the Open Science Framework prior to data collection (https://osf.io/a9wjy).

Participants

Participants were recruited through Prolific (prolific.co). The study was administered with Pavlovia (https://pavlovia.org/). Participants were screened on predefined questions:

normal or corrected-to-normal vision (YES), hearing loss (NO), English as a first language (YES), age (18 years and older), and Autistic Spectrum Disorder (NO). We used Prolific to distribute our study equally across men and women. The reason for excluding people with autistic spectrum disorder is that our study concerns face perception, among other things, and people on the autism spectrum tend to process faces differently from neurotypical people (Griffin et al., 2021). We excluded people with hearing impairments as they might possibly affect reading (Moeller et al., 2007) and understanding of our verbal instructions. Additionally, by accepting the consent form, participants confirmed that they did not have suspected or confirmed brain damage (which can affect object, face, or word recognition), and that they were using a laptop or a desktop computer with sound. Our final sample included 97 participants (56 women, 40 men, 1 other; mean age: 35 years, range 18–74 years; education: 33 high school graduates, 12 technical/community college graduates, 34 bachelor's degree holders, 18 graduate degree holders).

Session 1

The experiment included two sessions. Session 1 was a 5-point Likert scale survey to evaluate how participants rated their ability to recognize visual words and faces (very poor, below average, average, above average, excellent). Participants also answered background questions on age, gender (male, female, other) and education level (1. Completed high school or less; 2. Completed technical/community college; 3. Completed an undergraduate degree (BA/BSc/other); 4. Completed a graduate degree or higher (MA/Msc/MPhil/PhD/other). Session 1 took approximately 2 minutes. Participants were compensated £0.20 for their participation in this session. We recruited participants until 50 participants were identified who rated their ability to recognize visual words and faces as average or above. Our focus was on typical face and word processing, so the sample was restricted to people who did not report problems with recognizing faces or words.

Session 2

At the beginning of session 2, we conducted sound checks to ensure that participants could hear the verbal instructions. To prevent participants who had unusual screen size from participating, we displayed numbers at the top, right, left, and bottom of the screen and asked them to report the numbers via keypress. Participants then performed a visual recognition task (approximately 40-45 minutes) where they had to remember featural or configural information for faces, houses, and words, as described in detail below. Participants received £5 for successful completion of session 2.

Visual recognition task

The visual recognition task measured featural and configural processing of faces, houses, and pseudowords. Face and house stimuli were part of a larger set with featural and configural manipulations developed by Collins et al., (2012; we thank Jane E. Joseph for providing the stimuli). Pseudowords were generated by feeding a five letter word list (http://www.yougowords.com/5-letters) to Wuggy (freely available at http://crr.ugent.be/programs-data/wuggy; see Keuleers & Brysbaert, 2010)

Before the task, participants learned what constituted a featural and configural change of faces, houses, and pseudowords through verbal and visual instruction. There were 96 trials for each combination (featural faces, configural faces, featural houses, configural houses, featural pseudowords, configural pseudowords; figure 16, panel A). There were 24 blocks of 24 trials, for a total of 576 trials. Block order was counterbalanced (Latin square), and trial order within each block was randomized and was then kept fixed for all participants. We had four blocks for each combination (featural face, featural house, featural pseudoword, configural face, configural house, and configural pseudoword). For instance, the first block included 24 featural face trials, the second contained 24 featural house trials, and so on. Participants were informed which combination would be included in each block before it began.

For each sample item (face/house/pseudoword), there were image pairs: a match and foil. The pair of images was unique but individual images could be used on other trials. The match was identical to the sample, but the foil differed featurally or configurally. For featural changes, some features differed between sample and foil, such as eyes in faces, windows in houses, or lower or upper cases in pseudowords. For configural changes, some distances between features differed between sample and foil; for example, eyes, windows, or letters could be at different distances within faces, houses, and pseudowords, respectively.

Our aim with using unfamiliar faces, houses, and words, as well as keeping letter identity constant in pseudowords, was to minimize the usefulness of verbal cues. Following Conway et al., (2017), featural and configural changes were always made within the word, the first and last letter never changing, to keep overall size constant and to increase similarity with the face and house trials where only internal features or configurations change (but never e.g. ears or roof). Random modifications were made to the inter-letter spacing and case of pseudowords (identical for all participants).

After listening to instructions, participants performed a practice task consisting of twenty distinct trials with twenty stimulus pairings of fundamental geometrical objects (e.g., circle, square). Participants could proceed to the main task once they correctly completed eight consecutive practice trials (the loop was repeated otherwise). The main task involved matching faces, houses, or pseudowords where foils differed from samples either featurally or configurally. Both practice and main tasks were two-alternative forced-choice delayed match-to-sample. Participants received feedback on accuracy during practice but not during the main task.

In the primary task (figure 16, panel B), each trial started with a 500 ms fixation cross at screen center. Once it disappeared, a sample (face, house, or pseudoword) was shown for 500 milliseconds at screen center, followed by a 200 ms circular dot mask. To prevent low-level template matching, the match and foil were smaller than the sample (60% of the height and width of the sample). The match and foil remained onscreen until participants indicated which one was identical to the sample by keypress (left arrow key for the left image and the right arrow for the right image). The stimuli disappeared after the response, and the next trial started after a 500 ms inter-trial interval. Participants were given a break after each block and pressed the space bar when ready to continue.

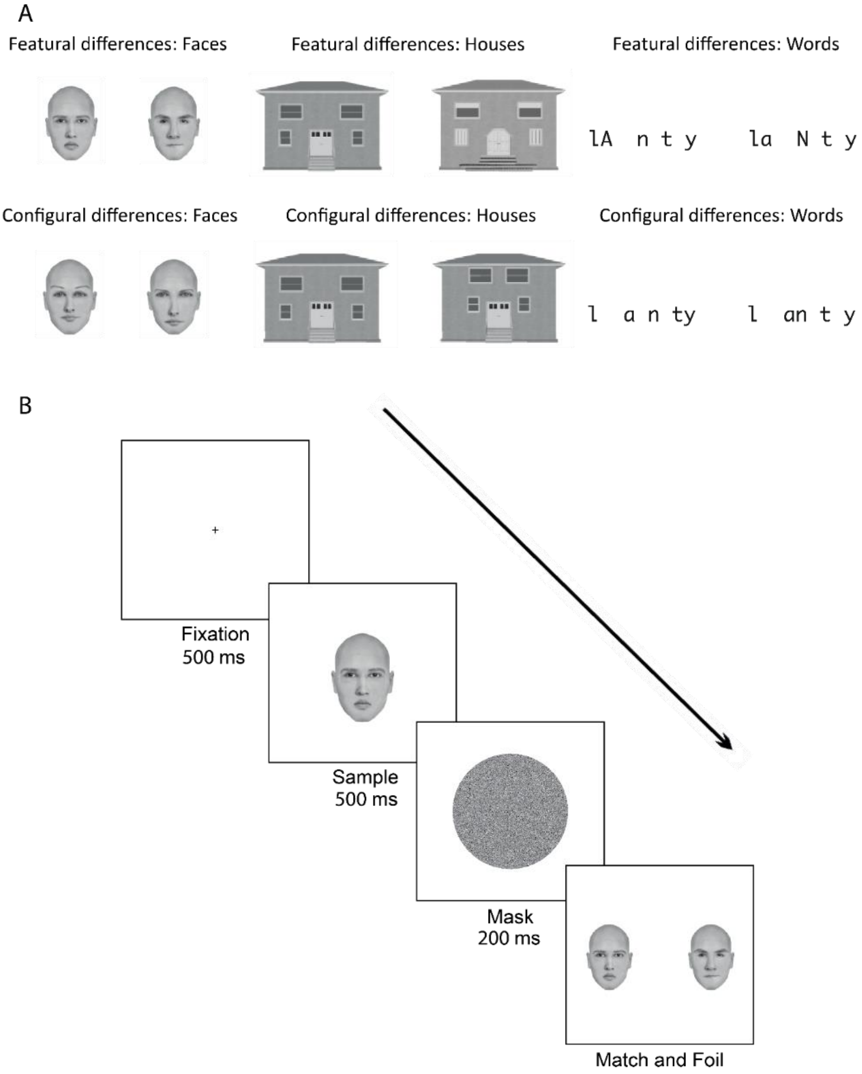


Figure 16. Visual recognition of featural and configural differences in unfamiliar faces, houses, and words. A. Examples of featural or configural changes of unfamiliar faces, unfamiliar houses, and unfamiliar words (pseudowords) of the visual recognition task. B. The experimental design. On each trial, a sample image (face, house, or pseudoword) appeared at screen center followed by

match and foil images displayed simultaneously to the left and right of screen center. The match was identical to the sample image, but the foil was different either featurally or configurally. Match and foil images remained onscreen until the participant indicated with a button press which of the images was identical to the sample image.

The visual recognition task included attention checks between blocks. For these, we used geometrical shapes after face and house blocks, for example, a square as a sample, a square as a match and a circle as a foil, and simple letters after pseudoword blocks, such as 'aaaa' as a sample, 'aaaa' as a match, and 'bbbb' as a foil. In the second session, we excluded and subsequently replaced participants from the first session who demonstrated a failure rate exceeding 33% on the attention checks. Our aim was to eliminate individuals who were evidently not fully engaged or committed to the task and recruit more suitable replacements, given that attentive and diligent participants would readily succeed in these attention tasks. The attention checks were otherwise excluded from further analysis.

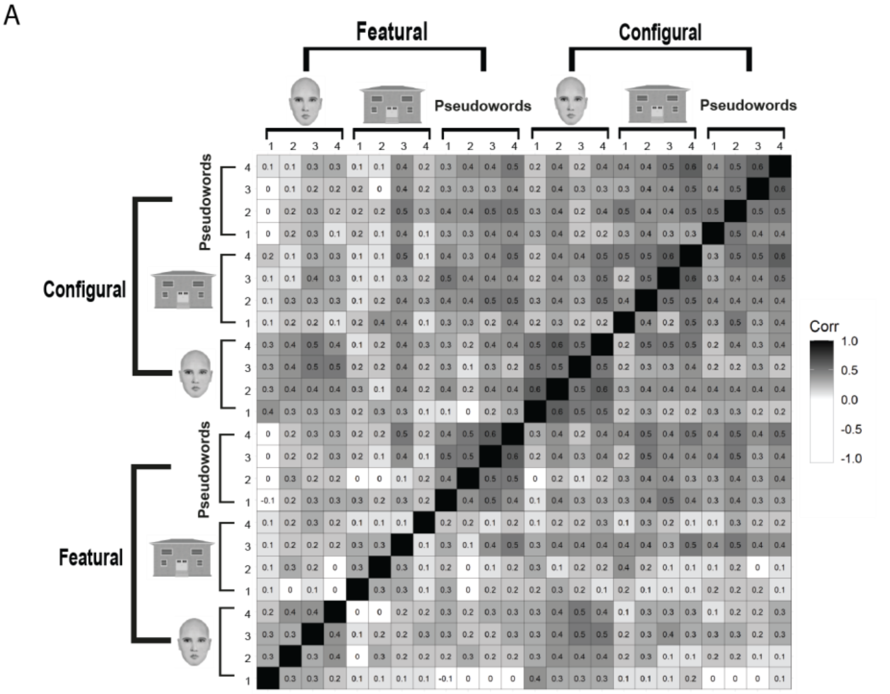
Data Analysis and Results

Reference model

The reference model is a 24 x 24 correlation matrix where each cell represents the correlation of the accuracy of two blocks of the visual recognition task (e.g., block one of featural faces and block two of configural pseudowords etc., figure 2, panel A). Visual recognition data for the 50 initially selected participants was divided into two randomly selected subsamples of 25 participants each, and the reference model of each subsample was calculated. Following that, we estimated the Pearson correlation between the upper triangles (excluding the diagonal) of the two subsample reference models. We repeated this 1000 times with different randomly selected subsamples and calculated the mean correlation. Next, we estimated the reliability of the reference model for the entire sample using the Spearman-Brown formula on the mean correlation: $R_{full} = 2(r_{half})/1+r_{half}$. The full-scale reliability was less than a preset value (0.75), so we recruited and assessed more individuals. As the full-scale reliability had not achieved the preset level by N=100, data collection was stopped. We deviate from our preregistration in the following ways: Once N=100 was reached, one additional participant was in the middle of doing the task and this person was therefore allowed to finish the study. Additionally, we excluded four out of 101 participants whose accuracies were at least 3 standard deviations from the mean of at least one of the conditions (accuracy around chance level). The reliability of this final reference model (figure 2, panel A) for the accuracy of the 97 participants was 0.63.

Conceptual models

We utilized representational similarity analysis (RSA), which originates in systems neuroscience, to estimate the separability of visual object recognition mechanisms. Our approach involved comparing the correlational matrix for accuracy (reference model) with predicted data patterns derived from conceptual models. We preregistered six conceptual models (preregistration: https://osf.io/a9wjy) based on possible predicted patterns: category-specific model, cost model, fine-tuning model, process model, holistic expertise model, and time model (see figure 17). We opted to exclude the holistic expertise model since it was solely based on data that we piloted prior to preregistration (N=10) which were later found to be unreliable (reliability estimate -0.06). We compared our reference model to the five remaining conceptual models as well as three additional theoretically important conceptual models: the face specialization model, the word specialization model, and the combination model (Figure 17, panel B). The following conceptual models were considered:



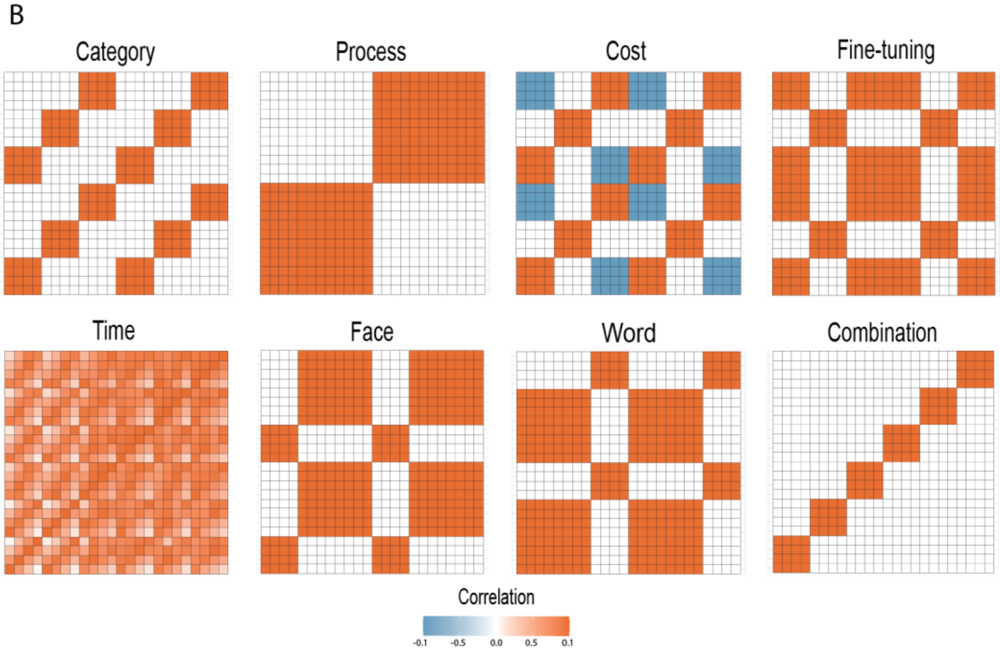


Figure 17. Observed and predicted data patterns of associations in performance. A. Reference model (correlational matrix for accuracy). B. Conceptual models (predicted data patterns). Orange reflects positive correlation, blue reflects negative correlations, and white reflects no correlation. Abbreviations: 1-4: block numbers.

Category model

In previous research (Jozranjbar et al., 2021) participants who were relatively good at identifying faces were not necessarily the same as those who were good at recognizing houses. Recognition of these objects appeared to rely on partly separable abilities. If each stimulus category (faces, houses, words) relies on a unique mechanism, the category model that assumes greater correlation between performance in same-category blocks than between-category blocks will predict our results. For the category model, when both blocks share a category type (e.g., both faces), the values for the conceptual model are "1"; otherwise, they are "0" (e.g., one face and another one pseudoword): If people's performance on one stimulus category is better predicted by their performance on other blocks from the same category than blocks from another category, a reference model will show a similar pattern.

Process model

Instead of being specialized for different categories, visual processing mechanisms may differ according to which type of information they handle (e.g. featural or configural; Lobmaier et al., 2010; Maurer et al., 2002; Rossion et al., 2000). To test this possibility, we manipulated featural (e.g., letters in words, windows in houses, eyes in faces) and configural information, referring to the spacing between features (e.g., absolute distance between nose and lip). If featural and configural processing are separable, the pattern in our data should fit the process model (see also Jozranjbar et al., 2021). This model assumes that performance in blocks where featural information is manipulated should best predict performance in other featural blocks, while performance in blocks where configural information is manipulated should best predict performance in other configural blocks. The values for the process model are "1" when both blocks involve the same manipulation (both featural or both configural), and "0" otherwise (one featural and one configural). If featural and configural processing are separable, performance with one type of process manipulation should better predict another block with that same process manipulation in the reference model.

Cost model

According to the cost model, words compete with faces for cortical resources (Dehaene et al., 2010) and better word recognition performance could be related to worse face recognition performance. The cost model assumes some category specificity, but additionally assumes negative links between word and face processing. The values for the cost models are "1" when both blocks share a category type; and if the blocks involve houses vs. faces, or houses vs. words the values are "0". But when one block involves faces and the other words, the values are "-1".

Fine-tuning model

The fine-tuning account (Hervais-Adelman et al., 2019) argues that training on orthographic stimuli does not result in a loss of responsiveness to faces but may rather induce an increase in shared aspects of the representation between words and faces. Better word recognition performance would thus go hand in hand with better face recognition. The fine-tuning model assumes some category specificity, but additionally assumes positive links between word and face processing. The fine-tuning model is equivalent to the cost model except that when one block involves faces and the other one words, the value is "1" in the fine-tuning model.

Time model

The time model will fit the data if performance drifts or fluctuates over time, where some participants may for example pay close attention to the task for a few blocks but then gradually lose concentration, leading to a positive correlation between performance in adjacent blocks. The time model could account for our empirical reference model together with other conceptual models or be the sole predictor of our results if there is complete domain-generality, i.e., no differences by category, process, or their combination and only a time difference. Values for the time model are based on the temporal proximity of the 24 blocks. For example, the first and second blocks have a value of " 0.96" since they are the closest in time, the first and third blocks have a value of " 0.92", and so on.

Face model

The pattern in our reference model should be consistent with the face model if faces tap into domain-specific processes (e.g., Kanwisher, 2000), distinct from house and word processing. Face model values are "1" when both blocks include faces and "0" when one block has faces and the other contains words or houses. Likewise, the values are "1" if both blocks include words or houses (e.g., one word, another one house, or both houses or words).

Word model

Our reference model should reflect the word model if words tap into domain-specific processes, distinct from face and house processing. Word model values are "1" when both blocks include words and "0" when one block has words and the other contains faces or houses. Likewise, the values are "1" if both blocks include faces or houses (e.g., one face, another one house, or both houses or faces).

Combination model

The combination model postulates the absence of domain-generality and hypothesizes high association only when blocks share both category and process. For the combination model, the values are "1" when blocks share both category and manipulation (e.g., both display configural houses); otherwise, they are "0".

Representational Similarity Analysis (RSA)

The Representational Similarity Analysis (RSA) assessed the similarity between the reference and conceptual models. The diagonal and the mirror version of the opposite off-diagonal triangle of each reference or conceptual model were removed and then the model was converted into a vector.

We first chose to consider the cost model over the fine-tuning model as the former fit the data better; they are in direct opposition to one another and create a multicollinearity issue (for details, see preregistration: https://osf.io/a9wjy). Then, using the vector of the reference model as a dependent variable and the vectors of the remaining conceptual models as independent variables, we ran Bayesian multiple regression to identify which model, or which combination of models, best predict the observed data. The Bayesian linear regression procedure begins by considering 127 models, including one null model in which no other model has any association with the reference model and seven models in which only a single conceptual model (e.g., the time model) is associated with the reference model. The remaining models include various combinations of the conceptual models. We analyzed the data with JASP (JASP Team, 2022). A .jasp file, including plots, data, and input options, is available at https://osf.io/6kjyq/. A Jeffreys-Zellner-Siow (JZS) prior and a beta binomial model prior were implemented. The Bayes factors were calculated relative to the best model.

The statistical model that best fit the actual data patterns in the reference model included the time model, the face model, and the combination model. Table 5 shows that the five best models all contain these three conceptual models. This suggests that these predictors play an important role in producing the observed data patterns.

Table 5. The five best models from the Bayesian linear regression for the visual recognition task. P(M) are the prior model probabilities; P(M|data) are the posterior model probabilities; BFM is the change from prior to posterior model odds; BF10 is the Bayes factor of the best model over the

model in that row; and R^2 is the explained variance of each model. Results for all 127 models are presented in the .jasp file.

Models	P(M)	P(M|data)	BF_M	BF_{10}	R^2
Time + Face + Combination	0.004	0.235	85.470	1.000	0.275
Time + Cost + Face + Word + Combination	0.006	0.097	17.894	0.248	0.278
Process + Category + Time + Cost + Face + Word + Combination	0.125	0.096	0.745	0.012	0.279
Process + Time + Cost + Face + Word + Combination	0.018	0.066	3.893	0.056	0.279
Time + Cost + Face + Combination	0.004	0.058	17.200	0.248	0.278

With this approach, parameter estimates depend on the chosen model. Bayesian model averaging can be used to determine the relevance of individual predictors. Instead of estimating parameters using a single model, this algorithm weighs each model's contribution by its posterior probability (Bergh et al., 2021). We used a Bayes factor's categorization scheme where Bayes factors between 1 and 3 constitute anecdotal evidence for a hypothesis, 3 to 10 = moderate, 10 to 30 = strong, 30 to 100 = very strong, and greater than 100 = extreme evidence (Jeffreys, 1998). Table 6 summarizes predictor inclusion probabilities and posterior distributions for all predictors. The evidence for the inclusion of the time and combination conceptual models is extreme, and moderate for the face model.

Table 6. Model-averaged posterior summary for linear regression coefficients for the visual recognition task. The leftmost column denotes the predictor. P(incl|data) denotes the posterior inclusion probability. The change from prior to posterior inclusion odds is given by the inclusion Bayes factor (BFinclusion). The columns 'mean' and 'sd' represent the respective posterior mean and standard deviation of the parameter after model averaging. The columns 'lower and 'upper' indicate the 95% central credible intervals.

Coefficient	P(incl|data)	$BF_{inclusion}$	Mean	SD	Lower	Upper
Intercept	1.000	1.000	0.289	0.008	0.274	0.304
Process	0.334	0.502	0.004	0.012	-0.010	0.041
Category	0.314	0.459	0.003	0.018	-0.030	0.056
Time	1.000	1.129e+6	0.204	0.035	0.131	0.275
Cost	0.543	1.190	0.035	0.040	0.000	0.106
Face	0.881	7.383	0.042	0.040	0.000	0.099
Word	0.537	1.158	-0.025	0.037	-0.100	0.024
Combination	0.997	355.917	0.114	0.029	0.053	0.171

The figures in table 6 also argue against the relevance of process, category, cost, and word models where the data decreased the inclusion probability (see figure 18 for shift from prior to posterior inclusion probability). The posterior distribution is defined as updated knowledge regarding the relative plausibility of parameter values given the observed data. The proximity of the posterior inclusion probabilities of time, face, and combination models (1.000, 0.881, and 0.997, respectively) to 1 demonstrates that each of these predictors is useful for predicting our observed data.

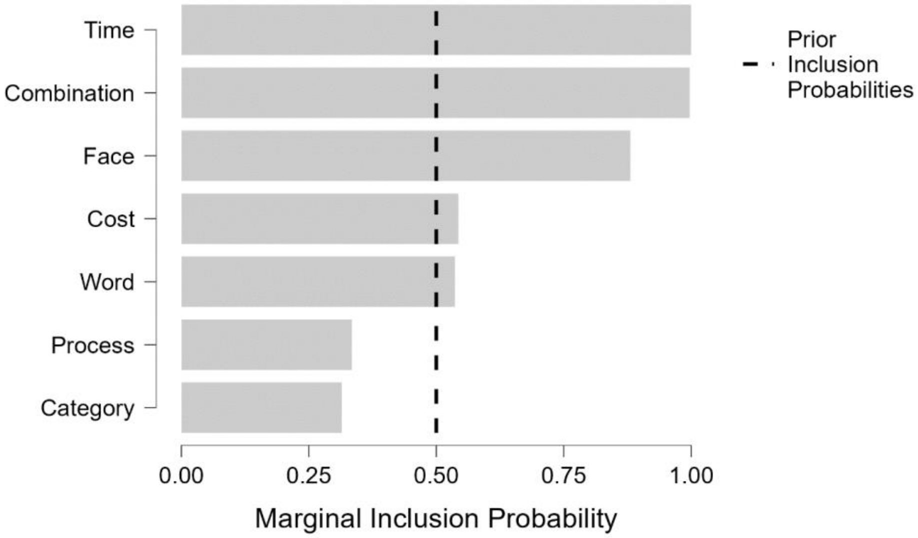


Figure 18. Posterior inclusion probabilities for the Bayesian linear regression. The dashed line represents the prior inclusion probabilities.

Our results strongly support the combination model, which assumes high association between blocks only when they share both category and process. Note that all combinations need not adhere to this pattern. We therefore performed an exploratory multidimensional scaling on the reference model. As illustrated in figure 19, the data clearly separated into face vs. non-face blocks, in alignment with the face model, and certain category-process combinations additionally clustered together, in alignment with the combination model. Featural houses also separated from the other combinations.

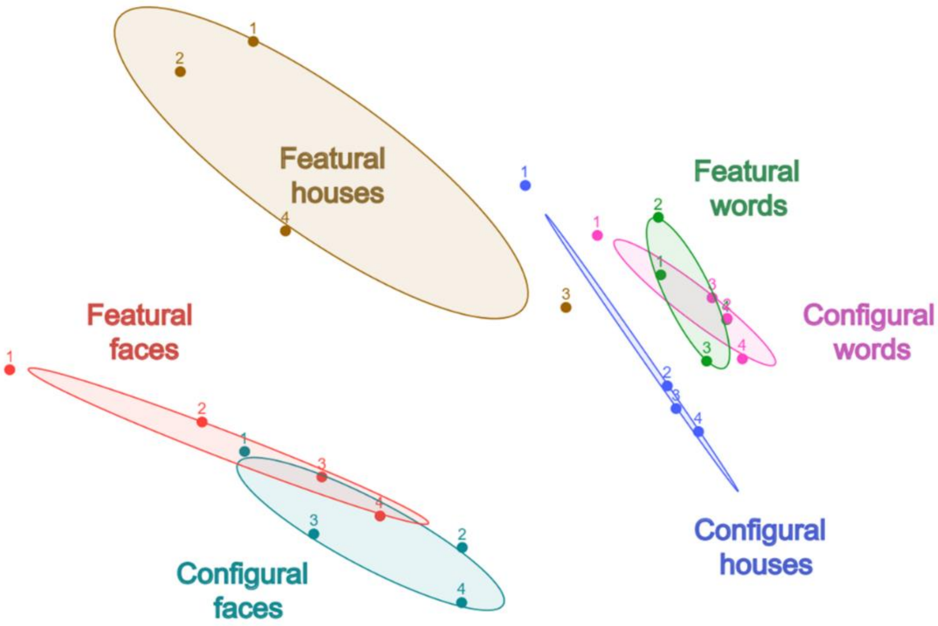


Figure 19. Multidimensional scaling of the accuracy on the visual recognition task. Two-dimensional scaling projects 24-dimensional data to a 2-dimensional space such that similar conditions in the 24-dimensional space will be close together on the two- dimensional plot. Dots represent blocks, numbers are block numbers. Ellipses show the confidence ellipses around barycenter of each cluster.

Discussion

Several neurodevelopmental conditions involve unusual recognition abilities or disabilities, and studying such conditions has contributed to our understanding of the processes involved in visual object recognition. Studies on typical populations may, however, also advance our understanding of neurodiversity, by identifying which mechanisms that might be shared or differ between neurodevelopmental conditions affecting visual recognition. In the current study, we examined whether visual object recognition relies on dissociable or shared mechanisms by asking people to recognize faces, houses, and words by their features or feature configurations. According to proposals of domain-general object identification abilities (*o* or *VG*), visual recognition accuracy for one object category should strongly predict performance for another category (Hendel et al., 2019; Richler et al., 2019). Visual recognition blocks were positively correlated, with a few exceptions, consistent with the proposal of such a shared component.

In accordance with arguments that face processing is unique (Kanwisher et al., 1997, 1997; Robbins & McKone, 2007; Tanaka & Farah, 1993; Young et al., 2013), our results furthermore indicate that face recognition employs specific mechanisms different from mechanisms supporting word and house processing. Faces are objects of expertise, and experience with objects may weaken the correlation between objects that are otherwise closely related (Gauthier et al., 2014; Ryan & Gauthier, 2016; Van Gulick et al., 2016). This may not fully explain the separability of faces, however, as performance for words – which are also objects of expertise – was quite closely connected to configural processing of houses, albeit less with featural houses. Our participants have significant experience with word recognition, so it is possible that words, regardless of our specific manipulations, are processed configurally due to their status as objects of expertise. Other visual categories should be evaluated before drawing firm conclusions, however.

We found no specific relationship between faces and words, above and beyond that for houses. We therefore reject the cost model which assumes that visual word representations specifically compete for neural resources with faces and negatively impact face recognition abilities. We also reject the fine-tuning model as operationally defined in this study. Reading development may still induce an increase in shared aspects of the representation between words and faces, but our results suggest that this is not *specific* to faces. We also note that visual expertise does not seem to moderate the shared variance between object categories as faces and words (objects of expertise for most people) were not more correlated with each other than with houses (generally not an expertise category). Visual expertise for faces and words may therefore be relatively independent.

The processing of features and the processing of configurations may reflect separable processes that operate across categories. For example, global priming with compound figures boosted the configural processing (more specifically, holistic processing) of words and faces comparably, suggestive of similar mechanisms (Ventura et al., 2021, see also Ventura et al., 2022, but see Ventura et al., 2017; 2019). We did not observe separable featural and configural processing independent of visual categories. Performance in blocks where configural information was manipulated did not necessarily predict performance in other configural blocks (over and above that for featural blocks) and blocks where featural information was manipulated did not necessarily best predict other featural blocks (over and above configural blocks). One possible reason is that manipulating "features" may result in unintended modifications of "configurations", and vice versa (Rakover, 2002). For instance, changing the distance between the eyes – a "configural" manipulation – may be considered a featural modification of the nasal bridge's size. This explanation however seems inconsistent with our previous study where featural and configural processes were apparently separable (Jozranjbar et al., 2021). In that study, trials were not blocked so participants were unaware of the upcoming category of a to-be-remembered sample, and furthermore did not know whether featural or configural information would be task-relevant. Separating the conditions into blocks may have allowed participants to adjust their strategy to specific *combinations* of category and processing that are optimal for each visual category.

Accordingly, we found that at least some combinations of manipulation (featural/configural) and category (unfamiliar houses/words/faces) are "special", having distinct variance not shared with other combinations; for instance, people's accuracy for configural information in unfamiliar faces was more related to their performance for configural differences in other faces than to any other combination of process and category.

Featural processing of a specific category may have little overlap with featural processing of other categories, as their high-level features are largely distinct (e.g., words never have windows, houses never have eyes). Configural processes could also be category-specific as configural detectors, such as potential bigram detectors, may also be sensitive to information that is distinct for a particular category. Not all combinations however appear to be "special". Blocks of featural houses were not noticeably more associated with other blocks with this combination versus other combinations (figure 2A). All object recognition may partially rely on a common, domain-general mechanism (*o/VG)* in addition to other mechanisms, but featural houses may almost exclusively rely on *o/VG*. Furthermore, in the multidimensional scaling (MDS, figure 4), featural and configural words cluster together and overlap even more than featural vs. configural faces or featural vs. configural houses. Keeping in mind that MDS is a simplification of the true pattern in our data, this may suggest that distinguishing featural and configural processing is less relevant for the visual processing of words.

Potential limitations

Some potential limitations should be considered. One such limitation pertains to the intrinsic differences between stimuli from different categories that could affect the task demands, even when the experimental task is similar (Robotham & Starrfelt, 2017a; 2017b). For instance, configural sensitivity may vary between words and faces, with even minor changes in facial features leading to significant differences in identity perception. In contrast, the identity of a word can be invariant to changes in font or kerning. Despite high sensitivity to subtle changes in perception of faces and facial expression, face recognition can also exhibit significant tolerance towards major differences in appearance: we can perceive a face as belonging to the same person despite changes in lighting, makeup, hairstyle, and even age. As a result, matching faces with other stimuli can be a challenging and intricate task.

We made the deliberate decision to use pseudowords instead of real words due to the unique challenges that would have arisen in the latter case. Specifically, if we had used familiar words, a fair comparison would have been with familiar faces and familiar houses. This would have required technically challenging manipulations of featural and configural properties of real-world stimuli. Additionally, this would have led to the unmanipulated face/house being the only familiar stimulus, while any form of manipulation would have rendered the stimulus unfamiliar. Moreover, not all participants may be equally experienced with all supposedly familiar stimuli, which could have further complicated the study. It is nonetheless important to acknowledge that people's expertise with pseudowords may not be equivalent to our proficiency with unfamiliar faces. For example, Ventura et al. (2017) showed that there is no composite effect for pseudowords, an effect often interpreted as a signature of perceptual expertise. This, however, does not imply a complete lack of expertise for pseudowords. For example, letter identification is not only superior within words compared to nonwords, but also better within pseudowords than in consonant strings (pseudoword superiority effect, e.g., Adams, 1979; Baron & Thurston, 1973; Grainger et al., 2003; Grainger & Jacobs, 1994; Spoehr & Smith, 1975). However, it remains possible that using real words as stimuli would have rendered the word condition separable from the face and house conditions, reflecting the fast, automatic, interactive processing that characterizes expertise in visual word recognition. It is also still debated whether we possess expertise in recognizing unfamiliar faces (e.g., Ritchie et al., 2023; Young & Burton, 2018).

We should also note that there are differences in the manipulations of pseudowords and faces/houses. For example, featural changes were quite predictable for pseudowords, as they always involved upper-lower case modifications, while features of the other categories could change in a number of ways. It was, however, crucial for us to ensure that our changes to the word features did not affect the pronunciation of the pseudowords used in the study. If we had made letter changes, it could have caused the participants to rely more on their verbal working memory. Additionally, in the configural change conditions, pseudowords always involved horizontal modifications, while the changes occurred in two dimensions for faces and houses. We chose this manipulation as vertically altering pseudowords would have made the stimuli appear artificial and unrealistic. We however also recognize that horizontal modifications in pseudowords may possibly lead to unexpected ways of processing these stimuli as the presence or absence of spaces may impact the bigrams extracted from a given text. However, if these choices had greatly influenced our data, we would have expected a particular relationship between faces and houses, which were similarly manipulated, which would have been different from words, but this was not the case. Therefore, we find it unlikely that our choices in the way in which we manipulated pseudowords played a strong role in explaining the observed patterns in our data.

A potential problem in our approach is that the conceptual models are not orthogonal to each other. We were aware of this and prioritized selecting between the cost model and the fine-tuning model, as they were directly opposed to each other and created a multicollinearity issue (for details, see preregistration: https://osf.io/a9wjy). We recognize that some multicollinearity between other models may be unavoidable, as the models are related in a way that is beyond our control. However, following Occam's razor, we found that the simplest model is typically the most effective. Despite some models performing equally well, the least complex model tended to outperform the other top five models.

Implications for developmental neurodiversity in recognition

Our results shed light on developmental conditions that involve unusual abilities or weaknesses in visual recognition. The special status of faces may fit with claims of developmental prosopagnosia being distinct from other developmental disorders that involve recognition deficits, such as developmental dyslexia and developmental object agnosia (Germine et al., 2011; Gray & Cook, 2018). It could also fit with claims of face processing being specifically affected in other developmental disorders, including autism (Griffin et al., 2021) and Williams syndrome (Järvinen et al., 2013), and being selectively enhanced in super-recognizers (Bobak et al., 2016; Davis et al., 2016). However, we find little evidence for such a special status of visual pseudoword processing in our data. Also, since we found partial support for domain-general mechanisms, certain conditions, such as developmental object agnosia, may reflect difficulties with o/VG-like mechanisms, possibly with spared abilities in domains that rely on more specialized mechanisms. A domain-general mechanism may possibly be at play in other cases of neurodiversity, including people with superior face recognition skills and even in some people diagnosed with developmental dyslexia. The former group has been found to be better than controls not just at face processing, but also at objects and words (Hendel et al., 2019), indicating that some such cases may reflect superior o/VG. Developmental dyslexia may be a "mixed bag", with some cases showing no problems with visual face or object processing (e.g., due to a non-visual etiology of the disorder such as a phonological processing deficit, Snowling, 1998), some showing a visual word recognition problem that generalizes to objects such as houses but with spared accuracy for faces (Jozranjbar et al., 2021), and some that show visual recognition problems with both objects and faces (Sigurdardottir et al., 2015) which could be indicative of a relatively generalized high-level visual processing disorder.

We also think that a more general methodological point can be taken from our results. We believe that our results clearly demonstrate the usefulness of representational similarity analyses (RSA) for the investigation of diversity in visual representations in neurodevelopmental conditions. The insights provided by RSA findings can prove invaluable in gaining a better understanding of the conditions mentioned in this subsection and could address ongoing theoretical and methodological debates pertaining to neurodiversity.

Overall Conclusions

Our representational similarity analysis of individual differences in visual recognition accuracy indicates that face processing draws on unique mechanisms that are different from those supporting processing of pseudowords and houses. We found no specific association between faces and words above and beyond that for houses, suggesting that visual expertise does not moderate the shared variance across objects of expertise like faces and words. A general mechanism (*o/ VG*) could partially support object recognition and especially the processing of features of non-expert object categories such as houses. Additionally, visual processing of features and configurations appears to be categorically modulated, where featural or configural processing of one category is separate from the featural or configural processing of another category. Overall, the results of our study offer valuable insights into the mechanisms underlying visual recognition and their association with neurodevelopmental conditions.

Acknowledgements

We would like to thank Jane E. Joseph for providing us with visual stimuli which were a part of a larger set developed by Collins et al. (2012). This work was supported by The Icelandic Research Fund (Grant No. 174013-051) and the University of Iceland Research Fund. Part of this paper was presented as poster at The European Conference on Visual Perception (ECVP) 2022, Nijmegen, the Netherlands.

References

Adams, M. J. (1979). Models of word recognition. Cognitive Psychology, 11(2), 133–176. https://doi.org/10.1016/0010-0285(79)90008-2

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia*, *49*(9), 2541–2552.

Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. Cognitive Psychology, 4(2), 207–228. https://doi.org/10.1016/0010-0285(73)90012-1

Barton, J. J., Albonico, A., Susilo, T., Duchaine, B., & Corrow, S. L. (2019). Object recognition in acquired and developmental prosopagnosia. *Cognitive Neuropsychology*, *36*(1–2), 54–84.

Behrmann, M., & Plaut, D. C. (2013). Distributed circuits, not circumscribed centers, mediate visual recognition. *Trends in Cognitive Sciences*, *17*(5), 210–219.

Behrmann, M., & Plaut, D. C. (2015). A vision of graded hemispheric specialization. *Annals of the New York Academy of Sciences*, *1359*(1), 30–46.

Biotti, F., Wu, E., Yang, H., Jiahui, G., Duchaine, B., & Cook, R. (2017). Normal composite face effects in developmental prosopagnosia. *Cortex*, *95*, 63–76.

Bishop, D. V. (1997). Cognitive neuropsychology and developmental disorders: Uncomfortable bedfellows. *The Quarterly Journal of Experimental Psychology Section A*, *50*(4), 899–923.

Bobak, A. K., Bennetts, R. J., Parris, B. A., Jansari, A., & Bate, S. (2016). An in-depth cognitive examination of individuals with superior face recognition skills. *Cortex, 82*, 48-62.

Brachacki, G. W., Nicolson, R. I., & Fawcett, A. J. (1995). Impaired recognition of traffic signs in adults with dyslexia. *Journal of Learning Disabilities*, *28*(5), 297–301.

Brady, N., Darmody, K., Newell, F., & Cooney, S. M. (2020). *Holistic processing of words and faces in dyslexia*.

Brady, N., Darmody, K., Newell, F. N., & Cooney, S. M. (2021). Holistic processing of faces and words predicts reading accuracy and speed in dyslexic readers. *Plos One*, *16*(12), e0259986.

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences*, *10*(4), 159–166.

Burns, E. J., Bennetts, R. J., Bate, S., Wright, V. C., Weidemann, C. T., & Tree, J. J. (2017). Intact word processing in developmental prosopagnosia. *Scientific Reports*, *7*(1), 1–12.

Burns, E. J., & Bukach, C. M. (2021). Face processing predicts reading ability: Evidence from prosopagnosia. *Cortex*, *145*, 67–78.

Collins, E., Dundas, E., Gabay, Y., Plaut, D. C., & Behrmann, M. (2017). Hemispheric organization in disorders of development. *Visual Cognition*, *25*(4–6), 416–429.

Collins, Zhu, X., Bhatt, R. S., Clark, J. D., & Joseph, J. E. (2012). Process and domain specificity in regions engaged for face processing: An fMRI study of perceptual differentiation. *Journal of Cognitive Neuroscience*, *24*(12), 2428–2444.

Conway, A., Brady, N., & Misra, K. (2017). Holistic word processing in dyslexia. *Plos One*, *12*(11), e0187326.

Davis, J. P., Lander, K., Evans, R., & Jansari, A. (2016). Investigating predictors of superior face recognition ability in police super-recognisers. *Applied Cognitive Psychology, 30*(6), 827-840.

De Martino, S., Espesser, R., Rey, V., & Habib, M. (2001). The "temporal processing deficit" hypothesis in dyslexia: New experimental evidence. *Brain and Cognition*, *46*(1–2), 104–108.

DeGutis, J., Cohan, S., & Nakayama, K. (2014). Holistic face training enhances face processing in developmental prosopagnosia. *Brain*, *137*(6), 1781–1798.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. Trends in Cognitive Sciences, 15(6), 254–262.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, *330*(6009), 1359–1364.

D'Souza, D., & Karmiloff-Smith, A. (2011). When modularization fails to occur: A developmental perspective. *Cognitive Neuropsychology*, *28*(3–4), 276–287.

Duchaine, B. C., & Nakayama, K. (2006). Developmental prosopagnosia: A window to content-specific face processing. *Current Opinion in Neurobiology*, *16*(2), 166–173.

Duchaine, B., & Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience*, *17*(2), 249–261.

Dundas, E. M., Plaut, D. C., & Behrmann, M. (2013). The joint development of hemispheric lateralization for words and faces. *Journal of Experimental Psychology: General*, *142*(2), 348.

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin & Review*, *2*(4), 460–493.

Gabay, Y., Dundas, E., Plaut, D., & Behrmann, M. (2017). Atypical perceptual processing of faces in developmental dyslexia. *Brain and Language*, *173*, 41–51.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition*, *103*(2), 322–330.

Gauthier, I., McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Van Gulick, A. E. (2014). Experience moderates overlap between object and face recognition, suggesting a common ability. *Journal of Vision*, *14*(8), 7–7.

Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. *Current Opinion in Neurobiology*, *11*(2), 219–224.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*(2), 191–197.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 431.

Gerlach, C., Barton, J. J. S., Albonico, A., Malaspina, M., & Starrfelt, R. (2022). Contrasting domain-general and domain-specific accounts in cognitive neuropsychology: An outline of a new approach with developmental prosopagnosia as a case. Behavior

Gerlach, C., Klargaard, S. K., & Starrfelt, R. (2016). On the relation between face and object recognition in developmental prosopagnosia: No dissociation but a systematic association. *PloS One*, *11*(10), e0165561.

Gerlach, C., & Starrfelt, R. (2022). Face processing does not predict reading ability in developmental prosopagnosia: A commentary on Burns & Bukach (2021). *Cortex*.

Germine, L., Cashdollar, N., Düzel, E., & Duchaine, B. (2011). A new selective developmental deficit: Impaired object recognition with normal face recognition. *Cortex*, *47*(5), 598–607. https://doi.org/10.1016/j.cortex.2010.04.009

Geskin, J., & Behrmann, M. (2018). Congenital prosopagnosia without object agnosia? A literature review. *Cognitive Neuropsychology*, *35*(1–2), 4–54.

Giofrè, D., Toffalini, E., Provazza, S., Calcagnì, A., Altoè, G., & Roberts, D. J. (2019). Are children with developmental dyslexia all the same? A cluster analysis with more than 300 cases. *Dyslexia*, *25*(3), 284–295.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10.

Grainger, J., Bouttevin, S., Truc, C., Bastien, M., & Ziegler, J. (2003). Word superiority, pseudoword superiority, and learning to read: A comparison of dyslexic and normal readers. Brain and Language, 87(3), 432–440. https://doi.org/10.1016/S0093-934X(03)00145-7

Grainger, J., & Jacobs, A. M. (1994). A dual read-out model of word context effects in letter perception: Further investigations of the word superiority effect. Journal of Experimental Psychology: Human Perception and Performance, 20(6), 1158–1176. https://doi.org/10.1037/0096-1523.20.6.1158

Gray, K. L. H., & Cook, R. (2018). Should developmental prosopagnosia, developmental body agnosia, and developmental object agnosia be considered independent neurodevelopmental conditions? *Cognitive Neuropsychology*, *35*(1–2), 59–62. https://doi.org/10.1080/02643294.2018.1433153

Griffin, J. W., Bauer, R., & Scherf, K. S. (2021). A quantitative meta-analysis of face recognition deficits in autism: 40 years of research. *Psychological Bulletin*, *147*(3), 268.

Hendel, R. K., Starrfelt, R., & Gerlach, C. (2019). The good, the bad, and the average: Characterizing the relationship between face and object processing across the face recognition spectrum. *Neuropsychologia*, *124*, 274–284. https://doi.org/10.1016/j.neuropsychologia.2018.11.016

Hervais-Adelman, A., Kumar, U., Mishra, R. K., Tripathi, V. N., Guleria, A., Singh, J. P., Eisner, F., & Huettig, F. (2019). Learning to read recycles visual cortical networks without destruction. *Science Advances*, *5*(9), eaax0262.

Hills, C. S., Pancaroglu, R., Duchaine, B., & Barton, J. J. (2015). Word and text processing in acquired prosopagnosia. *Annals of Neurology*, *78*(2), 258–271.

Holmes, D. R., & McKeever, W. F. (1979). Material specific serial memory deficit in adolescent dyslexics. *Cortex*, *15*(1), 51–62.

Huestegge, L., Rohrßen, J., van Ermingen-Marbach, M., Pape-Neumann, J., & Heim, S. (2014). Devil in the details? Developmental dyslexia and visual long-term memory for details. *Frontiers in Psychology*, *5*, 686.

Jozranjbar, B., Kristjansson, A., & Sigurdardottir, H. M. (2021). Featural and configural processing of faces and houses in matched dyslexic and typical readers. *Neuropsychologia*, 108059.

JASP Team. (2022). JASP (Version 0.16.1)[Computer software]. https://jasp-stats.org/

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*(8), 759.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627–633.

Kleinschmidt, A., & Cohen, L. (2006). The neural bases of prosopagnosia and pure alexia: Recent insights from functional neuroimaging. Current Opinion in Neurology, 19(4), 386–391. https://doi.org/10.1097/01.wco.0000236619.89710.ee

Kristjansson, A., & Sigurdardottir, H. M. (2022). *The role of visual factors in dyslexia* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/n8xer

Kühn, C. D., Gerlach, C., Andersen, K. B., Poulsen, M., & Starrfelt, R. (2020). Face recognition in developmental dyslexia: Evidence for dissociation between faces and words. *Cognitive Neuropsychology*, 1–9.

Le Grand, R., Cooper, P. A., Mondloch, C. J., Lewis, T. L., Sagiv, N., de Gelder, B., & Maurer, D. (2006). What aspects of face processing are impaired in developmental prosopagnosia? *Brain and Cognition*, *61*(2), 139–158.

Lieder, I., Adam, V., Frenkel, O., Jaffe-Dax, S., Sahani, M., & Ahissar, M. (2019). Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia. *Nature Neuroscience*, *22*(2), 256–264.

Liu, T. T., & Behrmann, M. (2014). Impaired holistic processing of left-right composite faces in congenital prosopagnosia. *Frontiers in Human Neuroscience*, *8*, 750.

Lobmaier, J. S., Bölte, J., Mast, F. W., & Dobel, C. (2010). Configural and featural processing in humans with congenital prosopagnosia. *Advances in Cognitive Psychology*, *6*, 23.

Maratos, F. A., Chu, K., Lipka, S., Stupple, E. J. N., & Parente, F. (2022). Exploring pattern recognition: What is the relationship between the recognition of words, faces and other objects? *Cognitive Processing*. https://doi.org/10.1007/s10339-022-01111-3

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260.

Mayseless, N., & Breznitz, Z. (2011). Brain activity during processing objects and pseudo-objects: Comparison between adult regular and dyslexic readers. *Clinical Neurophysiology*, *122*(2), 284–298.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences*, *11*(1), 8–15.

McKone, E., & Robbins, R. (2007). The evidence rejects the expertise hypothesis: Reply to Gauthier & Bukach. *Cognition*, *103*(2), 331–336.

Moeller, M. P., Tomblin, J. B., Yoshinaga-Itano, C., Connor, C. M., & Jerger, S. (2007). Current state of knowledge: Language and literacy of children with hearing impairment. *Ear and Hearing*, *28*(6), 740–753.

Moses, P., & Stiles, J. (2002). The lesion methodology: Contrasting views from adult and child studies. *Developmental Psychobiology*, *40*(3), 266–277. https://doi.org/10.1002/dev.10031

Nasr, S., Echavarria, C. E., & Tootell, R. B. (2014). Thinking outside the box: Rectilinear shapes selectively activate scene-selective cortex. *Journal of Neuroscience*, *34*(20), 6721–6735.

Norton, E. S., Beach, S. D., & Gabrieli, J. D. (2015). Neurobiology of dyslexia. *Current Opinion in Neurobiology*, *30*, 73–78.

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia*, *49*(5), 1226–1235.

Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature*, *423*(6941), 752.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, *331*(6157), 585–589.

Rakover, S. S. (2002). Featural vs. Configurational information in faces: A conceptual and empirical analysis. *British Journal of Psychology*, *93*(1), 1–30.

Reid, A. A. (2018). Neuroimaging reveals heterogeneous neural correlates of reading deficit in individuals with dyslexia consistent with a multiple deficit model. In *Neuroimaging-Structure, Function and Mind*. IntechOpen.

Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of Cognitive Neuroscience*, *16*(2), 189–203.

Richler, J. J., Tomarken, A. J., Sunday, M. A., Vickery, T. J., Ryan, K. F., Floyd, R. J., Sheinberg, D., Wong, A. C.-N., & Gauthier, I. (2019). Individual differences in object recognition. *Psychological Review*, *126*(2), 226–251. https://doi.org/10.1037/rev0000129

Ritchie, K. L., Flack, T. R., & Maréchal, L. (2023). Unfamiliar faces might as well be another species: Evidence from a face matching task with human and monkey faces. Visual Cognition, 1–6. https://doi.org/10.1080/13506285.2023.2184894

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, *103*(1), 34–79. https://doi.org/10.1016/j.cognition.2006.02.008

Roberts, D. J., Ralph, M. A. L., Kim, E., Tainturier, M.-J., Beeson, P. M., Rapcsak, S. Z., & Woollams, A. M. (2015). Processing deficits for familiar and novel faces in patients with left posterior fusiform lesions. *Cortex*, *72*, 79–96.

Robotham, J., & Starrfelt, R. (2017). Comparing word and face recognition: An insoluble conundrum. Journal of Vision, 17(10), 1002. https://doi.org/10.1167/17.10.1002

Robotham, R. J., & Starrfelt, R. (2017). Face and word recognition can be selectively affected by brain injury or developmental disorders. *Frontiers in Psychology*, *8*, 1547.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, *21*(2), 139–253.

Rossion, B., Dricot, L., Devolder, A., Bodart, J.-M., Crommelinck, M., Gelder, B. de, & Zoontjes, R. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *12*(5), 793–802.

Rubino, C., Corrow, S. L., Corrow, J. C., Duchaine, B., & Barton, J. J. (2016). Word and text processing in developmental prosopagnosia. *Cognitive Neuropsychology*, *33*(5–6), 315–328.

Rüsseler, J., Johannes, S., & Münte, T. F. (2003). Recognition memory for unfamiliar faces does not differ for adult normal and dyslexic readers: An event-related brain potential study. *Clinical Neurophysiology*, *114*(7), 1285–1291.

Ryan, K. F., & Gauthier, I. (2016). Gender differences in recognition of toy faces suggest a contribution of experience. *Vision Research*, *129*, 69–76. https://doi.org/10.1016/j.visres.2016.10.003

Sasaki, Y., Vanduffel, W., Knutsen, T., Tyler, C., & Tootell, R. (2005). Symmetry activates extrastriate visual cortex in human and nonhuman primates. *Proceedings of the National Academy of Sciences*, *102*(8), 3159–3163. https://doi.org/10.1073/pnas.0500319102

Shaywitz, S. E. (1998). Dyslexia. *New England Journal of Medicine*, *338*(5), 307–312.

Sigurdardottir, H. M., Arnardottir, A., & Halldorsdottir, E. T. (2021). Faces and words are both associated and dissociated as evidenced by visual problems in dyslexia. *Scientific Reports*, *11*(1), 1–15.

Sigurdardottir, H. M., Arnardottir, A., Halldorsdottir, E. T., Omarsdottir, H. R., & Valgeirsdottir, A. S. (2019). *Faces and words are both associated and dissociated: Evidence from visual problems in dyslexia*.

Sigurdardottir, H. M., Danielsdottir, H. B., Gudmundsdottir, M., Hjartarson, K. H., Thorarinsdottir, E. A., & Kristjánsson, Á. (2017). Problems with visual statistical learning in developmental dyslexia. *Scientific Reports*, *7*(1), 1–12.

Sigurdardottir, H. M., Fridriksdottir, L. E., Gudjonsdottir, S., & Kristjánsson, Á. (2018). Specific problems in visual cognition of dyslexic readers: Face discrimination deficits predict dyslexia over and above discrimination of scrambled faces and novel objects. *Cognition*, *175*, 157–168.

Sigurdardottir, H. M., Ívarsson, E., Kristinsdóttir, K., & Kristjánsson, Á. (2015). Impaired recognition of faces and objects in dyslexia: Evidence for ventral stream dysfunction? *Neuropsychology*, *29*(5), 739.

Smith-Spark, J. H., & Moore, V. (2009). The representation and processing of familiar faces in dyslexia: Differences in age of acquisition effects. *Dyslexia*, *15*(2), 129–146.

Snowling, M. (1998). Dyslexia as a phonological deficit: Evidence and implications. *Child Psychology and Psychiatry Review, 3*(1), 4-11.

Spoehr, K. T., & Smith, E. E. (1975). The role of orthographic and phonotactic rules in perceiving letter patterns. Journal of Experimental Psychology: Human Perception and Performance, 1(1), 21–34. https://doi.org/10.1037/0096-1523.1.1.21

Starrfelt, R., Klargaard, S. K., Petersen, A., & Gerlach, C. (2018). Reading in developmental prosopagnosia: Evidence for a dissociation between word and face recognition. *Neuropsychology*, *32*(2), 138.

Starrfelt, R., & Robotham, R. J. (2018). On the use of cognitive neuropsychological methods in developmental disorders. *Cognitive Neuropsychology*, *35*(1–2), 94–97. https://doi.org/10.1080/02643294.2017.1423048

Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., Pidcock, M., Dawel, A., Jeffery, L., & Wilson, C. E. (2010). Face recognition impairments despite normal holistic processing and face space coding: Evidence from a case of developmental prosopagnosia. *Cognitive Neuropsychology*, *27*(8), 636–664.

Tanaka, J. W., & Farah, M. J. (1993). Parts and Wholes in Face Recognition. *The Quarterly Journal of Experimental Psychology Section A*, *46*(2), 225–245. https://doi.org/10.1080/14640749308401045

Towler, J., Fisher, K., & Eimer, M. (2018). Holistic face perception is impaired in developmental prosopagnosia. *Cortex*, *108*, 112–126.

Tso, R. V., Chan, R. T., Chan, Y., & Lin, D. (2021). Holistic processing of Chinese characters in college students with dyslexia. *Scientific Reports*, *11*(1), 1–12.

Tso, R. V. Y., Chan, R. T. C., & Hsiao, J. H. (2020). Holistic but with reduced right-hemisphere involvement: The case of dyslexia in Chinese character recognition. *Psychonomic Bulletin & Review*, 1–10.

Ulrich, P. I., Wilkinson, D. T., Ferguson, H. J., Smith, L. J., Bindemann, M., Johnston, R. A., & Schmalzl, L. (2017). Perceptual and memorial contributions to developmental prosopagnosia. *Quarterly Journal of Experimental Psychology*, *70*(2), 298–315.

Valdois, S., Bosse, M., & Tainturier, M. (2004). The cognitive deficits responsible for developmental dyslexia: Review of evidence for a selective visual attentional disorder. *Dyslexia*, *10*(4), 339–363.

Van Gulick, A. E., McGugin, R. W., & Gauthier, I. (2016). Measuring nonvisual knowledge about object categories: The semantic Vanderbilt expertise test. *Behavior Research Methods*, *48*(3), 1178–1196.

Ventura, P., Bulajić, A., Wong, A. C.-N., Leite, I., Hermens, F., Pereira, A., & Lachmann, T. (2021). Face and word composite effects are similarly affected by priming of local and global processing. *Attention, Perception, & Psychophysics*, *83*(5), 2189–2204. https://doi.org/10.3758/s13414-021-02287-0

Ventura, P., Delgado, J., Ferreira, M., Farinha-Fernandes, A., Guerreiro, J. C., Faustino, B., Leite, I., & Wong, A. C.-N. (2019). Hemispheric asymmetry in holistic processing of words. *Laterality: Asymmetries of Body, Brain and Cognition*, *24*(1), 98–112.

Ventura, P., Fernandes, T., Leite, I., Almeida, V. B., Casqueiro, I., & Wong, A. C.-N. (2017). The Word Composite Effect Depends on Abstract Lexical Representations But

Not Surface Features Like Case and Font. *Frontiers in Psychology*, *8*, 1036. https://doi.org/10.3389/fpsyg.2017.01036

Ventura, P., Liu, T. T., Cruz, F., & Pereira, A. (2022). The mechanisms supporting holistic perception of words and faces are not independent. *Memory & Cognition*. https://doi.org/10.3758/s13421-022-01369-0

Järvinen, A., Korenberg, J. R., & Bellugi, U. (2013). The social phenotype of Williams syndrome. *Current Opinion in Neurobiology, 23*(3), 414-422.

Wong, A. C.-N., & Gauthier, I. (2007). An analysis of letter expertise in a levels-of-categorization framework. *Visual Cognition*, *15*(7), 854–879.

Wong, A. C.-N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects: Becoming a Ziggerin expert—But which type? *Psychological Science*, *20*(9), 1108–1117.

Wong, A. C.-N., Palmeri, T. J., Rogers, B. P., Gore, J. C., & Gauthier, I. (2009). Beyond shape: How you learn about objects affects how they are represented in visual cortex. *PloS One*, *4*(12), e8405.

Young, A. W., & Burton, A. M. (2018). Are We Face Experts? Trends in Cognitive Sciences, 22(2), 100–110. https://doi.org/10.1016/j.tics.2017.11.007

Young, A. W., Hellawell, D., & Hay, D. C. (2013). Configurational Information in Face Perception. *Perception*, *42*(11), 1166–1178. https://doi.org/10.1068/p160747n

Yovel, G., & Duchaine, B. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience*, *18*(4), 580–593.

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, *44*(5), 889–898.

Yue, X., Pourladian, I. S., Tootell, R. B., & Ungerleider, L. G. (2014). Curvature-processing network in macaque visual cortex. *Proceedings of the National Academy of Sciences*, *111*(33), E3467–E3475.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science*, *21*(4), 551–559.

Paper III

The impact of visual working memory constraints on object recognition

Bahareh Jozranjbar[1] (baj18@hi.is), Árni Kristjánsson[1], Christian Gerlach[2], Heida Maria Sigurdardottir[1]

[1] Icelandic Vision Lab
Department of Psychology
University of Iceland

[2] Department of Psychology
University of Southern Denmark
Odense, Denmark

Abstract

We explored the impact of visual working memory (VWM) constraints on processing of complex objects. The research design involved a VWM task where participants (N = 75) adjusted the orientation of a bar to match a previously viewed one. Additionally, they engaged in a delayed match-to-sample task involving faces, houses, and pseudowords, wherein individual features or feature configurations were manipulated. Our results showed a robust association between individual differences in VWM precision for simple stimuli and memory for configural information in houses, over and above that for featural and configural information in faces and pseudowords. As faces and words are expert categories while houses are not, the results are consistent with perceptual experience leading to decreased taxing of general VWM resources. We did not find a specific relationship between memory for featural information in houses, a non-expert category, and VWM precision for simple stimuli. Featural processing of non-expert categories may be less demanding on visual working memory than configural processing, as configural processing of non-expert categories could involve memory of both the features and relationships among features. In comparison, when individuals have extensive experience with an object category, they may create integrated chunks of information that facilitate rapid recognition and processing of the object as a whole. Such chunks in VWM consist of closely interconnected features, likely resulting from learned associations. Overall, we show that configural processing of non-expert categories may place more demands on visual working memory than both featural processing of such categories as well as both featural and configural processing of expert categories.

Keywords:

Visual Working Memory, Simple Objects, Real-World objects, Featural vs Configural Processing, Visual Expertise

Introduction

Visual working memory (VWM) involves the ability to briefly memorize visual information (Luck & Vogel, 1997b). VWM aids in maintaining perceptual continuity but has limited capacity (Alvarez & Cavanagh, 2004; Awh et al., 2007; Kristjánsson, 2006; Luck & Vogel, 1997; Vogel et al., 2001; Zhang & Luck, 2008). VWM research has primarily focused on simple objects with easily parametrized features, such as geometrical shapes and colours, while the effects of VWM limitations on more realistic stimuli, that are not simply combinations of basic features, are less clear.

Studies show mixed results regarding VWM differences between simple and real-world objects (Asp et al., 2021b; T. F. Brady et al., 2016b; Li et al., 2020b). Some suggest that real-world objects, which are visually complex, require more VWM resources than simpler stimuli (Luria et al., 2010; see also Alvarez & Cavanagh, 2004). Others propose that perception of meaningful stimulichanges VWM limits by envoking supplementary resources, such as the real size of object classes (Konkle & Oliva, 2012a; Long et al., 2016b, 2019b), expected nearby object classes (Kaiser et al., 2015b; O'Donnell et al., 2018b), and expertise with certain categories (Curby et al., 2009b; Curby & Gauthier, 2007b; Xie & Zhang, 2017c).

Familiarity's impact on VWM is also controversial. Some evidence suggests that VWM capacity increases with the presence of long-term memory representations (Jackson & Raymond, 2008b; Ngiam et al., 2019b). Scolari et al., (2008) and Lorenc et al., (2014) suggested that perceptual expertise and familiarity enhance stored information resolution but not slot number in VWM. In contrast, some studies have found no beneficial VWM effects for familiar objects (Chen et al., 2006b; Pashler, 1988b).

VWM may differ for real-world and simple objects due to varied processing methods. These include featural processing, where individual visual features are encoded, and configural processing, which involves encoding relationships between features (Maurer et al., 2002a). Simple visual features such as color and orientation can be stored independently and may be represented by different neurons or structures (B. R. Conway, 2009; Markov et al., 2019; Paik & Ringach, 2011; Wang et al., 2017). Simple object classes are therefore likely to be processed featurally. Complex features of real-world objects are largely integrated and may be encoded holistically in the medial temporal lobe and high-level ventral visual regions, as their representations do not merely constitute the sum of individual features or stimuli they comprise (Erez et al., 2016; van den Honert et al., 2017). The way in which features and configurations are processed for one real-world object category can however differ from how they are processed for another category (Jozranjbar et al., 2023).

Experts process their objects of expertise holistically (Bukach et al., 2006; Gauthier & Bukach, 2007; Gauthier & Tarr, 2002; but see McKone & Robbins, 2007). They can compress multiple features of familiar objects into chunks, enabling quick recognition and processing. Curby & Gauthier, (2007) found that faces are stored more efficiently in VWM than other complex objects. Holistic processing of faces may provide an advantage for face VWM, given sufficient encoding time. Curby et al., (2009) demonstrated that car experts have a similar VWM advantage for cars. These findings suggest that visual expertise can lead to domain-specific increases in VWM capacity due to holistic processing or chunking, enabling efficient feature binding and representation of complex object classes. Such chunks in VWM may be optimally compressed codes (Brady et al., 2009) or cues for retrieving information from long-term memory (Huang, 2011b; Hulme et al., 1991b, 1997b; Jones & Farrell, 2018b; Kahneman et al., 1992b). Both ideas assume that chunk contents must be retrieved from long-term memory. This may take more time but yet lead to an increased number of remembered features (Huang & Awh, 2018)

The distinction between featural and fully integrated representations is not clear, suggesting a potential continuum. Brady et al. (2013) found that various features of real-world object classes were forgotten at different rates, indicating that such objects are not always stored as fully bound units in visual memory. Markov et al. (2021) suggested that real-world object classes' features are represented somewhat independently in VWM as people assigned observed object attributes to incorrect locations or items. A visual working memory chunk could therefore be a collection of features that are highly – but not fully – integrated with one another due to learned associations. How integrated they are may depend on how diagnostic the features are of stimulus identity (Jackson & Raymond, 2008) and whether chunking is task-relevant. In tasks requiring feature individuation, individual features may need to be processed separately.

Current aims

The current study takes an individual differences approach to explore the relationship between VWM precision for simple stimuli and VWM accuracy for featural and configural information within unfamiliar houses, unfamiliar faces, and pseudowords. We estimate how memory accuracy for such objects is associated with – and is therefore likely explained by – VWM precision estimated independently for simple stimuli. We did this by comparing different predictive models to see which best predicted VWM precision. The study was preregistered on the Open Science Framework (https://osf.io/c58k2) where we proposed three hypotheses. First, it was expected that recognition accuracy for featural information in houses would be the best single predictor for VWM precision, since memorizing non-expert categories like houses may demand more VWM resources, particularly when individual features must be remembered as they tend to be encoded by more bits of information. For the same reasons, it was hypothesized that a model including house conditions, but excluding face and pseudoword conditions, would be the best overall model for predicting VWM precision. Finally, we predicted that the least effective model would have configural face condition as a predictor due to the holistic processing of faces.

Method

Participants

Participants were recruited through Prolific (prolific.co), an online research platform, where participants sign up for the sole purpose of participating in studies. The study was administered with Pavlovia (https://pavlovia.org). Participants took part in three sessions and received a total of £9.20 for their completion. As noted in the preregistration, we excluded participants with a mean bar adjustment error larger than 8 degrees on control trials (see Method: Visual working memory task). Moreover, one participant's accuracy in visual recognition task conditions (featural or configural face/house/word conditions) deviated from the average score of other participants by more than four, five, or even six standard deviations depending on the condition. This could potentially skew our overall results, threatening validity and reliability. Consequently, we diverged from our preregistration plan by excluding this participant to strengthen the conclusions from the study. This left a final sample of 75 participants (39 women, 34 men, 1 other; mean age: 35 years, range 18–72).

Procedure

The study protocol was reviewed by the Internal Review Board of Icelandic Universities (Siðanefnd háskólanna um vísindarannsóknir; ID SHV2021-001). Participants were recruited from a pool of people who had completed two sessions of a separate study (Jozranjbar et al., 2023 see: https://osf.io/a9wjy). Briefly, session 1 assessed how participants judged their ability to identify visual words and faces, and gathered information on their age, gender, and level of education. In session 2, people who did not report difficulties with recognizing faces or words in session 1 participated in a visual recognition task that required them to recall featural or configural information about unfamiliar faces, unfamiliar houses, and pseudowords (see Method: Visual recognition task). Participants in session two who did not fail attention checks (N = 101) received an invitation to take part and gave informed consent if they wished to participate in an additional session. In this session 3, participants performed a visual working memory task with simple stimuli. Participants used a laptop or desktop computer to participate in the study.

Visual working memory task

We evaluated VWM using a delayed report design. Each trial started with a sample array consisting of four oriented bars for 1000 ms; orientations were chosen at random from the whole range of possible orientations (0°—180°). The bars were positioned in such a way that they formed a square that fit within the screen. Following the sample array, a square visual mask covering all four bars was displayed for 1000 milliseconds, followed by a probe display consisting of the reappearance of one randomly selected bar from the sample display with a new randomly selected orientation. Participants adjusted the orientation of the bar by clicking on clockwise and counterclockwise arrow symbols until they thought that the bar matched the orientation of the equivalent bar in the preceding sample array that was in the same location (figure 20). There was no time limit on responses.

Participants initially completed a brief practice session consisting of ten trials, during which they received feedback by displaying the orientation of the bar that they were supposed to report and the orientation that they reported. After completing practice trials, each participant underwent 198 experimental trials. There was no feedback in the main experiment. Additionally, participants completed 22 control trials to determine their visual sensitivity and ability to stay on task. One control trial was run after every 9 experimental trials. On control trials, participants were presented with a line with a random orientation on the left (test bar) and a line with a random orientation on the right (adjustable bar), which they had to adjust until it matched the left line's orientation. The difference between the adjusted and test bars indicated the visual sensitivity as well as their attention to the task, with a smaller difference demonstrating greater sensitivity/attention.

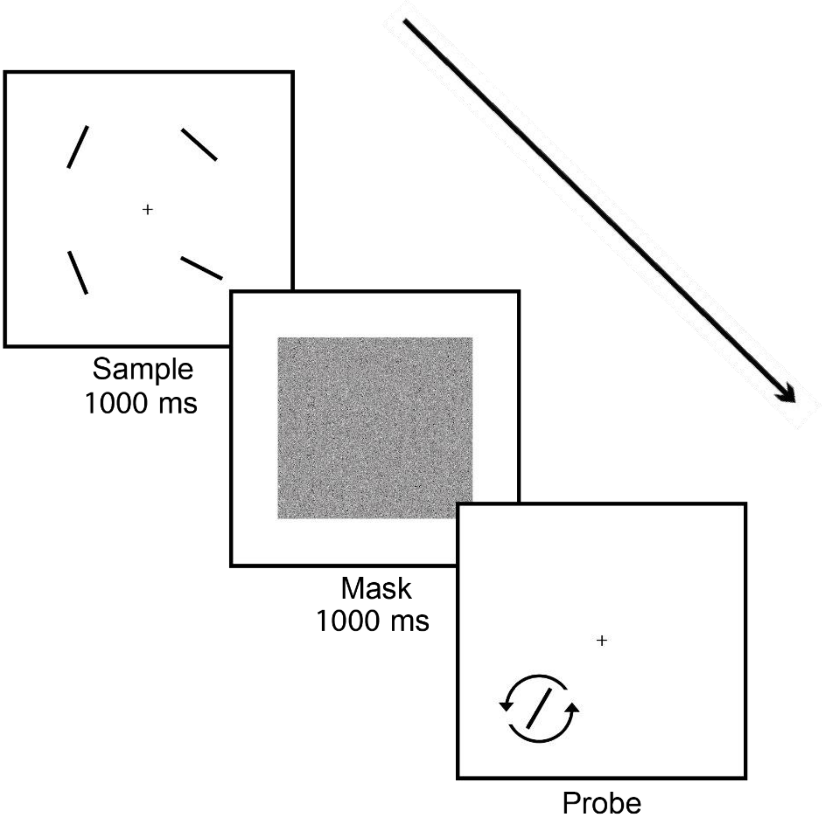


Figure 20. Visual working memory task. Four bars were displayed with random orientations for 1000 ms, followed by a 1000 ms mask. Subsequently, one of the items in the array was probed by location and the participants were asked to rotate it to indicate the remembered orientation. Arrows illustrate orientation adjustment and were not actually shown around the target.

Visual recognition task

The visual recognition task assessed memory for featural and configural information of faces, houses, and pseudowords. We reduced the impact of verbal cues by only using unfamiliar stimuli, including pseudowords where all manipulations kept letter identity and therefore pronunciation intact. Faces and houses were a part of a larger set made by Collins et al., (2012; we thank Jane E. Joseph for providing them). A five-letter word list (http://www.yougowords.com/5-letters) was fed to Wuggy (freely accessible at http://crr.ugent.be/programs-data/wuggy; see Keuleers & Brysbaert, 2010) to produce the set of pseudowords.

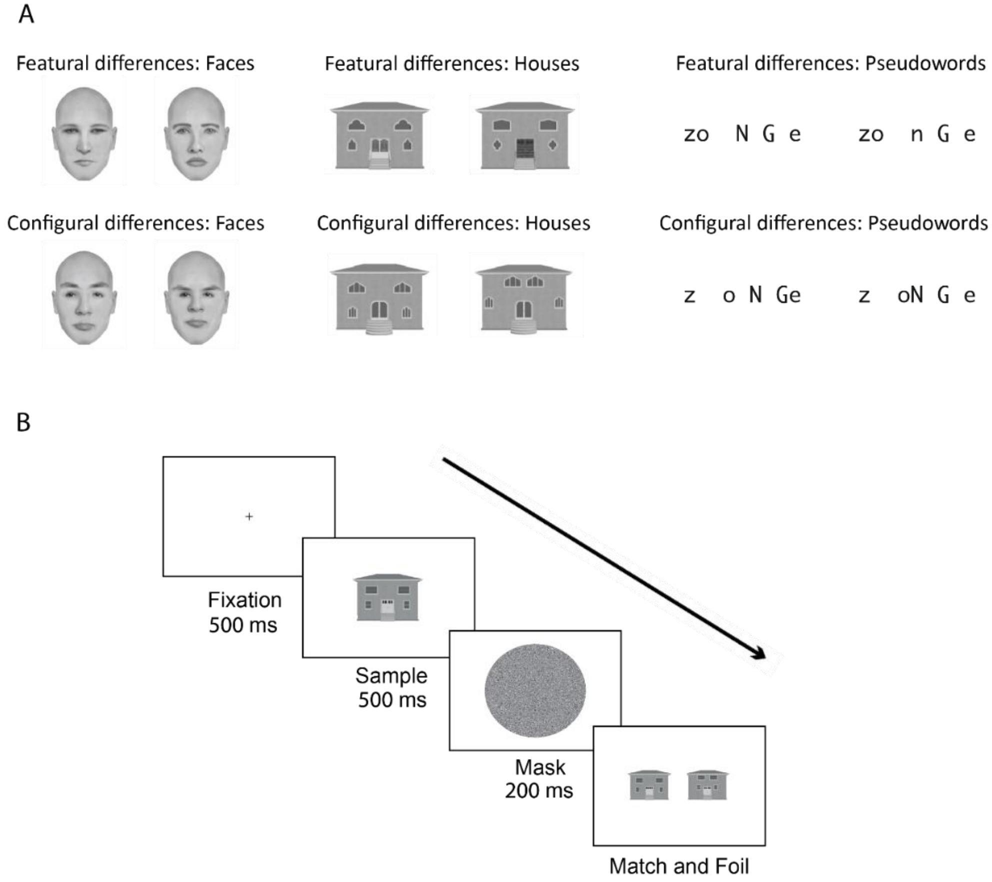


Figure 21. Visual recognition of featural and configural differences in faces, houses, and pseudowords. A. Examples of featural or configural changes of faces, houses, and pseudowords used in the visual recognition task. B. The experimental design. On each trial, a sample image (face, house, or pseudoword) appeared in the middle of the screen, followed by match and foil images concurrently presented to the left and right of the center. The match was identical to the sample image; however, the foil was either featurally or configurally different. Match and foil pictures remained onscreen until the participant identified which image was identical to the sample image by pressing a button.

The visual recognition task (figure 21) involved a two-alternative forced-choice delayed match-to-sample. A sample stimulus (face/house/pseudoword) was shown for 500 ms, and after a brief masked delay of 200 ms, a match and a foil stimulus were shown until the participants chose via keypress which stimulus they thought matched the sample. No feedback was given. The match was identical to the sample, but the foil had distinct featural or configural properties. Some features, such as noses in faces, doors in houses, and lower-uppercase letters in pseudowords, were varied between the sample and the foil for featural changes, and the distances between some features were altered for configural changes. Participants completed four blocks for each combination (featural face conditions, configural face conditions, featural house conditions, configural house conditions, featural pseudoword conditions, configural pseudoword conditions; figure 2, panel A), and each block had 24 trials (for details on this task, see Jozranjbar et al., (2023) and https://osf.io/a9wjy).

Results

Figure 22 shows the range of values for the six conditions of the visual recognition task and VWM precision. Different conditions of the visual recognition task may have slightly different levels of difficulty, but what is crucial for our current objectives is the substantial individual variability observed across all conditions.The precision of VWM was calculated as the reciprocal of the circular standard deviation of response error in experimental trials of the VWM task for oriented bars, where response error is the difference in degrees between the adjusted and test bars (Fisher, 1995). Precision measures response variability: the lower the response variability, the more precise the recall. We treated the visual working memory precision as a dependent variable and participants' accuracy in each of the six conditions of the visual object recognition task (memory for featural/configural information of unfamiliar faces/houses/pseudowords) as predictors in a Bayesian multiple regression. Note that we do this to estimate the degree of association between WWM and other variables and make no assumption about potential directions of causality by the choice of VWM as a dependent variable as opposed to an independent variable.

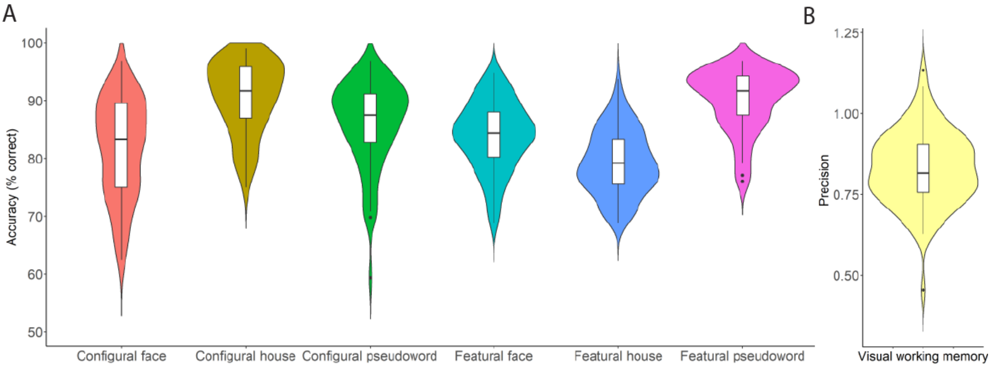


Figure 22. Distribution of task performance. A. Violin plot of total accuracies of the six conditions of the visual recognition task. B. Violin plot of VWM precision. The shape of each violin represents the density of the data. The box represents the middle 50% of the data (from the first quartile to the third quartile), with a line inside representing the median. The whiskers extend to the minimum and maximum values in the dataset, excluding any outliers which are plotted as individual points beyond the whiskers.

We ran Bayesian multiple regression analyses to identify which model, or which combination of models, best predicted the observed data. The Bayesian linear regression procedure begins by considering 64 models, including one null model which assumes no relationship between the predictor variables and VWM precision, six models in which only a single model (e.g., featural face accuracy) predicts VWM precision, and the remaining 57 models in which various combinations of models predict visual working memory precision. We analyzed the data with JASP (JASP Team, 2022). A .jasp file, including plots, data, and input options, is available at https://osf.io/m5h9u/. A Jeffreys-Zellner-Siow (JZS) prior and a beta binomial model prior were implemented. Table 7 shows that the 10 best models all contain configural house recognition accuracy as a predictor, and the configural house solo model was the best model of all, which suggests that this predictor is associated with precision of VWM. We hypothesized *a priori* that featural house recognition would be the best single predictor (Hypothesis 1) and that the best model would include featural house and configural house recognition as predictors (Hypothesis 2). Contrary to our first hypothesis and partially inconsistent with our second hypothesis, the best solo model and the best overall model for predicting VWM precision for simple lines was however the one with configural house accuracy as the only predictor. The worst solo model was the one with featural pseudowords as a predictor (see supplementary file for all models). This is also contrary to our third hypothesis, as we predicted that configural face conditions would be the worst predictor (Hypothesis 3). Bayes factor did not favour the null hypothesis of no association when configural face was a solo predictor (BF = 17.22). The results of all 64 models can be found in the ".jasp" file. Figure 4 depicts the relationships between VWM precision and configural house recognition accuracy.

Table 7. 10 best models from the Bayesian linear regression for the visual working memory and visual recognition task. P(M) are the prior model probabilities; P(M|data) are the posterior model probabilities; BFM is the change from prior to posterior model odds; BF10 shows the Bayes factors for each model. The Bayes factors were calculated relative to the best model. The first entry is always 1 since the best model is compared against itself. R^2 is the explained variance of each model.

Model Comparison - VWM precision

Models	P(M)	P(M|data)	BF_M	BF_{10}	R^2
Configural house	0.024	0.314	18.775	1.000	0.304
Featural house + Configural house	0.010	0.080	9.072	0.639	0.328
Configural face + Configural house	0.010	0.056	6.220	0.449	0.321
Featural pseudoword + Configural house	0.010	0.051	5.603	0.407	0.319

Model Comparison - VWM precision

Models	P(M)	P(M|data)	BF_M	BF_{10}	R^2
Featural face + Featural house + Featural pseudoword + Configural face + Configural house + Configural pseudoword	0.143	0.050	0.319	0.027	0.365
Featural house + Featural pseudoword + Configural house	0.007	0.043	6.303	0.460	0.350
Featural face + Configural house	0.010	0.035	3.754	0.277	0.311
Configural house + Configural pseudoword	0.010	0.029	3.142	0.233	0.307
Featural house + Featural pseudoword + Configural face + Configural house	0.010	0.026	2.791	0.208	0.360
Featural house + Featural pseudoword + Configural face + Configural house + Configural pseudoword	0.024	0.025	1.043	0.079	0.365

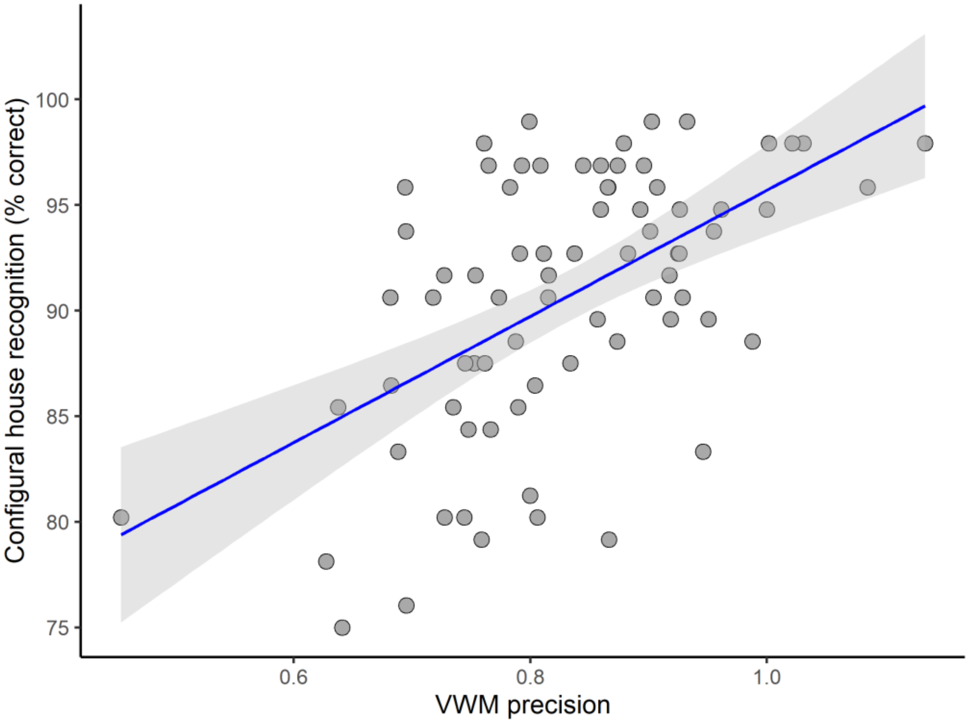


Figure 23. Scatter plot showing the relationship between VWM precision and configural house recognition accuracy. The line shows a linear fit (R2 = 0.30). The shaded areas are 95% confidence bands.

Given several candidate models, it is difficult to determine the importance of individual predictors. We therefore used Bayesian model averaging to determine the importance of individual predictors. Instead of estimating parameters using a single model, this algorithm weights the contribution of each model according to its posterior probability (Bergh et al., 2021). We used a Bayes factor classification method, which suggests anecdotal support for the hypothesis if the Bayes factor is between 1 and 3, moderate for 3 to 10, strong for 10 to 30, very strong for 30 to 100, and extreme for higher than 100 (Jeffreys, 1998). Table 2 provides a summary of the predictor inclusion probabilities and the posterior distributions averaged across all models.

Table 8. Model-averaged posterior summary for linear regression coefficients of the visual working memory and visual recognition task. The leftmost column denotes the predictor. P(incl|data) denotes the posterior inclusion probability. The change from prior to posterior inclusion odds is given by the inclusion Bayes factor (BFinclusion). The columns 'mean' and 'sd' represent the respective posterior mean and standard deviation of the parameter after model averaging. The columns 'lower and 'upper' indicate the 95%

Coefficient	P(incl|data)	$BF_{inclusion}$	Mean	SD	Lower	Upper
Intercept	1.000	1.000	0.830	0.011	0.806	0.851
Featural face	0.243	0.321	0.000	0.001	-0.002	0.003
Featural house	0.392	0.644	0.001	0.002	0.000	0.006
Featural pseudoword	0.359	0.560	-0.001	0.002	-0.007	0.000
Configural face	0.309	0.446	0.000	0.001	-0.001	0.004
Configural house	0.995	217.524	0.009	0.002	0.004	0.013
Configural pseudoword	0.251	0.335	0.000	0.001	-0.002	0.003

Table 8 confirms the conclusions from our initial analysis on the importance of configural house performance for predicting VWM precision, as indicated by the fact that the posterior inclusion probability is near 1 (0.995). The change from prior to posterior inclusion probabilities is visualized in the bar graph shown in Figure 24. In summary, the results of the Bayesian model-averaged analysis indicate strong support for the association between VWM precision for simple stimuli and configural house performance.

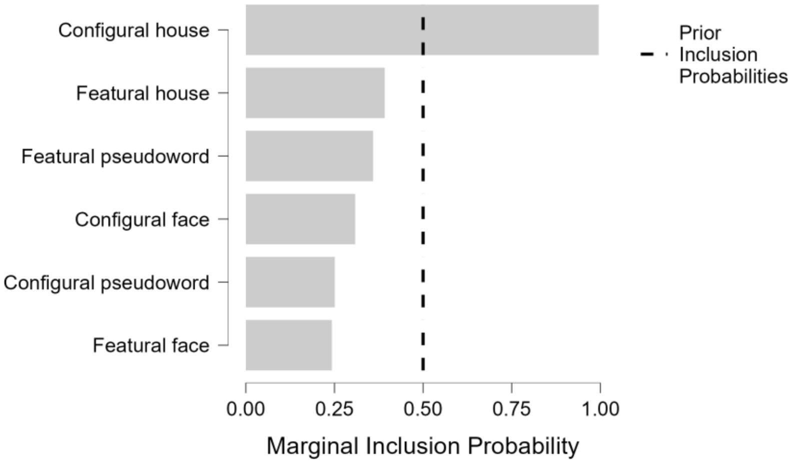


Figure 24. Bar graph of Posterior inclusion probabilities for the Bayesian linear regression. The dashed line represents the prior inclusion probabilities.

However, we found no evidence to suggest that models incorporating accuracy for featural faces, featural houses, or featural pseudowords, nor for configural face and configural pseudoword conditions, predict visual working memory precision for simple stimuli more effectively than models that do not include these predictors.

As illustrated in the network plot in figure 25, VWM precision is associated with memory for configural information for houses, but less with other stimulus categories.

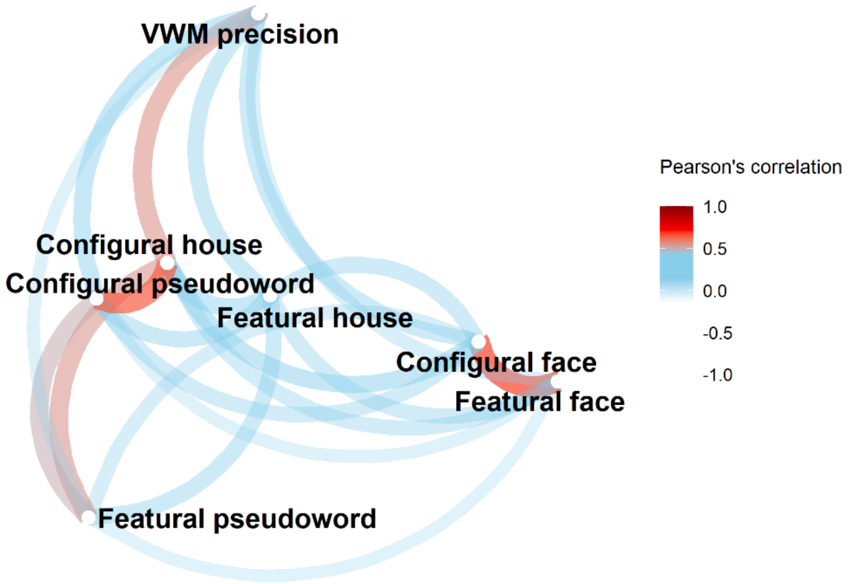


Figure 25. Network plot of recognition accuracy in conditions of the visual object recognition task and visual working memory precision for simple stimuli. The network plot shows the pattern in Pearson's correlations between the six object memory conditions and visual working memory precision. The graph's structure, obtained through multidimensional scaling (MDS), simplifies the high-dimensional correlation data for easier interpretation.

In an unregistered exploratory analysis, we utilized Bayesian regression to scrutinize the influence of visual sensitivity. Our goal was to assess its ability to explain data variability concerning VWM precision and configural house conditions, especially to rule out that the specific link between these two conditions could be explained by simple differences in perceptual abilities. First, we revisited our original Bayesian regression analysis, incorporating visual sensitivity as the seventh variable (figure 7A).

Subsequently, we examined configural house conditions as the dependent variable, with VWM precision and visual sensitivity serving as independent variables (figure 7B). This allowed us to gauge how well visual sensitivity and VWM factors explain individual differences in configural house trials. The results indicated that while visual sensitivity for oriented bars does seem to capture some of the variability in the data, it does not fully account for the link between VWM precision and memory for configural information in houses. The JASP file contains the Bayesian regression for the five remaining conditions of the visual recognition task.

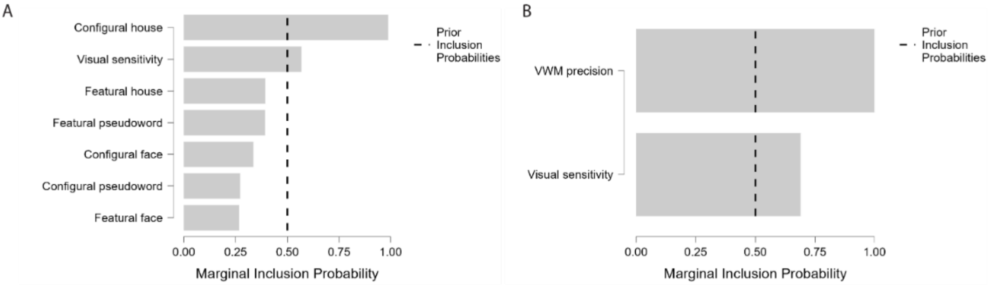


Figure 26. Bar graph of Posterior inclusion probabilities for the Bayesian linear regression. The dashed line represents the prior inclusion probabilities. A. Bayesian Linear Regression: Visual Working Memory Precision (Dependent Variable) vs. Visual Recognition Task Conditions and Visual Sensitivity (Independent Variables). B. Bayesian Linear Regression: Configural House Condition (Dependent Variable) vs. Visual Working Memory Precision and Visual Sensitivity (Independent Variables).

General Discussion

We investigated the association between VWM precision for simple features and memory accuracy for featural or configural information of different visual object categories (unfamiliar faces, unfamiliar houses, and pseudowords). While there may be ties between all conditions of the visual recognition task and VWM precision for simple stimuli (figure 6), our findings strongly support a stronger link between VWM precision for simple stimuli and memory accuracy for configural information in houses.

We hypothesized that VWM precision would be more associated with memory for object classes such as houses that people generally have limited experience with, especially if individual features need to be remembered as they are likely to require more bits of information for encoding. For expertise object classes such as faces and words, retaining less information might suffice due to established long-term memory representations, and memory for such object classes would therefore not be greatly associated with VWM for simple stimuli. When VWM stores object classes of expertise that have well-established representations in long-term memory, pointers may act as triggers for information recall from long-term memory (Huang, 2011b; Hulme et al., 1991b, 1997b; Jones & Farrell, 2018b; Kahneman et al., 1992b). Alternatively, these well-established representations could contribute to making the input more compressible (T. F. Brady et al., 2009b). In line with this general idea, memory accuracy for configural information in houses – a non-expert category – was associated with VWM precision over and above the latter's association with memory accuracy for object classes of expertise, faces and pseudowords (featural and configural conditions). This aligns with research that indicates enhanced VWM performance for expertise object classes through perceptual experience (Curby et al., 2009b; Curby & Gauthier, 2007b). Despite this, while houses represent a non-expertise category, we did not find a specific relationship between memory for featural information in houses and VWM precision for simple stimuli, suggesting that feature processing for non-expert categories may be less taxing on VWM than the processing of such object classes' configurations. This should be confirmed in future studies.

Our results suggest that when dealing with non-expert categories, the demands on VWM vary depending on information type. Featural processing involves encoding and remembering individual features independently, which may require fewer cognitive resources as each feature can be processed and stored separately. On the other hand, configural processing may require integrating multiple features and their spatial arrangements. This could tax VWM to a greater extent when remembering configurations for non-expert categories, as features *and* their relations may need to be encoded. Unitizing or chunking may be more effective for objects of expertise as experts often rely more on configural processing for their expertise classes (Bukach et al., 2006; Gauthier & Bukach, 2007b; Gauthier & Tarr, 2002; but see McKone & Robbins, 2007). When individuals possess extensive knowledge of a specific object or category, they can integrate chunks of information that enable them to rapidly recognize and process the object as a whole. Therefore, a VWM chunk could consist of features that are closely integrated with one another, possibly due to learned associations, and can be distinguished from less associated features. Consequently, the configural processing advantage may apply only to expert categories for which we have established feature associations.

It is unlikely that our results are due to the potential lower complexity of featural houses compared to configural houses, as average recognition accuracy for the former was lower than for the latter. Another possible explanation for the observed results could be the choice of VWM task. Using a different task, such as one involving colors instead of oriented lines, might have produced different outcomes, which could reflect the nature of configural processing, which focuses on spatial relationships. While orientation is a spatial dimension, color is a featural dimension that is less spatial. Configural processing includes aspects like distances and orientations, for example, the relative position of the windows to the door.

Visual sensitivity could possibly have played a role in our findings. This idea was hinted at in our exploratory analysis, where there was evidence of a relationship between VWM sensitivity and visual sensitivity, and between visual sensitivity and performance in the configural house condition. It is plausible that both VWM and visual sensitivity are tied to an individual's overall visual capabilities. A person with high visual acuity is likely able to capture detailed information, which could then be stored in VWM, providing some explanation for their shared variance. As for the second observed relationship, visual sensitivity could also be vital for recognizing or remembering houses based on their configuration. Presumably, those with higher visual sensitivity for orientations would be more adept at perceiving the fine details that make up the unique configuration of each house, which could lead to better performance, especially as configurations may be thought of as oriented information (e.g., windows are oriented 20 vs. 30 degrees from a door).

These results contribute to our understanding of how VWM operates in relation to different types of object categories and shed light on the varying cognitive demands associated with featural and configural processing. Further research can delve into the underlying mechanisms and neural processes involved in these different modes of processing, as well as explore the generalizability of these findings to other non-expert categories and cognitive tasks.

Conclusion

We investigated the association between VWM precision for simple stimuli and memory accuracy for featural or configural information of different visual object categories – faces, houses, and pseudowords. The findings suggest a strong link between VWM precision for simple stimuli and VWM accuracy for the non-expert category of houses, particularly when configural information needs to be encoded and remembered. The results suggest that configural processing of non-expert categories may place more demands on VWM than featural processing of such categories as well as both featural and configural processing of expert categories.

Acknowledgements

We would like to thank Jane E. Joseph for providing us with visual stimuli which were a part of a larger set developed by Collins et al. (2012). We are grateful to Paul Zerr for his help with coding and Randi Starrfelt for her feedback on this project. This work was supported by The Icelandic Research Fund (Grants No. 174013 and 228916) and the University of Iceland Research Fund.

References

Alvarez, G. A., & Cavanagh, P. (2004). The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects. *Psychological Science*, *15*(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Asp, I. E., Störmer, V. S., & Brady, T. F. (2021). Greater Visual Working Memory Capacity for Visually Matched Stimuli When They Are Perceived as Meaningful. *Journal of Cognitive Neuroscience*, *33*(5), 902–918. https://doi.org/10.1162/jocn_a_01693

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*(4), 487–502. https://doi.org/10.1037/a0016797

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2013). Real-world objects are not represented as bound units: Independent forgetting of different object details from visual memory. *Journal of Experimental Psychology: General*, *142*(3), 791–808. https://doi.org/10.1037/a0029649

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences*, *113*(27), 7459–7464. https://doi.org/10.1073/pnas.1520027113

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences*, *10*(4), 159–166.

Chen, D., Yee Eng, H., & Jiang, Y. (2006). Visual working memory for trained and novel polygons. *Visual Cognition*, *14*(1), 37–54. https://doi.org/10.1080/13506280544000282

Conway, B. R. (2009). Color Vision, Cones, and Color-Coding in the Cortex. *The Neuroscientist*, *15*(3), 274–290. https://doi.org/10.1177/1073858408331369

Curby, K. M., & Gauthier, I. (2007). A visual short-term memory advantage for faces. *Psychonomic Bulletin & Review*, *14*(4), 620–628.

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 94.

Erez, J., Cusack, R., Kendall, W., & Barense, M. D. (2016). Conjunctive Coding of Complex Object Features. *Cerebral Cortex*, *26*(5), 2271–2282. https://doi.org/10.1093/cercor/bhv081

Gauthier, I., & Bukach, C. (2007a). Should we reject the expertise hypothesis? *Cognition*, *103*(2), 322–330.

Gauthier, I., & Bukach, C. (2007b). Should we reject the expertise hypothesis? *Cognition*, *103*(2), 322–330.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 431.

Huang, L. (2011). Familiarity does not aid access to features. *Psychonomic Bulletin & Review*, *18*(2), 278–286. https://doi.org/10.3758/s13423-011-0052-6

Huang, L., & Awh, E. (2018). Chunking in working memory via content-free labels. *Scientific Reports*, *8*(1), 23. https://doi.org/10.1038/s41598-017-18157-5

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*(6), 685–701. https://doi.org/10.1016/0749-596X(91)90032-F

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), 1217–1232. https://doi.org/10.1037/0278-7393.23.5.1217

Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 556.

Jones, T., & Farrell, S. (2018). Does syntax bias serial order reconstruction of verbal short-term memory? *Journal of Memory and Language*, *100*, 98–122. https://doi.org/10.1016/j.jml.2018.02.001

Jozranjbar, B., Kristjansson, A., Starrfelt, R., Gerlach, C., & Sigurdardottir, H. M. (2022). *Using representational similarity analysis to reveal category and process specificity in visual object recognition* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/xh7kt

Jozranjbar, B., Kristjánsson, Á., Starrfelt, R., Gerlach, C., & Sigurdardottir, H. M. (2023). Using representational similarity analysis to reveal category and process specificity in visual object recognition. *Cortex*, S0010945223001387. https://doi.org/10.1016/j.cortex.2023.05.012

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219. https://doi.org/10.1016/0010-0285(92)90007-O

Kaiser, D., Stein, T., & Peelen, M. V. (2015). Real-world spatial regularities affect visual working memory for objects. *Psychonomic Bulletin & Review*, *22*(6), 1784–1790. https://doi.org/10.3758/s13423-015-0833-4

Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, *74*(6), 1114–1124.

Li, X., Xiong, Z., Theeuwes, J., & Wang, B. (2020). Visual memory benefits from prolonged encoding time regardless of stimulus type. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(10), 1998–2005. https://doi.org/10.1037/xlm0000847

Long, B., Konkle, T., Cohen, M. A., & Alvarez, G. A. (2016). Mid-level perceptual features distinguish objects of different real-world sizes. *Journal of Experimental Psychology: General*, *145*(1), 95–109. https://doi.org/10.1037/xge0000130

Long, B., Moher, M., Carey, S., & Konkle, T. (2019). Real-world size is automatically encoded in preschoolers' object representations. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(7), 863–876. https://doi.org/10.1037/xhp0000619

Lorenc, E. S., Pratte, M. S., Angeloni, C. F., & Tong, F. (2014). Expertise for upright faces improves the precision but not the capacity of visual working memory. *Attention, Perception, & Psychophysics*, *76*(7), 1975–1984. https://doi.org/10.3758/s13414-014-0653-z

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281.

Luria, R., Sessa, P., Gotler, A., Jolicœur, P., & Dell'Acqua, R. (2010). Visual Short-term Memory Capacity for Simple and Complex Objects. *Journal of Cognitive Neuroscience*, *22*(3), Article 3. https://doi.org/10.1162/jocn.2009.21214

Markov, Y. A., Tiurina, N. A., & Utochkin, I. S. (2019). Different features are stored independently in visual working memory but mediated by object-based representations. *Acta Psychologica*, *197*, 52–63. https://doi.org/10.1016/j.actpsy.2019.05.003

Markov, Y. A., Utochkin, I. S., & Brady, T. F. (2021). Real-world objects are not stored in holistic representations in visual working memory. *Journal of Vision*, *21*(3), 18. https://doi.org/10.1167/jov.21.3.18

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260.

McKone, E., & Robbins, R. (2007). The evidence rejects the expertise hypothesis: Reply to Gauthier & Bukach. *Cognition*, *103*(2), 331–336.

Ngiam, W. X. Q., Khaw, K. L. C., Holcombe, A. O., & Goodbourn, P. T. (2019). Visual working memory for letters varies with familiarity but not complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(10), 1761–1775. https://doi.org/10.1037/xlm0000682

O'Donnell, R. E., Clement, A., & Brockmole, J. R. (2018). Semantic and functional relationships among objects increase the capacity of visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(7), 1151–1158. https://doi.org/10.1037/xlm0000508

Paik, S.-B., & Ringach, D. L. (2011). Retinal origin of orientation maps in visual cortex. *Nature Neuroscience*, *14*(7), 919–925. https://doi.org/10.1038/nn.2824

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), 369–378. https://doi.org/10.3758/BF03210419

Scolari, M., Vogel, E. K., & Awh, E. (2008). Perceptual expertise enhances the resolution but not the number of representations in working memory. *Psychonomic Bulletin & Review*, *15*(1), 215–222. https://doi.org/10.3758/PBR.15.1.215

van den Honert, R. N., McCarthy, G., & Johnson, M. K. (2017). Holistic versus feature-based binding in the medial temporal lobe. *Cortex*, *91*, 56–66. https://doi.org/10.1016/j.cortex.2017.01.011

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92–114. https://doi.org/10.1037/0096-1523.27.1.92

Wang, B., Cao, X., Theeuwes, J., Olivers, C. N. L., & Wang, Z. (2017). Separate capacities for storing different features in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), 226–236. https://doi.org/10.1037/xlm0000295

Xie, W., & Zhang, W. (2017). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory & Cognition*, *45*(4), 677–689. https://doi.org/10.3758/s13421-016-0679-7

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235.