Exploring Pattern Recognition: What is the relationship between the recognition of words, faces and other objects?

Frances Maratos (f.maratos@derby.ac.uk)
University of Derby https://orcid.org/0000-0001-5738-6491
Kar-Man Chu
The Hong Kong Polytechnic University
Sigrid Lipka
University of Derby
Edward Stupple
University of Derby
Fabio Parente
University of Derby

Research Article

Keywords:

Posted Date: May 26th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-348237/v1

Exploring Pattern Recognition: What is the relationship between the recognition of words, faces and other objects?

Maratos, F.A.[1], Chu, K.[2], Lipka, S.[1], Stupple, E.J.N.[1] & Parente, F.[1]

[1]School of Psychology, College of Health, Psychology and Social Care, University of Derby

[2]Hong Kong Polytechnic University

Corresponding author:

Dr Frances A. Maratos
T: +44 (0) 1332 593053
E: f.maratos@derby.ac.uk

School of Psychology
College of Health, Psychology and Social Care
University of Derby,
Kedleston Road,
Derby
DE22 1GB
UK

ABSTRACT

Debate surrounds processes of visual recognition, with no consensus as to whether recognition of distinct object categories (faces, bodies, cars, and words) is domain specific or subserved by domain-general visual recognition mechanisms. Here, we investigated correlations between the performance of 74 participants on recognition tasks for words, faces and other object categories. Participants completed a counter-balanced test battery of the Cambridge Face, Car and Body Parts Memory tests, as well as a standard lexical decision task with response time (RT) and accuracy as dependent variables. We found that response times were consistently positively and significantly correlated between most object categories; for example, face recognition with body and car recognition, but also with word and non-word recognition. However, correlations between accuracy scores were more limited; for example, face recognition did not correlate with non-word recognition. Additionally, controlling for dyslexia resulted in some correlations becoming more robust (e.g., regular- and irregular-word recognition accuracy) or emerging (e.g., RT between body and real-word recognition), whilst others were partialled out (e.g., recognition accuracy between irregular words and confusable non-words). These results suggest some degree of functional overlap in the neural mechanisms subserving recognition of these different object categories, generally supporting a domain-general view of pattern recognition. This stated, observed relationships were complex. In order to further our understanding of pattern recognition, research investigating the recognition of words, faces and other objects in dyslexic individuals is recommended, as is research examining developmental trajectories of pattern recognition and research exploiting neuroimaging methodologies.

INTRODUCTION

Humans are unusual compared to many other species in their marked reliance on vision and visual experiences to interact with, understand, and even communicate the world around them (San Roque et al., 2015). It is therefore not surprising that a vast volume of research has attempted to investigate various aspects of our visual experience, from the neurobiology of low-level perception to higher-level object recognition and categorisation (for a review see Griffin & Motta-Mena, 2019). Much of the debate surrounding object recognition has centred around three potentially interrelated concepts: *functional specificity*, *anatomical stereotypy*, and *innateness*. In other words: i) Are recognition mechanisms different for different object domains? ii) If so, are these different mechanisms subserved by different brain regions that show anatomical consistency between individuals? iii) Does specificity and stereotypy develop under genetic control due to evolutionary selective pressures?

The idea that the brain is subdivided into functionally and anatomically distinct regions can be traced back to 19^{th}-century neuroanatomical investigations in patients with brain lesions (e.g., Broca, 1861; see Dronkers, Plaisant, Iba-Zizen and Cabanis, 2007 for a re-analysis of Paul Broca's original case studies). Slightly more recently, Konorski (1967; see also Fodor, 1983) proposed a model whereby populations of neurons (*gnostic sets*) sit at the top of a sensory hierarchy and are tuned to respond preferentially to combinations of low-level sensory features (from vision as well as other sensory modalities) that are diagnostic of very specific object categories or even category exemplars. However, alternative models of visual object recognition posited visual input can be used to construct representations that decompose objects into visual primitives (Biederman, 1986; Marr, 1982; see Figure 1). This would be achieved through general-purpose mechanisms applicable to any object in the visual field regardless of category.

*** Figure 1: About here ***

Importantly, modern neuroimaging has provided researchers with a means to directly investigate questions of functional specificity and anatomical consistency. For example, Malach and colleagues (1995) identified an occipital region, the lateral occipital complex (*LOC*), that preferentially responded to images of intact objects compared to scrambled images of the same objects, possibly demonstrating effects of object familiarity (see also Margalit et al., 2016), as well as category (Eger, Ashburner, Haynes, Dolan & Rees, 2008). Similarly, Sergent, Ohta and MacDonald (1992) identified an area of the fusiform gyrus, now termed the fusiform face area (*FFA*), which appeared to demonstrate selective responses to faces as compared to non-face objects. Kanwisher, McDermott and Chun (1997) observed this area to show stronger responses in the right hemisphere and argued, consistent with functional specificity, that it formed part of a larger network of face-processing regions including the occipital face area and the right anterior temporal lobe (see also Gauthier et al. 2000; Rajimehr, Young & Tootell, 2009). Almeida and colleagues (2020) have further described a case of a patient with a left splenium (posterior corpus callosum) lesion resulting in distorted processing of one side of human faces (hemi-prosopometamorphopsia or hemi-PMO), but not generalising to other classes of objects. Taken together, such studies support functional specificity for the identification of different object categories.

Indeed, shortly after the identification of the FFA, Cohen et al (2000) identified a region of the left fusiform gyrus showing selectivity for written words (Bolger, Perfetti, & Schneider, 2005; Cohen & Dehaene, 2004; Dehaene & Cohen, 2011; Dien, 2009), now known as the visual word form area (*VWFA*). Further studies have since identified additional occipitotemporal regions as having functional specialisation for: landmarks (e.g., Epstein, Harris, Stanley & Kanwisher, 1999), visual scenes (e.g., Kamps, Julian, Kubilius, Kanwisher & Dilks, 2016), and numbers (e.g., Hannagan, Amedi, Cohen, Dehaene-Lambertz & Dehaene, 2015). In sum, the functional selectivity and anatomical consistency observed within inferotemporal (IT) and occipital domains can be taken as evidence for innately programmed differential neural circuitry to process biologically important stimuli. For example, faces, given their importance to survival and bonding (e.g., McKone, Crookes, Jeffery & Dilks, 2012; Owen & Maratos, 2016).

This stated, a number of studies have challenged the notions of selectivity, consistency, and innateness on both theoretical and empirical grounds, putting forward arguments for functional and anatomical overlap, as well as experience in shaping neural organisation (e.g., Aguirre, Zarahn & D'Esposito, 1998; Maratos et al; 2007) or mechanisms (e.g., Beauchamp, Lee, Haxby & Martin, 2003; Chao, Weisberg & Martin, 2002; Downing, Chan, Peelen, Dodds & Kanwisher, 2006). To expand, the FFA has often been hailed as a clear example of functional and anatomical specialisation, and efforts have been made to trace its evolution by identifying homologous areas in social non-human primates, albeit with mixed results (Parr, 2011; Rossion & Taubert, 2019). However, a number of studies have shown that areas in the face network (including the FFA) appear to respond not only to faces, but also to other round objects (e.g., clocks or balls), potentially suggesting the existence of a common object recognition system based on primitive visual properties (Srihasam, Vincent & Livingstone, 2014; Tsao, Freiwald, Tootell & Livingstone, 2006; Yue, Pourladian, Tootell & Ungerleider, 2014) akin to earlier behavioural arguments of visual primitives (e.g., Marr, 1982).

Additionally, converging lines of research have provided evidence to challenge the idea of innate selectivity for faces. Studies in macaque (Livingstone et al., 2017) and human infants (Deen et al., 2017) found that face selectivity fully develops over several months after birth, and face expertise develops during early adolescence in humans (Hills & Lewis, 2018). Furthermore, face-stimulus deprivation in macaques prevents the formation of a face-processing area, which is then repurposed in a persistent fashion for general object recognition. Importantly, even when deprivation ceases, face-deprived macaques still show preferential gaze allocation towards hands rather than faces compared to controls (Arcaro, Schade, Vincent, Ponce & Livingstone, 2017). This is consistent with the Expertise Hypothesis (EH) put forward by Gauthier and colleagues (Gauthier, Tarr, Anderson, Skudlarski & Gore, 1999). According to EH, the FFA is part of a domain-general object recognition system (dependent on a domain-general visual ability) which is recruited for within-category object discriminations in domains of expertise. Under EH, face specialisation of the FFA gradually appears as infants/children develop more expertise in face processing and acquire familiarity with more and more individual faces to discriminate between. Consistent with this, using object categories in which expertise is acquired later (or as part of the experimental protocol itself), studies have shown that object recognition performance is more strongly correlated with face recognition performance, the higher the participant's expertise with those object categories (e.g., Bilalić, 2016; Bilalić, Grottenthaler, Nägele, & Lindig, 2016; Gauthier et al., 2014; Martens, Bulthé, van Vliet & de Beeck, 2018; see also Wang, Gauthier & Cottrell, 2016 for a modelling study using a biologically plausible neural network). In sum, these studies point to a shared object recognition mechanism guided by low-level visual properties, with the FFA as a *domain-general* expert discrimination system.

Much like the FFA, the domain-specificity and purpose of the VWFA has been challenged on multiple grounds, especially as the invention of written language is too recent in evolutionary terms for a dedicated 'written word recognition' region to have been selected for. Instead, a recycling hypothesis has been proposed (Dehaene & Cohen, 2007). Here, it is argued that cultural inventions such as reading take over evolutionarily older circuits that then become exclusively repurposed. For example, whilst the VWFA supports reading following literacy acquisition, it might also maintain other functions, such as the processing of groupings of visual stimuli characterised by high spatial frequencies and high complexity (Vogel, Petersen & Schlaggar, 2014; see also Price & Devlin, 2003). This would include, but not be limited to, groups of letter symbols and faces (Keil, Lapedriza, Masip & Vitria, 2008).

The possibility that the VWFA might be involved in the processing of different classes of visual objects therefore raises the possibility of associations between certain reading impairments and object recognition more generally. According with this, research has shown that individuals with developmental dyslexia not only demonstrate significant reading impairments, but also deficits in the rapid processing of object stimuli (Visser, Boden & Giaschi, 2004) and faces (Gabay, Dundas, Plaut & Behrmann, 2017; Sigurdardottir, Ívarsson, Kristinsdóttir & Kristjánsson, 2015). Additionally, Dundas, Plaut and Behrmann (2013) present evidence that the emergence of face lateralization is correlated with reading competence. As such, a number of studies have begun to explore the linkage/overlap between face and language processing more generally (e.g., Dundas, Plaut & Behrmann, 2014; Robinson, Plaut & Behrmann, 2017), as well as across patient populations (e.g., Asperud, Kuhn, Gerlach, Delfi & Starrfelt, 2019; Roberts et al., 2015). Asperud et al. (2019), for example, studied seven right-handed individuals with unilateral posterior cortex focal lesions and found evidence to support a bilaterally distributed network for the processing of words and faces. Importantly, however, they found that whilst in all cases word recognition deficits were paired with face deficits, the reverse was not true. Robinson et al. (2017) further observed face processing to interfere with word processing (but not vice-versa), suggesting a shared overlapping distributed network, especially for faces.

Arguably, therefore, object recognition would appear to rely on a broad network of functionally and anatomically overlapping brain regions, with many complex interactions between them, and whose behavioural correlates need to be better disentangled. Such research fits with the many-to-many theory of object recognition (MTMT; Behrmann & Plaut, 2013). In MTMT it is proposed that, while there may be brain regions that are ideally suited to the processing of certain properties of visual stimuli (see also Arcaro, Schade & Livingstone, 2019;

Jang, Song & Paik, 2020), these regions represent multiple object classes that share similar perceptual features.

Yet, a number of studies continue to find dissociations in recognition abilities between different object domains, indicating (at least partial) functional, and potentially neurological, dissociations, particularly in cases of visual agnosias (Behrmann, Moscovitch & Winocur, 1994; Farah, 1992; Moscovitch, Winocur & Behrmann, 1997; Starrfelt, Klargaard, Petersen & Gerlach, 2018; Susilo, Wright, Tree & Duchaine, 2015). It is also possible that the inconsistency between findings in the brain imaging literature may in part be due to insufficient spatial resolution or inherent between-subject variability in the size of the relevant regions of interest (Glezer & Riesenhuber, 2013). Furthermore, localisation of activation (or, indeed, lesions), particularly using proxy measures such as blood oxygen levels, may not on its own be sufficient to unambiguously identify function and predict behaviour (e.g., Van Horn et al., 2012). Indeed, the very premise of being able to directly infer mental states from brain activity is not without criticism both on theoretical (Poldrack, 2011) and methodological (Poldrack et al., 2017) grounds, particularly in the absence of a working hypothesis regarding the cognitive or behavioural phenomenon under investigation (Shallice & Cooper, 2011).

It is therefore important to further clarify whether performance in these different forms of visual pattern recognition is correlated at the behavioural level, using an appropriately powered sample and a variety of object classes. More specifically, in the present study we aimed to explore the relationship, if any, between the recognition of words, faces and other object categories in the general population. To this end we used the Cambridge Face, Car and Body Memory Tests (Dennett et al., 2012; Duchaine & Nakayama, 2006) as well as a standard lexical decision task as our measure of visual word recognition (Harley, 2014). In lexical decision tasks, participants need to decide if visually presented letter strings are known (English) words or made-up letter strings (i.e., non-words). Importantly, being able to process and store visual detail and visual units of various sizes is a key aspect of reading (Price, 2012). Thus, the lexical decision task allows investigation of detailed visual processing required for reading, including a participant's visual word recognition ability. This task therefore potentially constitutes a functional bridge linking word recognition to other types of visual processing such as that of faces and/or other object categories.

We hypothesized that if the recognition of words, faces, and other object categories (i.e., bodies and cars) share common neural underpinnings, then performance across tasks should be correlated, especially for the more familiar tasks (e.g., face and known-word recognition). However, if these different forms of pattern recognition are distinct dissociable ability domains, then performance across tasks would not be hypothesized to be correlated. Given emerging evidence of a relationship between reading, face processing and dyslexia, we also recorded whether participants had a dyslexia diagnosis (i.e., statement or equivalent).

METHOD

Design

We employed a correlational design. Sample size was calculated based on Pinel et al. (2014) and their non-word reading/FFA correlations. To obtain an expected correlation co-efficient of 0.32 between our key variables of (non-word) reading and face recognition with acceptable power (i.e., 0.8; with alpha set at 0.05), the calculated sample size required was 74.

Participants

Participants were recruited using opportunity sampling. Inclusion criteria included being a native English speaker and of age 18 or over. Consequently, 74 native English speakers were recruited and completed the study in full. 18 were male (24%) and all participants were either university students (undergraduate/ postgraduate) or staff (both academic and non-academic) from the University of Derby, United Kingdom. Participants ranged in age from

18.5 to 72.5 years, with an average age of 36.96 years. Ten participants (14%) reported having dyslexia (mean age = 30.6 years, non-dyslexia group mean age = 37.95 years). All participants gave informed written consent to participate in the study and received course credits (where applicable) or book tokens for their participation. The study received University of Derby Human Sciences Research Ethics Committee approval.

Materials

Visual and lexical recognition tasks

Cambridge Face Memory Test (CFMT), Duchaine & Nakayama (2006)

The CFMT measures participants' face recognition ability and has high internal reliability (Bowels et al., 2009). Items (and procedure) are the same as described in Duchaine and Nakayama (2006). Items consist of black and white photographs of neutral male faces taken in various poses (frontal viewing profile; 1/3 left profile; 1/3 right profile) and lighting conditions. ***Target faces*** consist of six male faces, with 12 images of each. ***Distractor faces*** consist of 46 different male images of the same age range, and taken in the same poses and lighting conditions as the target faces. All test items consist of a target face being presented alongside two distractor faces (randomly ordered), and below each a number (1, 2, 3) is presented. Scores can range from 0 to 72, with Duchaine and Nakayama (2006) reporting an average score of 57.9 in a sample of 50 University aged students.

Cambridge Car Memory Test (CCMT), Dennett et al. (2012)

The CCMT measures participants' ability in car recognition with high internal reliability (Dennett et al., 2012). The format and stages of the CCMT are the same as the CFMT described above, except that photographic stimuli are cars rather than faces. Whilst various types of car are used for the picture stimuli (e.g., sedans, sports cars, and wagons), they are presented in grayscale, with no identifying badges, logos and insignia. Scores can range from 0 to 72, with Dennett et al., reporting an average of 53.18 in a sample of 153 University aged students.

Cambridge Body Memory Test (CBMT), Duchaine & Nakayama (unpublished)

The CBMT measures participants' ability in body recognition and has the same format and stages as the CFMT and CBMT, with the exception that photographic stimuli are male bodies. Bodies are of various shapes and height, with only torso and limbs shown in grayscale. Scores can range from 0 to 72. To the authors' knowledge no descriptives are available for this task.

Lexical Decision Task (LDT)

The lexical decision task measures a participant's visual word recognition ability and is the most common behavioural method for investigating visual word recognition (Norris, 2013). Stimuli consisted of four trial types comprising two types of words and two types of made-up non-words. These were: 1) regularly spelt words (e.g., peach), which follow regular grapheme-to-phoneme correspondence rules; 2) irregularly spelt words (e.g., bough), which do not follow regular grapheme-to-phoneme correspondence rules and require visual memory of the correct orthography of the whole word (visual route); 3) pseudo-homophone non-words (e.g., floot), which are unfamiliar strings of letters that sound like real English words if grapheme-to-phoneme correspondence rules are applied (in this case: flute), and 4) visually confusable non-words (e.g., weeb; whem). This final category of non-words are unfamiliar strings of letters which do not sound like real English words but have high visual similarity to real words as they were created by either reversing letters that people with dyslexia often confuse (e.g., in weeb, the letter 'b' can be easily confused with 'd') or substituting letters with a visually similar letter (e.g., in whem, the letter 'm' has a high visual similarity with 'n'). Importantly, all non-words were created such that they were pronounceable and orthographically legal, in line with standard practice (Gabay et al., 2017).

In the current task, a total of 92 stimuli were used: 80 trial stimuli (20 per trial type), 8 practice stimuli and 4 warm-up stimuli. Number of letters was identical across stimulus type (average 4.5, range 3 to 6) and stimuli were matched on lexical frequency using the Elexicon megastudy norms (Balota et al., 2007), with an average of 5665.85 in the Hyperspace Analogue to Language (HAL) frequency norms and 7.72 in the log-transformed HAL frequency norms. Accuracy scores could range from: 0-20 per trial type; 0-40 per word type; 0-80 for overall task performance. Additionally, *correct answer* reaction time data was recorded.

Questionnaire measures

Waterloo Handedness Questionnaire-Revised (WHQ-R), Elias et al. (1998)

This questionnaire contains 36 items to measure participants' hand preference when executing various skilled (e.g., manipulation of objects) and less skilled (e.g., pick up small objects) activities. The test has high reliability (r= .88) (Steenhuis, Bryden, Schwartz & Lawson, 1990). Participants rate their hand preference according to a 5-point scale (Always left/ Usually left/ Equal/ Usually right/ Always right). Responses are scored by assigning values from -2 to 2 (Elias et al., 1998). Thus, total scores can range from -72 (strongly left-handed) to 72 (strongly right-handed).

Procedure

Participants were asked to sit at a distance of approximately 49.5 cm away from the computer screen with task presentation counter-balanced. That is, the face, car and body visual recognition tasks were counter-balanced; and either preceded or followed the LDT in 50% of cases, respectively. The Cambridge tasks were programmed using Java (https://www.java.com/), whereas the LDT was programmed using Inquisit (https://www.millisecond.com/).

In the CFMT, CCMT and CBMT participants were instructed to memorize the faces, cars or bodies that would follow. These tasks involved four phases, a Practice-, Introduction-, Novel- and Novel with Gaussian noise- phase, respectively. The latter three phases comprised 72 trials and the practice phase three trials. In a practice trial three study images of a target stimulus (a frontal viewing profile, a 1/3 left profile and a 1/3 right profile) were presented consecutively for three seconds each. Participants were then asked to discriminate the target stimulus (shown in one of the prior viewing profiles) from two distractor stimuli by pressing keyboard number '1', '2', or '3' that corresponded to the number below each stimulus image. Participants needed to score 100% (i.e., 3 out of 3 trials correct) before they could progress to the actual test. In the actual task, the 'Introduction' phase (comprising 18 trials with 6 targets x 3 presentations) had the same format as the practice trials. That is, participants viewed three images of a face, car or body part in three different orientations and then discriminated this target from two distractors (of the same category) with the same lighting and viewing profile. Following on from this, in the 'Novel' phase (comprising 30 trials with 6 targets x 5 presentations), six targets were presented simultaneously in a single review image from a frontal profile for twenty seconds. After which a participant was presented with three stimuli and asked to discriminate the target (i.e., one of the six prior stimuli) from two distractors stimuli. Importantly in this 'novel' version, distractor stimuli were novel images of the same category, but with different lighting and/or viewing profiles. Finally, in the 'Novel with Gaussian noise' phase (24 trials: 6 target faces x 4 presentations), trial events were the same as in the novel condition, but with the exception that all images were covered with Gaussian noise during the discrimination process to prevent ceiling effects.

In the LDT participants were asked to decide, as quickly and as accurately as possible, if a string of visually presented letters was an English word or not by pressing either the 'I' key on a computer keyboard (labelled Yes) or the 'E' key (labelled No). Stimuli remained on the screen until participants made a response. No feedback was provided. Stimuli were presented centrally on the screen using Arial font (12pt) in white colour against a grey background. Each stimulus was preceded by a central fixation cross presented for 250 ms. Participants started with eight practice trials to help them become familiar with the location of the response keys and trial format. In the experiment proper, there were 84 trials divided into 2 blocks. Each block of 42 trials included 2 warm-up trials and 40 experimental trials. These 40 trials consisted of 10 trials of each of the 4 stimulus types (i.e., regularly spelt words; irregularly spelt words; pseudo-homophone non-words; visually confusable non-words) presented pseudo-randomly. That is, trial order was randomised, but with the constraint that no two items of the same stimulus type were presented consecutively. A break was provided between the two blocks.

Following completion of the four visual recognition tasks, participants filled in the Waterloo Handedness Questionnaire-Revised, Elias et al., (1998) and provided demographic data pertaining to whether they had an official dyslexia diagnosis, their age, gender and occupation. In total, it took approximately 45 minutes to complete all aspects of the study.

Data Screening

Lexical Decision Task Accuracy

As the data were non-normally distributed, we used Friedman's test to establish differences in accuracy between the 4 types of lexical decision task (regularly spelt words median = 20.00, IQR=1; irregularly spelt words median = 19, IQR=3; pseudo-homophone non-words median = 19.5, IQR=1; visually confusable non-words median = 19.00, IQR=2). This revealed a main effect of trial type (χ^2=17.3, df=3, p=0.001). Bonferroni corrected pair-wise comparisons (threshold alpha level=.008) revealed differences reflected higher task accuracy on regularly compared with irregularly spelt words (p<.001), and pseudo-homophone non-words compared with visually confusable non-words (p=.001). No other comparisons were significant.

Lexical Decision Task Reaction Time

As the data were non-normally distributed, we used Friedman's test to establish differences in reaction time between the 4 types of lexical decision task (regularly spelt words median = 875.53ms, IQR=358.89ms; irregularly spelt words median = 833.55, IQR=327.92; pseudo-homophone non-words median = 1022.55, IQR=387.35; visually confusable non-words median = 1050.81, IQR=540.56). This revealed a main effect of trial type ($\chi 2$ =67.38, df=3, p<.001). Bonferroni corrected pair-wise comparisons (threshold alpha level=0.008) revealed significantly faster reaction times for regularly spelt words compared with i) pseudo-homophone non-words (p<.001), and ii) confusable non-words (p<.001), and significantly faster reaction times for irregularly spelt words compared with i) pseudo-homophone non-words (p<.001) and ii) visually confusable non-words (p<.001), in line with a typical lexicality effect. There were no significant differences between regularly and irregularly spelt words, nor between pseudo-homophone non-words and visually confusable non-words.

Handedness

Scores on the WHQ-R ranged from -53 to 72, with a mean of 43.22 (SD = 28.54), indicating a predominately right-handed sample.

Results

We hypothesized that if the recognition of words, faces, and other object categories share common neural underpinnings, then performance across tasks should be correlated. However, if these different forms of pattern recognition are distinct dissociable ability domains, then performance across tasks would not be hypothesised to be correlated. To investigate these differing hypotheses, we undertook two different types of exploratory correlations. Firstly, we explored recognition accuracy (in percentage correct) across tasks, then reaction time (for correct trials only) across tasks. Descriptive statistics pertaining to task performances per se are presented in Table 1.

Table 1 about here

1. Visual Recognition Accuracy Correlations Across Tasks

As screening checks for LDT accuracy revealed that the real words and the non-words did not group together in terms of accuracy, we entered all four lexical stimulus types into correlation analyses with accuracy for the three types of visual recognition: faces, body parts and cars. Non-parametric pairwise correlations revealed significant positive correlations between face and car recognition (r_s=.260, n=73, p=.026), car and body recognition (r_s=.243, n=74, p=.037), body recognition and regularly spelt words (r_s=.326, n=74, p=.005). Additionally, significant positive correlations were found between regularly and irregularly spelt words (r_s=.685, n=74, p<.001), between visually confusable non-words and irregularly spelt words (r_s=.249, n=74, p=.032), and between visually confusable and pseudo-homophone non-words (r_s=.685, n=74, p<.001). No other correlations met the threshold alpha level for significance (largest r_s =.223, p =.056).

When we controlled for dyslexia, significant partial correlations were found between car and body recognition (r=.240, n=71, p=.041), body recognition and regularly spelt words (r=.315, n = 71, p = .007), and between face recognition and irregularly spelt words (r=.260, n = 71, p = .027). In addition, significant positive partial correlations were found between regularly and irregularly spelt words (r=.706, n=71, p<.001), and between visually confusable and pseudo-homophone non-words (r=.788, n=71, p<.001). No other partial correlations met the threshold alpha level for significance (largest r=0.231). These partial correlations remained when controlling for handedness.

These results appear to suggest the existence of a relationship between the recognition of faces, words and other object categories, but also that this relationship is complex and potentially involves mediation/moderation effects of reading ability.

2. **Reaction Time Correlations Across Tasks**

As RTs for: i) the two types of real words and ii) the two types of non-words did not differ, for these analyses we combined both ‘real’ word types together (i.e., regularly and irregularly spelt words), and both non-word types together (i.e., pseudo-homophone and visually confusable non-words). Thus, two lexical stimulus type RTs alongside RTs for the three types of visual recognition, i.e., faces, body parts and cars, were entered into correlation analyses. This revealed significant positive correlations between face and body recognition (r_s=.716, n=73, p<.001), face and car recognition (r_s=.743, n=73, p<.001), body and car recognition (r_s=.783, n=74, p<.001), face recognition and non-word reaction times (r_s=.321, n=73, p=0.006), body recognition and non-word reaction times (r_s=.252, n=74, p=.030) and face recognition and real word reaction times (r_s=.251, n=73, p=.032). RTs for words and non-words were positively correlated (r_s=.652, n=74, p<.001). No other correlations met the threshold alpha level for significance (largest r_s=0.194).

When we controlled for dyslexia, the pattern of significant correlations reported above was replicated for the partial correlations: i.e. face and body recognition (r=.706, n=70, p<.001), face and car recognition (r=.788, n=70, p<.001), body and car recognition (r=.810, n=70, p<.001), face recognition and non-word reaction times (r=.346, n=70, p=.003), body recognition and non-words (r=.318, n=70, p=.006) and face recognition and real words (r=.284, n=70, p=.016). RTs for words and non-words remained significantly positively correlated (r=.690, n=71, p<.001). In addition, significant positive partial correlations between car recognition and non-word reaction times (r=.259, n=71, p=.027), as well as body recognition and real words reaction times (r=.267, n=70, p=.022) were observed. Only the partial correlation between car recognition and real-word recognition RTs was not significant (r=.207, n=71, p=.079). These correlations remained when controlling for handedness.

These full and partial RT correlation results broadly accord with the task accuracy correlations, revealing that RTs in the different tasks are significantly positively correlated, pointing to a relationship between them. However, unlike the accuracy performance, controlling for dyslexia statement did not change the pattern of results. Given the small size of the dyslexia subgroup in the current study (n=10), we are unable to conclusively opine whether the correlations between pattern recognition performance in the different domains was modulated by differences in reading ability. Nevertheless, see the Supplementary Materials for exploratory investigations of this possibility.

Discussion

Previous small sample behavioural and neuroimaging research has proven equivocal as to whether different forms of pattern recognition form distinct dissociable ability domains or share common neural underpinnings. Thus, the purpose of the present research was to explore relationships between body, face, word, and car recognition, through investigation of both accuracy and reaction time metrics. For the most part, we found that word recognition, face recognition and/or other object recognition (i.e., body and car) were significantly correlated, broadly supporting the theorem of common neural underpinnings. However, correlations differed in size from small to large, as discussed below.

In terms of response accuracy, a small but significant positive correlation was found between regular word recognition and body recognition accuracy, which remained significant when controlling for dyslexia status. However, only when controlling for dyslexia was a small, significant positive correlation between irregular word recognition and face recognition revealed. This suggests a functional overlap between these two abilities that dyslexia may interfere with. Further small, positive, significant correlations were also found between car and face recognition, and between car and body recognition. This stated, the former was no longer significant when partialling out dyslexia status. In general, these results appear to suggest that while functional overlap may exist between different object and word recognition domains, consistent with some research (e.g., Asperud et al., 2019; Dundas et al., 2013; Sigurdardottir et al., 2015), the patterns of overlap are likely complex. It is therefore possible that, for response accuracy, recognising an irregularly spelt word - i.e., a word for which the link between visual word form and phonological representation may be less entrenched – involves processing akin to that observed in face recognition. To expand, recognizing an irregularly spelt word depends on holistic pattern recognition processes: it cannot be achieved by applying rules linking individual letters to sounds, but requires the recognition of letter combinations or indeed the whole word as Gestalts or distinct patterns. This recognition of a pattern or Gestalt is potentially similar to configurational, holistic processes required in face recognition (e.g., Richler & Gauthier, 2014). However, given that a relationship between irregular word recognition and face recognition was only observed when controlling for dyslexia, we tentatively suggest that dyslexia may interfere with this shared mechanism, with such holistic pattern processing becoming dissociated. Conversely, recognising regularly spelt words can be achieved by piecemeal processing of individual letters and their regular mapping onto sounds, akin to processing individual features of (headless) bodies; with our results indicating this type of featural process does not seem to be affected by dyslexia status.

A further finding of the present study was that accuracy scores in the LDT were positively correlated between regular and irregular words, between irregular words and confusable non-words, and between confusable and pseudo-homophone non-words. However, the correlation between irregularly-spelt words and visually confusable non-words was no longer significant when controlling for dyslexia status, suggesting the latter may account for the shared variance between these two variables. This might indicate similarities in processing between real words and between non-words, but also that dyslexic participants may process visually confusable non-words by attempting to holistically build analogies between them and known irregular words. However, this possibility would require confirmation in a larger sample of this population.

Considering now the correlations between RTs (for correctly recognised words/faces/bodies/cars) within and across object categories, we found small but significant positive correlations, suggesting functional overlap between real-word recognition and face recognition, and between non-word recognition and face recognition (for full and partial correlations). When controlling for dyslexia, positive correlations were also found between real-word recognition and body recognition, and non-word recognition and car recognition. Response times within the LDT were significantly and positively correlated between words and non-words, and also when controlling for dyslexia status (see the Supplementary Analysis for a breakdown by word category). Similarly, response times across the object (face/body/car) recognition tasks were significantly, positively, and strongly correlated with each other, and these correlations also remained when controlling for dyslexia statement.

While the significant correlations between real-word recognition and face recognition RT and non-word recognition and face recognition RT are in partial agreement with the expertise hypothesis (Gauthier et al., 1999) and many-to-many theory of object recognition (MTMT; Behrmann & Plaut, 2013), it must be noted that correlations were more consistent between non-word recognition and face/body recognition (particularly when controlling for dyslexia). If, as per the expertise hypothesis, we might expect real words and faces to represent domains of relative expertise compared to non-words and bodies or cars, this pattern of results does not appear to necessarily support the idea that expertise is the dominant principle in establishing functional overlap. However, it must be pointed out that in this study we did not explicitly record participant category expertise. Furthermore, in our Supplementary Analysis, decomposing LDT response times produced a slightly clearer pattern of results, whereby the seemingly weaker correlations between real-word recognition and face recognition are partly due to the lack of a correlation between irregular-word and face recognition. Finally, processing non-words to reach a lexical decision arguably relies on the same expertise as processing real words (e.g., recognising letters and letter combinations, applying grapheme-phoneme correspondence rules, accessing the mental lexicon). Hence, the pattern of correlations reported here for non-words adds to understanding the relationship between visual processing of verbal and other object stimuli.

As discussed previously, our results are partly in agreement with Gauthier et al. (2014), but they also suggest that the visual processing of words is more like the visual processing of faces (and, to a lesser degree, bodies) than cars. However, it is also apparent that words and non-words are different in this respect. Similarly, the fact that accuracy (and to some extent RTs; see here supplementary Table's S2 and S4) was more strongly correlated for the two non-word categories and for the two word categories, than between words and non-words, suggests that these categories do indeed behave differently and recognising words compared with non-words involves partly distinct/differential processes, in line with established lexicality effects (e.g., Balota & Spieler, 1999).

Last but not least, it appears that dyslexia may be a useful test bed to investigate the presence and extent of functional overlap or dissociation between different domains of pattern recognition. More specifically, the results presented here (and in the Supplementary Materials) suggest that dyslexia accounted for a significant portion of the variance shared by car and face recognition accuracy, but not by regular word and body recognition accuracy. Given that the dyslexia group appeared to more substantially struggle with recognising irregular words (Supplementary Table S1), it is possible that these may be more functionally similar to cars and faces (with these three types of visual input all requiring more holistic processing) than to bodies, at least for this population. Future research should attempt to investigate these ideas with a sufficiently powered sample of individuals with dyslexia, so that individual differences in dyslexia presentation (e.g., phonological vs surface) may be more fully explored. In turn, a better understanding of dyslexia and of potentially comorbid deficits in both reading and object recognition domains could greatly inform future interventions and improve outcomes for such individuals.

To sum, whilst we set out to investigate if the recognition of words, faces and/or other object categories share common neural underpinnings, or if these different forms of pattern recognition represent distinct dissociable ability domains, the results presented provided a complex picture of domain interactions between words, faces and other object categories. For the most part, reaction times and accuracy correlated across task, which is partly consistent with proposals for domain-general object recognition mechanisms (e.g., Behrmann & Plaut, 2013; Gauthier et al., 2014). As such our findings lend behavioural support to neuroscientific research challenging domain-specificity of the FFA (e.g., Bilalić et al., 2016) or VWFA (e.g., Vogel et al., 2014). However, we did not find correlations between all recognition tasks (e.g., no correlation between car and real-word recognition), nor that correlations between stimuli were greater for those that are arguably better known (e.g., real words and faces compared with non-words and body parts). This hints that object recognition is potentially very complex involving both elements of domain-general and domain-specific processing; as also became evident from how controlling for dyslexia affected correlations differentially. In future research, in addition to more focused studies with individuals with dyslexia, any research in this area should also include measures of individuals' reading proficiency and familiarity with specific object categories, to further expand our understanding of the functional similarities in visual processing across a range of categories. The latter could also be explored by investigating developmental trajectories of word, face and other object recognition in typically developing children and participants/children with dyslexia. Finally, neuroimaging methods with high temporal resolution (such as magnetoencephalography) could be used to explore the time-course of visual processing of words, faces and other objects to investigate functional overlaps or otherwise.

Declarations

Funding
This work was partially supported by internal funding; a University of Derby (URSS067) award to Frances Maratos and Sigrid Lipka.

Conflicts of Interest
The authors have no relevant financial or non-financial interests to disclose.

Ethics Approval
All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Ethical approval was granted by the University of Derby Psychology Research Ethics Committee.

Consent to Participate
Informed consent was obtained from all individual participants included in the study.

Consent to Publish
Participants consented to the findings being published in journal articles. Figure 1 was reproduced with permission from the copyright holder.

Availability of Data and Materials – Open Practices
Raw data will be made available upon request. None of the experiments was pre-registered.

Author's Contributions
The study was incepted by FM, who oversaw all elements. SL, ES and KC contributed to study design, with SL designing the lexical decision task. KC collected all data, and analyses were progressed by FM, ES & FP. The first draft of the paper was progressed by FM, KC& FP, with SL and ES providing input/revision suggestions. All authors approved the final version of the manuscript for submission.

References

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron, 21*(2), 373-383.

Almeida, J., Freixo, A., Tábuas-Pereira, M., Herald, S. B., Valério, D., Schu, G., ... & Santana, I. (2020). Face-specific perceptual distortions reveal a view-and orientation-independent face template. *Current Biology, 30*(20), 4071-4077.

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.

Asperud, J., Kühn, C. D., Gerlach, C., Delfi, T. S., & Starrfelt, R. (2019). Word recognition and face recognition following posterior cerebral artery stroke: Overlapping networks and selective contributions. *Visual Cognition, 27*(1), 52-65.

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General*, 128(1), 32.

Behrmann, M., Moscovitch, M., & Winocur, G. (1994). Intact visual imagery and impaired visual perception in a patient with visual agnosia. *Journal of Experimental Psychology: Human Perception and Performance, 20*(5), 1068.

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

Bilalić, M. (2016). Revisiting the role of the fusiform face area in expertise. *Journal of Cognitive Neuroscience, 28*(9), 1345-1357.

Bilalić, M., Grottenthaler, T., Nägele, T., & Lindig, T. (2016). The faces in radiological images: Fusiform face area supports radiological expertise. *Cerebral Cortex, 26*(3), 1004-1014.

Biederman, I. (1986). Recognition by components: a theory of visual pattern recognition. *Psychology of Learning and Motivation*, *20*, 1-54.

Broca, P. (1861). Remarks on the Seat of the Faculty of Articulated Language, Following an Observation of Aphemia (Loss of Speech). *Bulletin de la Société Anatomique, 6*, 330-357.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). FMRI responses to video and point-light displays of moving humans and manipulable objects. *Journal of Cognitive Neuroscience, 15*(7), 991-1001.

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

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., & Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123*(2), 291-307.

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: the case for the visual word form area. *Neuroimage, 22*(1), 466-476.

Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., ... & Saxe, R. (2017). Organization of high-level visual cortex in human infants. *Nature Communications, 8*(1), 1-10.

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.

Dejerine J. (1892). Contribution à l'étude anatomo-pathologique et clinique des différentes variétés de cécité verbale. *Mémoires de la Société de Biologie, 4*, 61–90.

Demirci, M. F., Shokoufandeh, A., Keselman, Y., Bretzner, L., & Dickinson, S. (2006). Object recognition as many-to-many feature matching. *International Journal of Computer Vision, 69*(2), 203-222.

Dennett, H. W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Ewards, 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. *Behaviour Research Methods*, *44*, 587-605.

Devlin, J. T., Jamison, H. L., Gonnerman, L. M., & Matthews, P. M. (2006). The role of the posterior fusiform gyrus in reading. *Journal of Cognitive Neuroscience*, *18*(6), 911-922.

Dien, J. (2009). A tale of two recognition systems: Implications of the fusiform face area and the visual word form area for lateralized object recognition models. *Neuropsychologia*, *47*(1), 1–16.

Downing, P. E., Chan, A. Y., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2006). Domain specificity in visual cortex. *Cerebral Cortex*, *16*(10), 1453-1461.

Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: high resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, *130*(5), 1432-1441.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44*, 576-585.

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.

Dundas, E. M., Plaut, D. C., & Behrmann, M. (2014). An ERP investigation of the co-development of hemispheric lateralization of face and word recognition. *Neuropsychologia*, *61*, 315-323.

Eger, E., Ashburner, J., Haynes, J. D., Dolan, R. J., & Rees, G. (2008). fMRI activity patterns in human LOC carry information about object exemplars within category. *Journal of Cognitive Neuroscience*, *20*(2), 356-370.

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding?. *Neuron*, *23*(1), 115-125.

Farah, M. J. (1992). Is an Object an Object an Object? Cognitive and Neuropsychological Investigations of Domain Specificity in Visual Object Recognition. *Current Directions in Psychological Science*, *1*(5), 164–169.

Fodor, J. A. (1983). *The modularity of mind*. MIT press.

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., 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., 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), 568-573.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*(3), 495-504.

Glezer, L. S., and Riesenhuber, M. (2013) "Individual variability in location impacts orthographic selectivity in the "visual word form area"." *Journal of Neuroscience*. *33*(27), 11221-11226.

Griffin, J. W., & Motta-Mena, N. V. (2019). Face and object recognition. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.). *Encyclopedia of Evolutionary Psychological Science.* Springer, Cham.

Hannagan, T., Amedi, A., Cohen, L., Dehaene-Lambertz, G., & Dehaene, S. (2015).

Origins of the specialization for letters and numbers in ventral occipitotemporal cortex. *Trends in Cognitive Sciences*, *19*(7), 374-382.

Harley, T. A. (2014). *The Psychology of Language.* 4th ed. Hove: Psychology Press.

Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., ... & Rapoport, S. I. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *88*(5), 1621-1625.

Hills, P. J., & Lewis, M. B. (2018). The development of face expertise: Evidence for a qualitative change in processing. *Cognitive Development*, 48, 1-18.

Hirshorn, E. A., Li, Y., Ward, M. J., Richardson, R. M., Fiez, J. A., & Ghuman, A. S. (2016). Decoding and disrupting left midfusiform gyrus activity during word reading. *Proceedings of the National Academy of Sciences*, *113*(29), 8162-8167.

Hruby, G.G. & Goswami, U. (2011). Neuroscience and reading: A review for reading education researchers. *Reading Research Quarterly, 46*(2), 156–172.

Jang, J., Song, M., & Paik, S. B. (2019). Classification of columnar and salt-and-pepper organization in mammalian visual cortex. *Cell Reports*, *30*(10), 3270-3279.

Keil, M. S., Lapedriza, A., Masip, D., & Vitria, J. (2008). Preferred spatial frequencies for human face processing are associated with optimal class discrimination in the machine. *PloS One*, *3*(7).

Kamps, F. S., Julian, J. B., Kubilius, J., Kanwisher, N., & Dilks, D. D. (2016). The occipital place area represents the local elements of scenes. *Neuroimage*, *132*, 417-424.

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.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. *Neuroimage*, *21*(3), 946-953.

Livingstone, M. S., Vincent, J. L., Arcaro, M. J., Srihasam, K., Schade, P. F., & Savage, T. (2017). Development of the macaque face-patch system. *Nature Communications*, *8*(1), 1-12.

Malach R, Reppas J. B, Benson R. R, Kwong K. K, Jiang H, Kennedy W. A, Tootell R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences*, *92*(18), 8135–8139.

Margalit, E., Shah, M. P., Tjan, B. S., Biederman, I., Keller, B., & Brenner, R. (2016).

The Lateral Occipital Complex shows no net response to object familiarity. *Journal of vision*, *16*(11), 3-3.

Maratos F.A., Anderson S.J., Hillebrand A., Singh K.D & Barnes G. (2007). The spatial distribution and temporal dynamics of brain regions activated during the perception of object and non-object patterns. *Neuroimage*. 34, 371-383

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman, San Francisco.

Martens, F., Bulthé, J., van Vliet, C., & de Beeck, H. O. (2018). Domain-general and domain-specific neural changes underlying visual expertise. NeuroImage, *169*, 80-93.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293-299.

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.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, *9*(5), 555-604.

Owen, S., & Maratos, F. A. (2016). Recognition of subtle and universal facial expressions in a community□based sample of adults classified with intellectual disability. *Journal of Intellectual Disability Research*, *60*(4), 344-354.

Parr, L. A. (2011). The evolution of face processing in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1764-1777.

Poldrack, R. A. (2011). Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*, *72*(5), 692-697.

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., ... & Yarkoni, T. (2017). Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, *18*(2), 115.

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, *62*(2), 816-847.

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

Rajimehr, R., Young, J. C., & Tootell, R. B. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences*, *106*(6), 1995-2000.

Rauschecker, A. M., Bowen, R. F., Parvizi, J., & Wandell, B. A. (2012). Position sensitivity in the visual word form area. *Proceedings of the National Academy of Sciences*, *109*(24), E1568-E1577.

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, *140*(5), 1281-1302.

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.

Robinson, A. K., Plaut, D. C., & Behrmann, M. (2017). Word and face processing engage overlapping distributed networks: Evidence from RSVP and EEG investigations. *Journal of Experimental Psychology: General*, *146*(7), 943.

Rossion, B., & Taubert, J. (2019). What can we learn about human individual face recognition from experimental studies in monkeys?. *Vision Research*, *157*, 142-158.

San Roque, L., Kendrick, K. H., Norcliffe, E., Brown, P., Defina, R., Dingemanse, M., ... & Rossi, G. (2015). Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cognitive Linguistics*, *26*(1), 31-60.

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing: a positron emission tomography study. *Brain*, *115*(1), 15-36.

Shallice, T., & Cooper, R. (2011). *The Organisation of Mind*. (pp. 68-105). Oxford University Press.

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.

Srihasam, K., Vincent, J. L., & Livingstone, M. S. (2014). Novel domain formation reveals proto-architecture in inferotemporal cortex. *Nature Neuroscience*, *17*(12), 1776.

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.

Susilo, T., Wright, V., Tree, J. J., & Duchaine, B. (2015). Acquired prosopagnosia without word recognition deficits. *Cognitive Neuropsychology*, *32*(6), 321-339.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, *311*(5761), 670-674.

Van Horn, J. D., Irimia, A., Torgerson, C. M., Chambers, M. C., Kikinis, R., & Toga, A. W. (2012). Mapping connectivity damage in the case of Phineas Gage. *PloS One*, *7*(5).

Visser, T. A., Boden, C., & Giaschi, D. E. (2004). Children with dyslexia: evidence for visual attention deficits in perception of rapid sequences of objects. *Vision Research*, *44*(21), 2521-2535.

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.

Wernicke, C. (1970). The aphasic symptom-complex: a psychological study on an anatomical basis (1875). *Archives of Neurology*, *22*(3), 280-282.

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.

Figure 1: Example of how the visual recognition of complex objects (e.g., a human body) could be performed through decomposition into more basic visual primitives (i.e., volumes and shapes), a mechanism that could be domain-general. Figure reproduced with permission from Marr (1982).

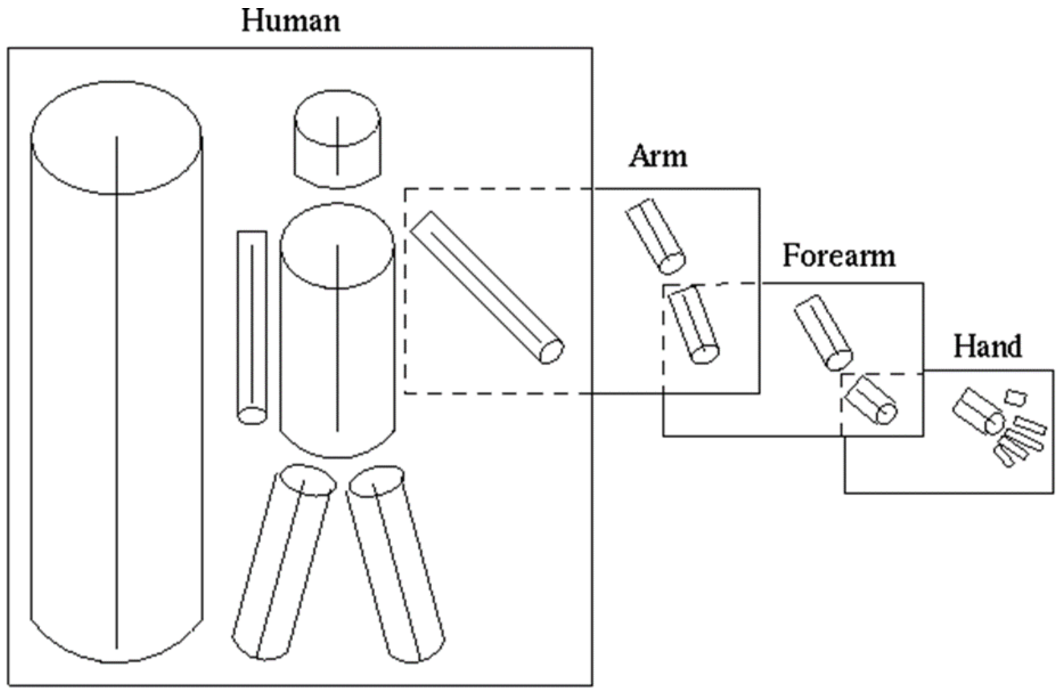

Table 1. *Mean (Standard Deviation) accuracy and reaction time for correct trials in Lexical Decision and Visual Recognition Tasks.*

		Accuracy (%)	Reaction Time (ms)
Lexical Decision Tasks			
	Regularly spelt words	95.45 (7.16)	938.27 (278.62)
	Irregularly spelt words	92.45 (9.49)	922.56 (286.62)
	Pseudo-homophone non-words	94.45 (8.91)	1164.55 (494.71)
	Visually confusable non-words	91.75 (12.74)	1171.59 (470.63)
Visual Recognition Tasks			
	Face	79.01 (15.86)	3461.80 (1024.96)
	Body	65.93 (11.35)	4308.98 (1455.05)
	Car	71.75 (17.07)	4932.80 (1502.43)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterialsFinal19.03.21.docx