The mechanisms supporting holistic perception of words and faces are not independent

Paulo Ventura[1] · Tina T. Liu[2] · Francisco Cruz[1] · Alexandre Pereira[3]

Abstract

The question of whether word and face recognition rely on overlapping or dissociable neural and cognitive mechanisms received considerable attention in the literature. In the present work, we presented words (aligned or misaligned) superimposed on faces (aligned or misaligned) and tested the interference from the unattended stimulus category on holistic processing of the attended category. In Experiment 1, we found that holistic face processing is reduced when a face was overlaid with an unattended, aligned word (processed holistically). In Experiment 2, we found a similar reduction of holistic processing for words when a word was superimposed on an unattended, aligned face (processed holistically). This reciprocal interference effect indicates a trade-off in holistic processing of the two stimuli, consistent with the idea that word and face recognition may rely on non-independent, overlapping mechanisms.

Introduction

Faces are generally considered under a perceptual expertise framework. In recent years, increased attention has also been devoted to visual word recognition under a perceptual expertise framework (e.g., Liu et al., 2016; Ventura, 2014; Wong & Gauthier, 2007). To unravel the characteristics of perceptual expertise that are either general or specific to various visual categories (e.g., Busey & Vanderkolk, 2005; Richler et al., 2011; Wong & Gauthier, 2007; Xu, 2005), studies on perceptual expertise have often investigated face recognition (e.g., Farah et al., 1998; Maurer et al., 2002) or the comparison between faces and non-face categories of expertise. It has been proposed that, rather than visual properties of stimuli (regardless of their resemblance to faces or not), it is the previous intense and continuous experience and task demands (i.e., individuation and fast processing of items composed by highly similar local elements) that drive perceptual expertise (e.g., Baker et al., 2002; Harel, 2016; Wong et al., 2012b).

Holistic processing has been regarded as one of the mechanisms underpinning the ability of the visual system to fulfill the task demand of fast individuation (Diamond & Carey, 1986; Gauthier & Bukach, 2007; Gauthier et al., 2003; Rossion, 2013; Wong et al., 2009; Young et al., 1987). Holistic processing can be loosely defined as the tendency to perceptually integrate parts into a unitary whole (Diamond & Carey, 1986; Gauthier & Bukach, 2007; Gauthier et al., 2003; Rossion, 2013; Wong et al., 2009; Young et al., 1987). Holistic processing has had different definitions in the literature. Three of the most studied definitions involve the representation of faces as undifferentiated wholes (e.g., Young et al., 1987; Farah et al., 1998; Maurer et al., 2002), a perceptual strategy of processing all parts together that becomes automatized with experience and/or due to a history of learned attention to diagnostic parts (Chua, Richler, & Gauthier, 2014, 2015; Richler et al., 2012; Richler, Wong, & Gauthier,

✉ Paulo Ventura
paulo.ventura@gmail.com

1   Faculdade de Psicologia, Universidade de Lisboa, Lisboa, Portugal

2   Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA

3   Centre for Research in Applied Communication, Culture and New Technologies (CICANT), Universidade Lusófona de Humanidades e Tecnologias, Lisboa, Portugal

2011), and concerning the third definition, the explicit representation of spatial relationships between features (e.g., Diamond & Carey, 1986; Leder & Bruce, 2000).

Face recognition is considered a unique form of object recognition because faces are processed more holistically than other types of objects (Farah et al., 1998). In contrast, word recognition is considered to rely on part-based processing because neither the number, order, nor configural relationships among letters reveal word identity (Grainger, 2008). Given these differences between faces and words, Farah and colleagues (Farah 1991, 1992; Farah et al., 1998) portrayed faces and words at two extremes on the object recognition continuum: holistic processing for faces and part-based processing for words.

Visual word processing is an acquired perceptual skill that allows us to rapidly identify words formed by a limited set of letters with high similarity (Kleinschmidt & Cohen, 2006; Wong et al., 2011a). Since holistic processing has been proposed as a general marker of perceptual expertise (Richler & Gauthier, 2014), one might expect both faces and words to be processed holistically. Indeed, there is recent evidence of holistic word processing (e.g., Liu et al., 2016; Ventura, 2014; Wong & Gauthier, 2007); the distinction between part-based word processing and holistic face processing may thus be oversimplified. Indeed, faces and words seem both to be processed holistically. Rather than undifferentiated wholes, face recognition involves representations of both the local elements (individual face parts) and their configuration (e.g., Farah et al., 1998; Maurer et al., 2002; Young et al.,1987). In a similar vein, in visual word recognition, it has long been shown that letter identities are not bypassed, and word holistic processing is not just about supra letter features (e.g., Paap et al., 1984). Holistic processing can thus be

defined as obligatory encoding of/attending to all object parts, which in turn are also encoded and represented independently (Richler & Gauthier, 2014).

The composite task is commonly used to assess holistic processing of faces (cf. Richler & Gauthier, 2014; Rossion, 2013). The composite task (cf. Fig. 1) is a perceptual task requiring participants to perform a same-different matching task on a specific visual part of two sequential visual stimuli. Two essential aspects of this task argue for holistic processing. First, the influence of the irrelevant part (e.g., the right half) on performance over the target part (e.g., the left half), i.e., a significant congruency effect: better performance when the irrelevant part is congruent in response to the one induced by the relevant part than when incongruent (i.e., "same" for both target and distractor parts or "different" for both target and distractor parts vs. "same" for target part and "different" for distractor part or "different" for target part and "same" for distractor part). Second, the congruency effect is modulated by alignment, i.e., it is drastically reduced when the two parts of the visual stimuli are misaligned (e.g., the right part is moved down relative to the left) rather than aligned, likely because the entire percept is disrupted. This interaction between alignment and congruency is more indicative of holistic processing than the observation of a congruency effect, which is tainted by response compatibility and decision-making processes.

For the task with words there is also an influence of the irrelevant part (e.g., the right half) on performance over the target part (e.g., the left half) which have been reported in recent studies of visual words in alphabetic and logographic scripts (Chen et al., 2013; Ventura et al., 2017; Wong, Bukach, et al., 2012a; Wong, Zhiyi, et al., 2011c). That is, a significant congruency effect: better performance when the irrelevant part is congruent in response to the one induced

		Same Trial		Different Trial	
		Congruent	Incongruent	Congruent	Incongruent
Aligned	Study	A B	A B	A B	A B
	Test	A B	A C	D C	D B
Misaligned	Study	A / B	A / B	A / B	A / B
	Test	A / B	A / C	D / C	D / B

Fig. 1 Illustration of the composite task with left-right face composites. Adapted from Liu and Behrmann (2014)

by the critical part (in same-response trials: e.g., LANE -LANE, as the critical and irrelevant parts are the same; in different-response trials: e.g., LANE – COZY, as both the critical and irrelevant parts induce a different-response) than when incongruent (in same-response trials: e.g., LANE -LADY, because the critical part of the two words is the same but the irrelevant part is different; in different-response trials: LANE - CONE, as the critical part of the words is different but the irrelevant part is the same).

The fact that holistic processing has been shown for both faces and words does not mean that the exact same mechanism underlies the effects of the two types of stimuli. For example, the word composite effect can occur at an abstract lexical level of representation that is not linked to a specific visual structure of the word (Ventura et al., 2017), and in another study the composite effect was only found when word pairs had syllables with a unique grapheme-to-phoneme mapping, indicating an automatic phonological modulation (Ventura et al., 2019).

However, evidence from a recent study hinted at some similarities between holistic processing of faces and words. Ventura et al., (2021) investigated the influence of global or local priming (Navon matching task using compound hierarchical figures) for holistic faces and words. In the Navon task, participants are asked to match two simultaneously presented letters or figures (Kinchla, 1974; Navon, 1977) – compound hierarchical figures with both a local and a global structure, i.e., larger letters/figures composed of smaller ones. In counterbalanced blocks instruction requires attention to either the global (large letters/figures relevant) or the local level (smaller letters/figures relevant), while ignoring the irrelevant level. This design evokes a robust global advantage.

Ventura et al., (2021) replicated the effects obtained by Gao et al. (2011), who previously showed that local or global processing in a Navon task primes local or global processing in a subsequently presented composite task for faces. Second, similar stronger global priming effects were found for faces and words, suggesting that holistic face and word processing were susceptible to attention manipulations to similar degrees.

The present study aimed to directly address whether faces and words share holistic processing mechanisms. We modified a task designed to assess overlap in early perceptual stages (cf. Curby and Moerel, 2019) to test whether a trade-off between holistic processing of faces and words can be found when they are superimposed. Participants were asked to ignore one stimulus category and focus on the other category. They were told to focus on the left side of the attended category (either face or word) and compare the relevant sides of the two stimuli in a trial, ignoring the irrelevant sides.

Words were divided between the second and third letters. Consequently, we used matched left-right face composites instead of the more common upper half-lower half face composites. If words and faces engage shared holistic processing,

we would expect a trade-off in holistic processing when they are superimposed and aligned. For example, faces would be processed less holistically when superimposed with aligned than misaligned words because aligned words are also processed holistically. Alternatively, if holistic processing of words and faces are relatively more independent of each other (i.e., can be processed in parallel with less interference), then we would expect minimal differences in holistic processing of one when the other stimuli are aligned or misaligned. We examined these two possibilities in Experiment 1. Experiment 2 used the exact same paradigm and displays as in Experiment 1, except that participants made judgments about the words while ignoring faces.

Experiment 1: Attended faces and unattended words

Method

Design

In Experiment 1, the task was to attend to faces and ignore words; the corresponding within-subject factors were face alignment, face congruency, and word alignment.

Holistic processing of the attended stimuli, faces, is reflected by an interaction between the first two factors (face alignment × face congruency). The question of interest is whether the task-irrelevant word alignment interferes with the holistic processing of faces. This would be revealed by a three-way interaction between word alignment, face alignment, and face congruency. To investigate the underlying source of a hypothetical three-way interaction, the data from the trials where the words were aligned and those where they were misaligned were analyzed separately. The presence or absence of an interaction between face alignment and face congruency indicates whether holistic processing of the faces is occurring or not.

Holistic processing of faces was complementarily evaluated by the following subtraction: (congruent aligned - incongruent aligned trials) minus (congruent misaligned – incongruent misaligned trials), that is the congruency effect under aligned and misaligned word conditions. The question of interest is whether the task-irrelevant word alignment interferes with the holistic processing of faces. Thus, the critical comparison is the magnitude of holistic processing of faces under different word contexts (word-aligned vs. word-misaligned).

Participants

Prior to the study, we performed a power analysis based on results from Curby and Moerel (2019). Specifically, we leveraged the critical three-way interaction between line pattern

alignment, face alignment, and face Congruency ($\eta_p^2 = .26$) in Experiment 1 of Curby and Moerel (2019). Using MorePower 6.0.4 (Campbell & Thompson, 2012), a sample size of 32 would be required to find a comparable effect at $\alpha = 0.05$ with a power of 0.9 for a $2 \times 2 \times 2$ repeated-measures ANOVA.

Despite the predetermined minimum sample size of 32, all students enrolled in a psychology course in Faculdade de Psicologia of Universidade de Lisboa were invited to participate due to anticipated data exclusion considering our exclusion criteria defined a priori. Sixty-five participants took part in Experiment 1. Data from 12 participants were excluded (see below for details of the exclusion criteria).

The study's protocol adhered to the guidelines of the Declaration of Helsinki and the Portuguese deontological regulation for Psychology and was approved by the Deontological Committee of Faculdade de Psicologia of Universidade de Lisboa. All participants provided informed consent.

Stimuli

Composite words 168 disyllabic Consonant-Vowel Consonant-Vowel (CV CV) Portuguese words in Tracker font were used (see Fig. 2). A thin vertical blue line (2-pixel wide) between the second and third letter divides each word into a left and a right half (average word-aligned: 242 × 184 pixels; average word-misaligned: 242 × 264 pixels). Misaligned words were created by moving down the right half of words by approximately 80 pixels on average.

Composite faces Because the words superimposed on the faces had a horizontal orientation and were divided into a left and a right half, we elected to use matched left-right face composites. The face composites were the same as the Caucasian face subset used in Liu and Behrmann (2014) and Liu et al., (2014) studies. The 20 faces were subdivided into five groups of four similar faces based on prior, independently
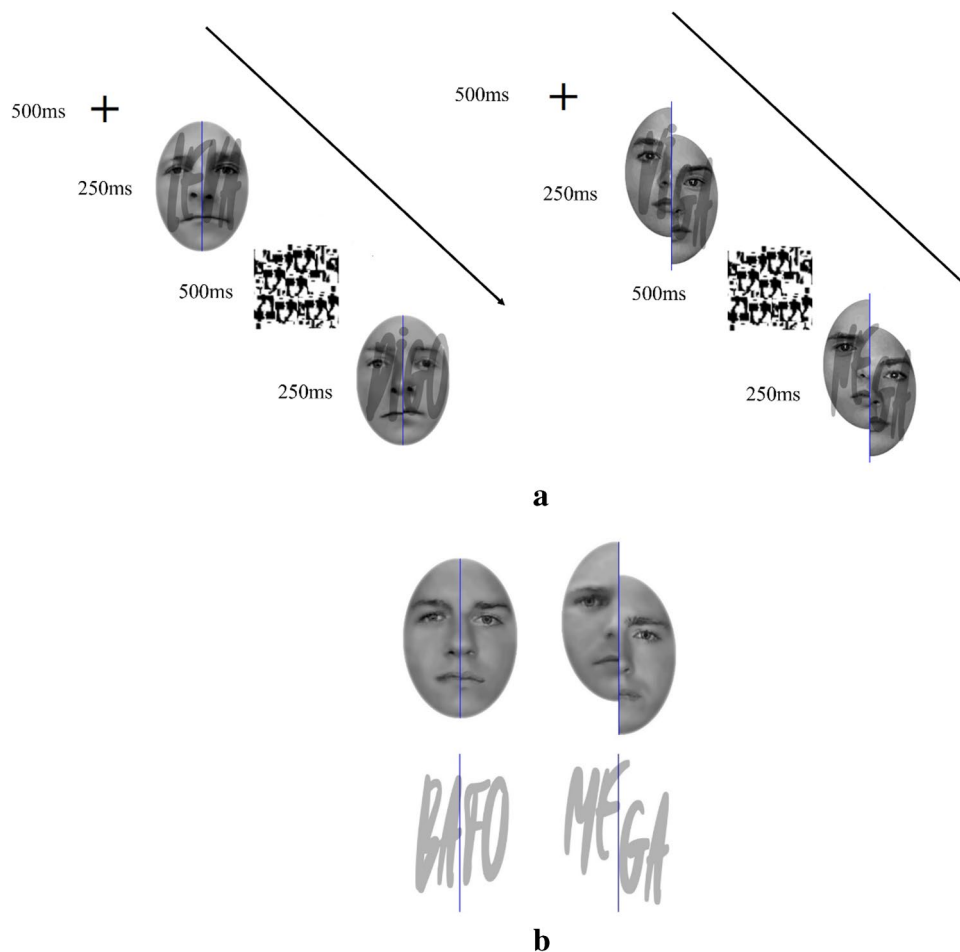

Fig. 2 a Procedure used for the modified composite task. Aligned or misaligned composite word stimuli were overlaid on top of either aligned (**left**) or misaligned (**right**) faces. This figure illustrates two sample trials – one face aligned, and word aligned trial, and one face misaligned, and word misaligned trial – of the task. In Experiment 1, participants made same/different judgments on the left halves of the study and test faces. **b** Examples of the faces and words in Tracker font that were superimposed in the Experiment

assessed ratings. This ensured that the task could not be performed based purely on facial symmetry. Each composite face was then created by pairing the left half of one face with the right half of another face from the same group (274 × 384 pixels). Each misaligned composite face was created by moving the right half down by approximately one-third of the face (274 × 464 pixels).

Procedure

Participants were tested on-line using E-Prime Go (https://pstnet.com/eprime-go/). They could only use a PC to participate (neither smartphones nor tablets were accepted). Timing, sequence of events within-trial, and data collection (accuracy and response time (RT) from target onset) were controlled by E-Prime 3.04 (www.pstnet.com/eprime). Participants were free to choose the best time to run the experiment. Participants could not have neurological disorders (e.g., epilepsy), could not be taking psychiatric medication, and could not have any developmental disorders (e.g., dyslexia, attention deficit disorder). Participants were instructed to put their cell phones on silent and prepare the environment to carry out the study alone, in a calm and uninterrupted context. They were asked to close other programs that they may have running on the computer and close other windows in the browser (only keep open the window corresponding to the link sent). Participants were also instructed not to interrupt participation (complete the requested tasks until the end). They were also reminded that if they need glasses to be at the computer, they must put them on before starting.

Each participant completed a total of 384 trials divided over four blocks. In each trial, face composite images were presented with word composite images overlaid on top (see Fig. 2). A 2-pixel wide vertical blue line was placed on the midline of each word (i.e., between the second and third letters) and the midline of each face. Each trial proceeded as follows (see Fig. 2): (1) fixation screen (500 ms), (2) study stimulus (i.e., a face composite with a word composite overlaid; 250 ms), (3) pattern mask (500 ms), (4) test stimulus (i.e., a face composite with word composite overlaid; 250 ms). Participants were instructed to make same-different judgments on the left halves of the two sequentially displayed faces while ignoring the right halves of faces and the overlaid words. Each participant completed 16 practice trials prior to the experiment.

The left and right halves of word and face stimuli were either aligned or misaligned, resulting in four stimulus configurations ("face aligned, and word aligned," "face aligned, and word misaligned," "face misaligned, and word aligned," and "face misaligned, and word misaligned"). Trials of the same stimulus configurations were blocked, and the block order was randomized (96 trials per block). The correct response to the left half of the face image (same or different) and the congruency of the relationship between the task-irrelevant right halves with the correct response were counterbalanced within a block. The congruency for the task-irrelevant words was also counterbalanced with the congruency for the faces within each block. The stimuli were exactly the same across both experiments.

Results and discussion

Four outlier participants, with mean RT > 2.5 SD from the group mean were identified and removed from further analysis. Three participants had low sensitivity (mean d′ < 0) and were dropped from the analysis. Five other participants had poor sensitivity performance (mean d′ < 0) in at least one of the eight conditions of the design (Face Alignment × Face Congruency × Word Alignment) and were also excluded. Thus, the results of 53 participants were analyzed. All exclusion criteria were determined a priori.

Sensitivity analysis

For the computation of d′, we used for hits the correct responses to "different" responses, and for the false alarms, we used 1-Hit ("same" responses). A 2 (face alignment: aligned, misaligned) × 2 (face congruency: congruent, incongruent) × 2 (word alignment: aligned, misaligned) ANOVA performed on the sensitivity (d′) scores revealed a three-way interaction between word alignment, face alignment and congruency, $F(1, 52) = 4.21$, $p = .045$, $\eta_p^2 = .08$. This three-way interaction is the crucial effect.

To investigate the underlying source of the three-way interaction, the data from the trials where the words were aligned and those where they were misaligned were analyzed separately. The presence or absence of an interaction between face alignment and face congruency indicates whether holistic processing of the faces is occurring or not. Within the misaligned word condition, the 2 (face congruency) × 2 (face alignment) ANOVA revealed a main effect of face alignment, $F(1, 52) = 4.93$, $p = .03$, $\eta_p^2 = .09$, but no main effect of face congruency, $F < 1$. There was an interaction between face congruency and face alignment, $F(1, 52) = 14.68$, $p < .002$, $\eta_p^2 = .17$. Thus, holistic face processing was found in the presence of misaligned words (see Fig. 3 and Table 1).

In contrast, the 2 (face congruency) × 2 (face alignment) ANOVA of the aligned word trials showed no interaction between face alignment and face congruency, $F < 1$. Thus, whereas holistic face processing was found in the presence of misaligned words, no holistic processing was found in the context of aligned words (see Fig. 3 and Table 1).

In sum, we found a difference between holistic processing of faces under word aligned vs. word misaligned conditions,
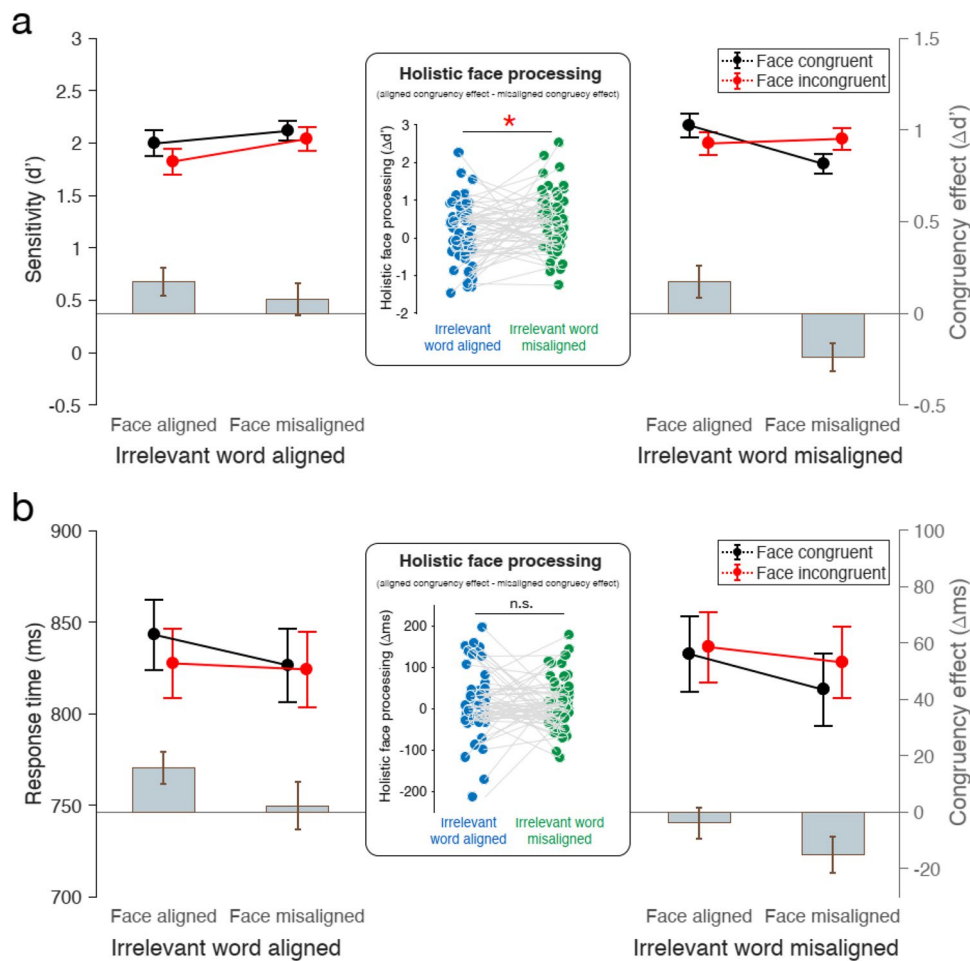
Fig. 3 Mean sensitivity (d′) for the face congruent and incongruent conditions, and the resulting index of holistic perception (congruency effect, filled bars, reflecting the difference between the congruency conditions) for the faces overlaid with aligned (**left panel**) and misaligned (**right panel**) words in Experiment 1. The mean response time (ms) for accurate trials for the faces overlaid with aligned (**left panel**) and misaligned (**right panel**) words is also shown. Error bars represent standard error values. The central panels show holistic face processing (defined as the subtraction of face aligned congruency effect – face misaligned congruency effect) for word aligned and word misaligned conditions

Table 1 Mean and standard error (d') for the congruent and incongruent face conditions, for aligned and misaligned faces overlaid with aligned and misaligned words in Experiment 1

Aligned words				Misaligned words			
Face aligned		Face misaligned		Face aligned		Face misaligned	
Face Congruent	Face Incongruent	Face Congruent	Face Incongruent	Face Congruent	Face Incongruent	Face Congruent	Face Incongruent
1.99 (.12)	1.82 (.12)	2.12 (.09)	2.04 (.11)	2.17 (.11)	2.0 (.10)	1.8 (.09)	2.04 (.10)

suggesting a specific interference with holistic face processing due to word alignment.

We also compared the magnitude of holistic face processing when irrelevant words were aligned or misaligned. To help understand this logic we present a numerical example.

The magnitude of holistic face processing when irrelevant words were aligned is a subtraction of congruent and incongruent trials for word-aligned and face-aligned trials (1.74–1.0) minus the subtraction of congruent and incongruent trials for word-aligned and face-misaligned trials

(2.13–2.30). In this example the magnitude of holistic face processing when irrelevant words were aligned is .91.

Holistic face processing was significantly smaller in the presence of aligned words (mean = .1, SD = .8) than misaligned words (mean = .41, SD = .78), $t(52) = 2.05$, $p = .045$ (cf. Fig. 3). In sum, we found a difference between holistic processing of faces under word aligned vs. word misaligned conditions, suggesting a specific interference with holistic face processing due to word alignment.

Response-time analysis

Data from the same twelve participants were excluded from the analysis. Trials with RT < 200 ms or > 1,750 ms (similar to the criteria used in Curby & Moerel, 2019) were removed from the analysis (< 1%). Only trials with a correct response were analyzed. A 2 (face alignment: aligned, misaligned) × 2 (face congruency: congruent, incongruent) × 2 (word alignment: aligned, misaligned) ANOVA was performed on the remaining RT data from correct trials.

The three-way interaction between face alignment, face congruency, and word alignment was not significant (F < 1; cf. Fig. 3 and Table 2). Holistic face processing was equivalent in the presence of aligned words (mean = 13.46, SD = 80.17) and misaligned words (mean = 11.12, SD = 61.06), $t(52) = .16$, $p = .87$.

Although previous studies have reported holistic processing using the RT measure, our finding of significant holistic face processing in sensitivity analysis, but not in RT analysis, was not unusual. In fact, it was similar to the significant three-way interaction in d´ but not in RT in Experiment 1 of Curby and Moerel (2019).

Experiment 2: Unattended faces, attended words

Method

Design

We reasoned that, if word and face processing recruit partially overlapping resources, then one's interference in holistic processing of the other should be reciprocal. To test the degree to which face alignment interferes with holistic word processing, Experiment 2 was identical to Experiment 1 except that the task was to attend to words and ignore faces. The overlaid aligned and misaligned faces create high and low interference conditions, respectively.

Here, the corresponding within-subject factors are word alignment, word congruency, and face alignment. Holistic processing of the attended stimuli, words, is reflected by an interaction between the face alignment and face congruency. The question of interest is whether the task-irrelevant face alignment interferes with the holistic processing of words. This would be revealed by a three-way interaction between face alignment, word alignment, and word congruency. To investigate the underlying source of a hypothetical three-way interaction, the data from the trials where the faces were aligned and those where they were misaligned were analyzed separately. The presence or absence of an interaction between word alignment and word congruency indicates whether holistic processing of the words is occurring or not.

Holistic processing of words was complementarily evaluated by the following subtraction: (congruent aligned – incongruent aligned trials) minus (congruent misaligned – incongruent misaligned trials), that is the congruency effect under aligned and misaligned face conditions. The question of interest is whether the task-irrelevant face alignment interferes with the holistic processing of words. Thus, the critical comparison is the magnitude of holistic processing of words under different face contexts (word-aligned vs. word-misaligned).

Participants

We retained the indication of a sample size of 32 from the power analysis in Experiment 1. Nevertheless, all students enrolled in a psychology course in Faculdade de Psicologia of Universidade de Lisboa were invited to participate due to anticipated data exclusion considering our exclusion criteria defined a priori. Fifty-four participants accepted the invitation and took part in Experiment 2. None of them had participated in Experiment 1. Seven participants were excluded (see below for detailed exclusion criteria).

Table 2 Mean and standard error (RT) for the congruent and incongruent face conditions, for aligned and misaligned faces overlaid with aligned and misaligned words in Experiment 1

Aligned words				Misaligned words			
Face aligned		Face misaligned		Face aligned		Face misaligned	
Face Congruent	Face Incongruent	Face Congruent	Face Incongruent	Face Congruent	Face Incongruent	Face Congruent	Face Incongruent
843.07 (18.2)	827.41 (17.88)	826.33 (18.81)	824.13 (19.59)	832.51 (19.21)	836.36 (18.0)	813.03 (18.46)	828.0 (18.48)

This study's protocol adhered to the guidelines of the Declaration of Helsinki and the Portuguese deontological regulation for Psychology and was approved by the Deontological Committee of Faculdade de Psicologia of Universidade de Lisboa. All participants provided written informed consent.

Procedure

The stimuli and procedure were identical to Experiment 1, except that participants were instructed to make judgments about the words instead of the faces.

Results and discussion

Mean RTs were examined for outliers (mean RT > 2.5 SD from group mean), resulting in two participants being excluded from further analysis. Two participants had poor sensitivity performance (mean d′ < 0) and were excluded from further analysis. Three additional participants had poor sensitivity performance (mean d′ < 0) in at least one of the eight conditions of the design (word alignment × word congruency × face alignment) and were also excluded. Thus, the results of 47 participants were analyzed. All exclusion criteria were determined a priori.

Sensitivity analysis

A 2 (face alignment: aligned, misaligned) × 2 (word alignment: aligned, misaligned) × 2 (word congruency: congruent, incongruent) ANOVA performed on the sensitivity (d′) scores (cf. Fig. 4 and Table 3) revealed a non-significant three-way interaction, $F(1, 46) = 3.14$, $p = .08$, $\eta_p^2 = .06$.

Holistic word processing was equivalent in the presence of aligned faces (mean = .15, SD = .72) and misaligned faces (mean = -.05, SD = .40), $t(46) = 1.77$, $p = .08$.

Response-time analysis

The same seven participants were removed from the analysis. Trials with a response time < 200 ms or > 1,750 ms; < 1% (similar to the criteria in Curby & Moerel, 2019) were removed from the data. We analyzed response times from correct trials only. A 2 (face alignment: aligned, misaligned) × 2 (word alignment: aligned, misaligned) × 2 (word congruency: congruent, incongruent) ANOVA was performed on the remaining RT data from correct trials (Fig. 4 and Table 4).

The three-way interaction was significant, $F(1, 46) = 16.55$, $p < .001$, $\eta_p^2 = .27$. To probe the underlying source of the three-way interaction, the data from the trials where the faces were aligned and those where they were misaligned were analyzed separately. Holistic processing is indicated by the presence of an interaction between word alignment and word congruency, with a greater congruency effect for

aligned than for misaligned words. Thus, this interaction indicates whether holistic processing of the words is occurring.

When faces are aligned, and thus processed holistically, there is a significant interaction between word alignment and congruency, $F(1, 46) = 5.98$, $p = .018$, $\eta_p^2 = .12$. This interaction, however, does not reflect a holistic processing effect for words; indeed, it is for misaligned words that the congruent condition is faster than the incongruent condition.

For misaligned faces, not processed holistically, we find evidence of holistic processing for words. The interaction of word alignment and congruency was significant, $F(1, 46) = 7.35$, $p < .009$, $\eta_p^2 = .14$. In aligned trials, RTs for congruent trials were faster than RTs for incongruent trials. This difference disappeared for misaligned trials.

Holistic word processing was significantly smaller in the presence of misaligned faces (mean = -15.17, SD = 38.63) than aligned faces (mean = 16.49, SD = 46.24), $t(46) = 4.06$, $p = < .0001$ (cf. Fig. 4). Note that holistic processing of faces was complementarily evaluated by the following subtraction: (congruent aligned – incongruent aligned trials) minus (congruent misaligned – incongruent misaligned trials), that is the congruency effect under aligned and misaligned words conditions. Considering RTs, congruency effect aligned should be smaller than congruency effect misaligned.

The absence of a three-way interaction for d′ scores with a significant three-way interaction in RTs obtained here matches the results reported in Curby and Moerel's (2019) Experiment 2.

General discussion

The question of whether word and face recognition rely on shared or dissociable neural resources and cognitive processes is under heated discussion (for recent reviews, see Burns & Bukach, 2021; Burns & Bukach, 2022; Gerlach & Starrfelt, 2022; Rossion & Lochy, 2021). Evidence supporting a dissociable mechanism comes from human neuroimaging studies that report category-selective regions (e.g., faces, bodies, places, and words) in the human ventral temporal cortex (Cohen & Dehaene, 2004; Kanwisher et al., 1997) suggesting a modular view of the functional architecture of the mind (Burns et al., 2017; Rubino et al., 2016; Saygin et al., 2015; Starrfelt et al., 2018; Susilo & Duchaine, 2013; Susilo et al., 2015). In addition, word and face processing are typically lateralized to opposite hemispheres (i.e., left and right hemispheric lateralization of word and face recognition, respectively), so traditionally it has been considered that they depend on different neural resources and distinct cognitive processes (for a recent example, cf. Hagen et al., 2021).

We found a trade-off between the holistic processing of faces and words, which is difficult to explain by the
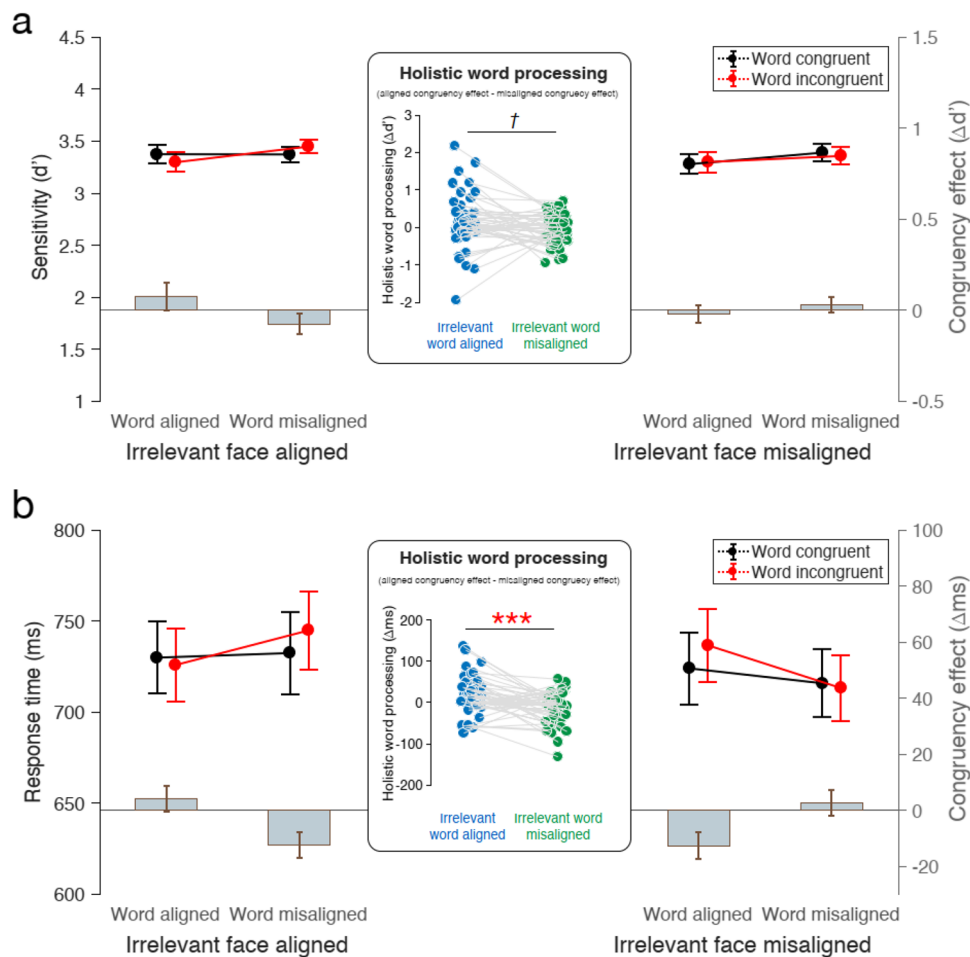
Fig. 4 Mean sensitivity (d′) for the word congruent and incongruent conditions, and the resulting index of holistic perception (congruency effect, filled bars, reflecting the difference between the congruency conditions) for the words overlaid with aligned (**left panel**) and misaligned (**right panel**) faces in Experiment 1. The mean response time (ms) for accurate trials for the words overlaid with aligned (**left panel**) and mis-aligned (**right panel**) faces is also shown. Error bars represent standard error values. The central panels show holistic word processing (defined as the subtraction of word aligned congruency effect – word misaligned congruency effect) for face aligned and face misaligned conditions

Table 3 Mean and standard error (d') for the congruent and incongruent word conditions, for aligned and misaligned words overlaid with aligned and misaligned faces in Experiment 2

Aligned faces				Misaligned faces			
Word aligned		Word misaligned		Word aligned		Word misaligned	
Word Congruent	Word Incongruent	Word Congruent	Word Incongruent	Word Congruent	Word Incongruent	Word Congruent	Word Incongruent
3.37 (.08)	3.30 (.09)	3.37 (.08)	3.45 (.07)	3.28 (.09)	3.30 (.08)	3.39 (.08)	3.36 (.08)

modular position. That is, if faces and words were processed independently, holistic processing of both would not have produced the interference effects we found. Instead, the interference between holistic processing of faces and words is in line with the view that word and face recognition rely on partly shared cortical resources. For example, according to the many-to-many hypothesis

(Behrmann & Plaut, 2013, 2014, 2015, 2020; Plaut & Behrmann, 2011) during the course of literacy acquisition, competition between word and face representations emerges (Behrmann & Plaut, 2015; Dehaene et al., 2015, Liu et al., 2018) and reductions in holistic face processing ensue (Ventura et al. 2013). According to this distributed account with the systems supporting face and word

Table 4 Mean and standard error (RT) for the congruent and incongruent word conditions, for aligned and misaligned words overlaid with aligned and misaligned faces in Experiment 2

Aligned faces				Misaligned Faces			
Word aligned		Word misaligned		Word aligned		Word misaligned	
Word congruent	Word incongruent	Word congruent	Word incongruent	Word congruent	Word incongruent	Word congruent	Word incongruent
729.95 (19.75)	725.89 (19.91)	732.47 (22.56)	744.91 (21.51)	724.07 (19.79)	736.71 (20.0)	715.85 (16.69)	713.22 (18.14)

recognition exhibiting graded and overlapping functional specialization both within and, especially, between hemispheres (Behrmann & Plaut, 2013, 2014, 2015, 2020; Plaut & Behrmann, 2011), in addition to the possible overlap of neural regions, the behavioral signatures typically associated with either holistic- or part-based processing may apply to both faces and words. This is very interesting given that the image properties of faces and words are completely distinct – whereas faces comprise a three-dimensional (3D) structure with more curved features and with parts which are not easily separable (e.g., eyes, nose, and mouth), words are composed of two-dimensional (2D) structures with individual letters that occur independently in their own right and are made of mostly straight edges (Behrmann & Plaut. 2020).

The many-to-many hypothesis postulates that cortical regions in the brain are involved in processing a variety of visual stimuli as opposed to a single stimulus. Thus, it is asserted that the fusiform gyri of both hemispheres are involved in both visual word processing and face processing. Lateralization, driven by the need to make connections between visual word and other linguistic processing as efficient as possible, results in the left hemisphere processing predominantly visual words and the right hemisphere processing predominantly faces. However, this lateralization is incomplete, and there is substantial overlap between the regions activated by faces and visual words in each hemisphere. This has been interpreted as competition for and sharing of limited perceptual resources (Behrmann & Plaut, 2013; Dundas et al., 2013, 2014).

Collins et al., (2017) tested congenital prosopagnosics and developmental dyslexics. Behaviorally, the dyslexic group exhibited clear deficits in both word and face processing relative to controls, while the prosopagnosia group showed a specific deficit in face processing only. This pattern was mirrored in the evoked response potential (ERP) data too. These findings are consistent with the hypothesis that the typical hemispheric organization for words can develop in the absence of typical hemispheric organization for faces but not vice versa, supporting the many-to-many account.

The study by Furubacke et al., (2020) provide some support for the prediction of the many-to-many hypothesis, that face and visual word processing share neural resources.

However, the networks in the brain for face and visual word recognition likely show hemispheric differences in the perceptual operations performed on face and visual word stimuli, in accordance with a view that hemispheric functions are complementary rather than equivalent.

In addition, word impairments in prosopagnosia (individuals with difficulty in face recognition) and face impairments in pure alexia (individuals with an impairment in reading) have been found (Behrmann & Plaut, 2014), and the severity of impairments in face recognition can predict the severity of impairments in word processing (Burns & Bukach, 2021). Importantly, studies with dyslexics show that they have impaired facial recognition (Sigurdardottir et al., 2015). Sigurdardottir et al., (2018) showed that people who were worse at face matching had greater reading problems. In a second experiment, matched dyslexic and typical readers were tested, and difficulties with face matching were consistently found to predict dyslexia over and above novel object matching and general object perception mechanisms. Sigurdardottir et al., (2018) speculate that dyslexic readers have specific visual problems with individuating visually homogeneous objects, such as faces and words, with which people have prior experience. Further evidence for a concurrent impairment in face perception in dyslexics is provided by Gabay et al., (2017; cf. also Collins et al., 2017) Relative to controls, dyslexic individuals matched faces more slowly, and showed disproportionate cost in performance when target and distractor faces differed in viewpoint, and discriminated faces more poorly, particularly as the faces were increasingly alike perceptually. Brady et al. (2021) evaluated whether holistic processing is anomalous in dyslexia. Specifically, they compared holistic processing of words and of faces in participants with dyslexia and age-matched controls. Holistic processing of faces is comparable in dyslexic and typical readers, but dyslexic readers show greater holistic processing of words. Brady et al. (2021) also showed that holistic processing of both faces and words predicts reading performance in the dyslexic group (with a more holistic style associated with better accuracy and speed scores) but not in the typical reader group.

However, other authors (Robotham & Starrfelt, 2017) have found evidence that reading can be preserved in acquired and developmental prosopagnosia and also

evidence (though weaker) that face recognition can be preserved in acquired or developmental dyslexia, suggesting that face and word cognition are at least in part supported by independent processes.

It has been suggested that reading acquisition drives the lateralization of the face processing system to the right hemisphere of the brain (Dehaene et al., 2010). Kuhn et al. (2021) investigated whether this developmental co-dependency has a behavioral cost, at least in the short term, by testing whether learning to read diminished face recognition ability. In a longitudinal study, 82 children aged 5–7 years were evaluated twice, at the beginning and end of their first school year. Visual letter recognition, word recognition, sentence reading, and immediate face recognition were tested during both sessions. Kuhn et al. (2021) found no evidence of a negative correlation between literacy acquisition and performance in face recognition. In a subsequent study involving dyslexics, Gerlach et al. (2022) found little evidence in support of a correlation between the acquisition of reading skills and the lateralization of face recognition.

Our findings suggest that the processes underlying face and word expertise may at least share some similar properties, namely holistic processing, and hint that both forms of expertise may employ a single shared mechanism. However, other evidence regarding inversion effects on efficiency (Albonico, Furubacke, Barton, & Oruc, 2018) suggest that, despite these similar properties, there might be some key differences in the quality of face and word processing. While inversion profoundly reduced face processing efficiency, it had a markedly lesser effect on visual word processing efficiency (Albonico et al., 2018).

Furthermore, Feizabadi et al., (2021) reviewed the evidence for experimental effects in word recognition that parallel those used to support holistic face processing, namely inversion effects, the part-whole task, and composite effects. The observations support some parallels in whole object influences between face and visual word recognition, but they do not necessarily imply similar expert mechanisms. of visual stimuli processing for faces and words. Both may show whole object effects but for different reasons. In the case of visual words, they may reflect a top-down modulation of the primary interactive hierarchical mechanism, while in the case of faces they may derive from a primary holistic mechanism, in which early

bottom-up processing is based on holistic face representations (Feizabadi et al., 2021).

Regarding the question of shared versus independent cognitive processes, behavioral evidence suggests that word and face recognition may rely, at least in part, on similar cognitive processes. For example, Ventura et al., (2021) investigated the role of global or local priming on the composite face or word effect under aligned and misaligned conditions. Global priming using compound figures promoted holistic

processing of words and faces, and to a similar extent, suggesting that holistic processing of these two different types of stimuli may rely on similar mechanisms.

The present investigation takes a step further and directly shows a reciprocal interference between the holistic processing of word and face stimuli. Faces were processed less holistically when an aligned word was superimposed, while words were processed less holistically when an aligned face was superimposed. This finding evidenced a trade-off in holistic processing of the two stimuli, suggesting that faces and word stimuli rely, at least in part, on similar holistic processing mechanisms. It is important to note that these interference effects take place at the holistic processing level, as they are contingent on the stimuli's potential for holistic processing (i.e., their alignment). Given that misalignment disrupts holistic perception of both faces and words, this finding is consistent with a trade-off in holistic processing between the stimulus classes. That is, a reduction in holistic processing of the overlaid face or word stimulus frees up processing resources and thereby results in an increase in holistic processing of the other, concurrently processed stimulus. These findings suggest that the mechanisms supporting holistic perception of faces and words are not independent. Nevertheless, we discuss below other interpretations of our pattern of results.

An alternative explanation could be that the alignment itself, as opposed to competing or shared holistic processes, is at stake; both words and faces are easier to process when they are aligned, without necessarily requiring holistic processing. It is essential to note, however, that the interference between the processing of faces and words is specifically a function of holistic processing, as it is modulated by the extent to which the stimuli engage holistic processing capacity – that is, the alignment of the stimuli (Curby & Moerel, 2019). Consequently, these findings cannot be explained by the expected overlap between the processing of all visual stimuli. If this were the case, both aligned and misaligned stimuli would have been equally effective at interfering with the concurrent processing of stimuli from the other stimulus class (Curby & Moerel, 2019).

One may ask whether automatic processes are responsible for the pattern of our results instead of holistic processing. This does not seem reasonable. Interference between holistic processing of words and faces is present even though the processing of only one of the stimuli is task relevant. This suggests that participants were obliged to process both stimuli, and potentially did so automatically (Curby & Moerel, 2019). This provides further support for the locus of this interference being in early, rather than late, processing stages. This automaticity allows the overlap between the processing of face and words which might occur via the mechanisms responsible for learned attentional strategies (Chua et al., 2014, 2015) which may be common to

faces and non-face objects of expertise (Curby & Moerel, 2019). Holistic processing can be defined as obligatory encoding of/attending to all object parts, which in turn are also encoded and represented independently (Richler & Gauthier, 2014). Attending to all parts is applied to face recognition and visual word recognition under a perceptual expertise framework (e.g., Liu et al., 2016; Ventura, 2014; Wong & Gauthier, 2007). This shared mechanism might be at the origin of our pattern of results. For words, this may be supported by the posterior-to-anterior organization of the VWFA (Dehaene et al., 2004; Thesen et al., 2012; Vinckier et al., 2007). The VWFA, as it is generally agreed, intervenes in the efficient identification of orthographic stimuli (Dehaene et al., 2001) and enables quick association of such stimuli with phonological and lexical information. In expert alphabetic readers, the VWFA is organized in a posterior-to-anterior hierarchy (Dehaene et al., 2004; Thesen et al., 2012; Vinckier et al., 2007): posterior parts respond to individual letters (thus underpinning sublexical representations), irrespective of case (Dehaene et al., 2004; Thesen et al., 2012), and as such letters are abstract units at this level. The lateral anterior region is sensitive to lexical properties, underpins holistic word representations, and has greater connectivity to language and conceptual neural networks (Bouhali et al., 2019; Lerma-Usabiaga et al., 2018). Holistic processing may intervene to bind together individual letters that activate the posterior part of the VWFA, providing the input that activates more anterior parts of the VWFA.

Our interpretation of the results as reflecting a trade-off between the holistic processing of faces and words should not be seen as an end to the debate between a modular and an interactive hypothesis and future studies should be run. For example, including a control condition/control stimulus type, such as aligned/misaligned objects for example, houses, or another category that is unrelated to faces and words, would allow determining whether what is occurring is unique to faces and words.

A further and important question is to evaluate the degree of interference between the holistic processing of words and faces. Words can be thought of as more cohesive units than faces for three reasons. First, from a bottom-up, stimulus-driven perspective, word recognition and reading involve scanning letters or chunks of letters (graphemes) in quick succession and grouping them into words, which is a more demanding task than face recognition (e.g., a limited set of facial features with broadly similar spatial arrangement and faces are not presented in quick spatial and temporal succession). Second, feedback from lexical processing to early perceptual processing contributes to word cohesiveness (Reicher, 1969; Wheeler, 1970). For example, the advantage of word context is interpreted as a top-down influence of whole-word representations on letter recognition (McClelland & Rumelhart, 1981). Third, words have phonology and semantics

that might help the perceived cohesiveness of words through reentrant feedback to the orthographic level both directly and indirectly (Seidenberg and McClelland, 1989).

One limitation of the present study is that it does not allow for a fair comparison of the two stimulus categories. First, in our experiments, words and faces were both divided into left and right halves. While the words we used are read from left to right, this may not be the universal scanning direction when viewing faces. There is, however, a slight tendency to look at the left-side of the face (from the viewer's perspective; Hsiao & Cottrell, 2008). In particular, the left eye was shown as the earliest diagnostic feature from the eye movement data (Vinette, Gosselin, & Schyns, 2004). Indeed, face processing ability, according to Royer et al. (2018), is associated with a systematic increase in the use of the eye area, particularly the left eye from the observer's perspective. Second, the faces we used in the present study are unfamiliar to the participants and thus lack phonology (i.e., recognizable identity) and person-specific semantic information. To allow for a fairer comparison between words and faces, a future study could use famous faces, since famous faces contain both phonological and semantics information.

In our study, participants were instructed to focus on the left side of stimuli, thus the experiments were not counterbalanced with respect to lateralization of the attended part of the stimulus. This is the design we have adopted in our many studies with words and faces and we always found composite effects. The left part of the stimuli is first processed by the right hemisphere, suggested by some (Rossion et al., 2000) to rely more on "holistic processing". This does not seem to have influenced the results. We also found a composite word effect in the left but not the right hemifield, which would be consistent with localizing whole word holistic effects also to the right hemisphere (Ventura et al., 2019).

In Experiment 1, the hallmark of holistic processing for faces revealed in the composite task was attenuated when they were processed in the context of aligned (holistically processed) words, relative to when they were processed in the context of (non-holistically processed) words. This was observed in sensitivity analysis, but not in RT analysis. In Experiment 2, words processed in the context of aligned faces (holistically processed) failed to show hallmarks of being holistically processed, while those processed in the context of misaligned faces (non-holistically processed) did. This was observed in RT analysis, but not in sensitivity analysis. It is not unusual for effects in the composite task to be found in RT or d′ or both (e.g., Curby et al., 2016; Curby & Moerel, 2019). Importantly, there was no evidence of a speed-accuracy trade-off in any of the Experiments as the d′ data and RT data showed the same general pattern.

In conclusion, through two experiments, we show a reciprocal interference between holistic processing of

word stimuli and face stimuli, implying that the mechanisms supporting the holistic perception of words are not independent of those supporting the holistic perception of faces. The findings have important implications for our understanding of a hypothetical shared mechanism between face and word processing, and more broadly, the functional architecture of the mind.

Data availability The data for both experiments are available through Figshare: https://doi.org/10.6084/m9.figshare.20923204.

Code availability Can be provided upon reasonable request.

Declarations

Conflicts of interest/Competing interests None.

Ethics approval This study's protocol adhered to guidelines of the Declaration of Helsinki, the Portuguese deontological regulation for Psychology, and was approved by the Deontological Committee of Faculdade de Psicologia of Universidade de Lisboa.

Consent to participate All participants provided written informed consent. Participants received course credit.

Consent for publication Not applicable.

References

Albonico, A., Furubacke, A., Barton, J., & Oruc, I. (2018). Perceptual efficiency and the inversion effect for faces, words and houses. *Vision Research, 153*, 91–97. https://doi.org/10.1016/j.visres.2018.10.008

Baker, C. I., Behrmann, M., & Olson, C. R. (2002). Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nature Neuroscience, 5*(*11*), 1210–1216. https://doi.org/10.1038/nn960

Behrmann, M., & Plaut, D. C. (2013). Distributed circuits, not circumscribed centers, mediate visual cognition. *Trends in Cognitive Sciences, 17*(*5*), 210–219. https://doi.org/10.1016/j.tics.2013.03.007

Behrmann, M., & Plaut, D. C. (2014). Bilateral hemispheric processing of words and faces: Evidence from word impairments in prosopagnosia and face impairments in pure alexia. *Cerebral Cortex, 24*(4), 1102–1118. https://doi.org/10.1093/cercor/bhs390

Behrmann, M., & Plaut, D. C. (2015). A vision of graded hemispheric specialization. *Annals of the New York Academy of Sciences, 1359*(1), 30–46. https://doi.org/10.1111/nyas.12833

Behrmann, M., & Plaut, D. C. (2020). Hemispheric Organization for Visual Object Recognition: A Theoretical Account and

Empirical Evidence. *Perception, 49*(4), 373–404. https://doi.org/10.1177/0301006619899049

Bouhali, F., Bézagu, Z., Dehaene, S., & Cohen, L. (2019). A mesial-to-lateral dissociation for orthographic processing in the visual cortex. *Proceedings of the National Academy of Sciences, 201904184*. https://doi.org/10.1073/pnas.1904184116

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. https://doi.org/10.1371/journal.pone.0259986

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. https://doi.org/10.1038/s41598-017-01917-8

Burns, E. J., & Bukach, C. M. (2021). Face processing predicts reading ability: Evidence from prosopagnosia. *Cortex, 145*, 67–78. https://doi.org/10.1016/j.cortex.2021.03.039

Burns, E. J., & Bukach, C. M. (2022). Face processing still predicts reading ability: Evidence from developmental prosopagnosia. A reply to Gerlach and Starrfelt (2022). *Cortex, 154*, 340–347. https://doi.org/10.1016/j.cortex.2022.06.008

Busey, T., & Vanderkolk, J. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research, 45*, 431–448. https://doi.org/10.1016/j.visres.2004.08.021

Campbell, J. I., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods, 44*, 1255–1265. https://doi.org/10.3758/s13428-012-0186-0

Chen, H., Bukach, C. M., & Wong, A. C. (2013). Early electrophysiological basis of experience-associated holistic processing of Chinese characters. *PloS one, 8*(4), e61221. https://doi.org/10.1371/journal.pone.0061221

Chua, K. W., Richler, J. J., & Gauthier, I. (2014). Becoming a Lunari or Taiyo expert: Learned attention to parts drives holistic processing of faces. *Journal of Experimental Psychology: Human Perception and Performance, 40*(3), 1174–1182. https://doi.org/10.1037/a0035895

Chua, K. W., Richler, J. J., & Gauthier, I. (2015). Holistic processing from learned attention to parts. *Journal of Experimental Psychology: General, 144*(4), 723–729. https://doi.org/10.1037/xge0000063

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. *NeuroImage, 22*(1), 466–476. https://doi.org/10.1016/j.neuroimage.2003.12.049

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. https://doi.org/10.1080/13506285.2017.1370430

Curby, K. M., Entenman, R. J., & Fleming, J. T. (2016). Holistic face perception is modulated by experience-dependent perceptual grouping. *Attention, Perception, & Psychophysics, 78*(5), 1392–1404. https://doi.org/10.3758/s13414-016-1077-8

Curby, K. M., & Moerel, D. (2019). Behind the face of holistic perception: Holistic processing of Gestalt stimuli and faces recruit overlapping perceptual mechanisms. *Attention, Perception & Psychophysics, 81*(8), 2873–2880. https://doi.org/10.3758/s13414-019-01749-w

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience, 16*(4), 234–244. https://doi.org/10.1038/nrn3924

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., & Cohen, L. (2004). Letter binding and invariant recognition of masked words: Behavioral and neuroimaging

evidence. *Psychological Science, 15*(5), 307–313. https://doi.org/10.1111/j.0956-7976.2004.00674.x

Dehaene, S., Naccache, L., Cohen, L., Bihan, D., LeMangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience, 4*(7), 752.

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. https://doi.org/10.1126/science.1194140

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*(2), 107–117. https://doi.org/10.1037/0096-3445.115.2.107

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*, 348–358. https://doi.org/10.1037/a0029503

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. https://doi.org/10.1016/j.neuropsychologia.2014.05.006

Farah, M. J. (1991). Patterns of cooccurrence among the associate agnosias: Implications for visual object representation. *Cognitive Neuropsychology, 8*, 1–19. https://doi.org/10.1080/02643299108253364

Farah, M. J. (1992). Agnosia. *Current Opinion in Neurology, 2*, 162–164. https://doi.org/10.1016/0959-4388(92)90005-6

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*(3), 482–498. https://doi.org/10.1037/0033-295X.105.3.482

Feizabadi, M., Albonico, A., Starrfelt, R., & Barton, J. (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

Furubacke, A., Albonico, A., & Barton, J. S. (2020). Alternating dual-task interference between visual words and faces. *Brain Research, 1746*. https://doi.org/10.1016/j.brainres.2020.147004

Gabay, Y., Dundas, E., Plaut, D., & Behrmann, M. (2017). Atypical perceptual processing of faces in developmental dyslexia. *Brain and language, 173*, 41–51. https://doi.org/10.1016/j.bandl.2017.06.004

Gao, Z., Flevaris, A. V., Robertson, L. C., & Bentin, S. (2011). Priming global and local processing of composite faces: Revisiting the processing-bias effect on face perception. *Attention, Perception, & Psychophysics, 73*(5), 1477–1486. https://doi.org/10.3758/s13414-011-0109-7

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition, 103*(2), 322–330. https://doi.org/10.1016/j.cognition.2006.05.003

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience, 6*(4), 428–432. https://doi.org/10.1038/nn1029

Gerlach, C., Kühn, C. D., Poulsen, M., Andersen, K. B., Lissau, C. H., & Starrfelt, R. (2022). Lateralization of word and face processing in developmental dyslexia and developmental prosopagnosia. *Neuropsychologia, 170*, 108208. https://doi.org/10.1016/j.neuropsychologia.2022.108208

Gerlach, C., & Starrfelt, R. (2022). Face processing does not predict reading ability in developmental prosopagnosia: A commentary on Burns & Bukach (2021). *Cortex, 154*, 421–426. https://doi.org/10.1016/j.cortex.2022.02.003

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes, 23*(1), 1–35. https://doi.org/10.1080/01690960701578013

Hagen, S., Lochy, A., Jacques, C., Maillard, L., Colnat-Coulbois, S., Jonas, J., & Rossion, B. (2021). Dissociated face-and word-selective intracerebral responses in the human ventral occipito-temporal cortex. *Brain Structure and Function, 226*(9), 3031–3049. https://doi.org/10.1007/s00429-021-02350-4

Harel, A. (2016). What is special about expertise? Visual expertise reveals the interactive nature of real-world object recognition. *Neuropsychologia, 83*, 88–99. https://doi.org/10.1016/j.neuropsychologia.2015.06.004

Hsiao, J. H. W., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science, 19*(10), 998–1006. https://doi.org/10.1111/j.1467-9280.2008.02191.x

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of neuroscience, 17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Kinchla, R. A. (1974). Detecting target elements in multielement arrays: A confusability model. *Perception & Psychophysics, 15*(1), 149–158. https://doi.org/10.3758/BF03203263

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

Kühn, C. D., Wilms, I. L., Dalrymple, K. A., Gerlach, C., & Starrfelt, R. (2021). Face recognition in beginning readers: Investigating the potential relationship between reading and face recognition during the first year of school. *Visual Cognition, 29*(4), 213–224. https://doi.org/10.1080/13506285.2021.1884151

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 53A*(2), 513–536. https://doi.org/10.1080/027249800390583

Lerma-Usabiaga, G., Carreiras, M., & Paz-Alonso, P. M. (2018). Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *Proceedings of the National Academy of Sciences, 115*(42), E9981–E9990. https://doi.org/10.1073/pnas.1803003115

Liu, T. T., & Behrmann, M. (2014). Impaired holistic processing of left-right composite faces in congenital prosopagnosia. *Frontiers in Human Neuroscience, 8*, 750. https://doi.org/10.3389/fnhum.2014.00750

Liu, T. T., Chuk, T. Y., Yeh, S.-L., & Hsiao, J. H. (2016). Transfer of perceptual expertise: The case of simplified and traditional Chinese character recognition. *Cognitive Science, 40*(8), 1941–1968. https://doi.org/10.1111/cogs.12307

Liu, T. T., Hayward, W. G., Oxner, M., & Behrmann, M. (2014). Holistic processing for left–right composite faces in Chinese and Caucasian observers. *Visual Cognition, 22*(8), 1050–1071. https://doi.org/10.1080/13506285.2014.944613

Liu, T. T., Nestor, A., Vida, M. D., Pyles, J. A., Patterson, C., Yang, Y., Yang, F. N., Freud, E., & Behrmann, M. (2018). Successful reorganization of category-selective visual cortex following occipito-temporal lobectomy in childhood. *Cell Reports, 24*(5), 1113–1122. https://doi.org/10.1016/j.celrep.2018.06.099

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*, 255–260. https://doi.org/10.1016/S1364-6613(02)01903-4

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*, 375–407. https://doi.org/10.1037/0033-295X.88.5.375

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*(3), 353–383. https://doi.org/10.1016/0010-0285(77)90012-3

Paap, K. R., Newsome, S. L., & Noel, R. W. (1984). Word shape's in poor shape for the race to the lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 10*(3), 413–428. https://doi.org/10.1037//0096-1523.10.3.413

Plaut, D. C., & Behrmann, M. (2011). Complementary neural representations for faces and words: A computational exploration. *Cognitive Neuropsychology, 28*, 251–275. https://doi.org/10.1080/02643294.2011.609812

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

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin, 140*(5), 1281–1302. https://doi.org/10.1037/a0037004

Richler, J. J., Wong, Y. K., & Gauthier, I. (2011). Perceptual expertise as a shift from strategic interference to automatic holistic processing. *Current Directions in Psychological Science, 20*(2), 129–134. https://doi.org/10.1177/0963721411402472

Richler, J. J., Palmeri, T. J., Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Perception Science.* https://doi.org/10.3389/fpsyg.2012.00553

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. https://doi.org/10.3389/fpsyg.2017.01547

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition, 21*(2), 139–253. https://doi.org/10.1080/13506285.2013.772929

Rossion, B., & Lochy, A. (2021). Is human face recognition lateralized to the right hemisphere due to neural competition with left-lateralized visual word recognition? A critical review. *Brain Structure and Function, 1-31.* https://doi.org/10.1007/s00429-021-02370-0

Rossion, B., Dricot, L., Devolder, A., Bodart, J. M., Crommelinck, M., De Gelder, B., & 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. https://doi.org/10.1162/089892900562606

Royer, J., Blais, C., Charbonneau, I., Déry, K., Tardif, J., Duchaine, B., et al. (2018). Greater reliance on the eye region predicts better face recognition ability. *Cognition, 181*, 12–20. https://doi.org/10.1016/j.cognition.2018.08.004

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. https://doi.org/10.1080/02643294.2016.1204281

Saygin, Z., Scott, T., Feather, J., Youssoufian, D., Fedorenko, E., & Kanwisher, N. (2015). The VWFA and FFA have sharply contrasting functional selectivities and patterns of connectivity. *Journal of Vision, 15*(12), 914. https://doi.org/10.1167/15.12.914

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523–568. https://doi.org/10.1037/0033-295X.96.4.523

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. https://doi.org/10.1016/j.cognition.2018.02.017

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), 739–750. https://doi.org/10.1037/neu0000188

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), 138–147. https://doi.org/10.1037/neu0000428

Susilo, T., & Duchaine, B. (2013). Advances in developmental prosopagnosia research. *Current Opinion in Neurobiology, 23*(3), 423–429. https://doi.org/10.1016/j.conb.2012.12.011

Susilo, T., Wright, V., Tree, J. J., & Duchaine, B. (2015). Acquired prosopagnosia without word recognition deficits. *Cognitive Neuropsychology, 32*(6), 321–339. https://doi.org/10.1080/02643294.2015.1081882

Thesen, T., McDonald, C. R., Carlson, C., Doyle, W., Cash, S., Sherfey, J., Felsovalyi, O., Barr, W., Devinsky, O., ,Kuzniecky, R., & Halgren, E. (2012). Sequential then interactive processing of letters and words in the left fusiform gyrus. *Nature Communications, 3*, 1284. https://doi.org/10.1038/ncomms2220

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron, 55*, 143–156. https://doi.org/10.1016/j.neuron.2007.05.031

Ventura, P. (2014). Let's face it: Reading acquisition, face and word processing. *Frontiers in Psychology, 5*, 787. https://doi.org/10.3389/fpsyg.2014.00787

Ventura, P., Fernandes, T., Leite, I., Pereira, A., & Wong, A. C. N. (2019). Is holistic processing of written words modulated by phonology? *Acta Psychologica, 201*, 102944. https://doi.org/10.1016/j.actpsy.2019.102944

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.

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., 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*, 2189–2204. https://doi.org/10.3758/s13414-021-02287-0

Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes. *Cognitive Science, 28*(2), 289–301. https://doi.org/10.1207/s15516709cog2802_8

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1*, 59–85. https://doi.org/10.1016/0010-0285(70)90005-8

Wong, A. C.-N., & Gauthier, I. (2007). An analysis of letter expertise in a levels-of-categorization framework. *Visual Cognition, 15*(7), 854–879. https://doi.org/10.1080/13506280600948350

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. https://doi.org/10.1111/j.1467-9280.2009.02430.x

Wong, A. C. N., Qu, Z., McGugin, R. W., & Gauthier, I. (2011b). Interference in character processing reflects common perceptual expertise across writing systems. *Journal of Vision, 11*(1), 15, 1-8. https://doi.org/10.1167/11.1.15

Wong, A. C. N., Bukach, C. M., Yuen, C., Yang, L., Leung, S., & Greenspon, E. (2011a). Holistic processing of words modulated by reading experience. *PLoS ONE, 6*(6). https://doi.org/10.1371/journal.pone.0020753

Wong, A. C. N., Bukach, C. M., Hsiao, J., Greenspon, E., Ahern, E., Duan, Y., & Lui, K. F. H. (2012a). Holistic processing as a hallmark of perceptual expertise for nonface categories including Chinese characters. *Journal of Vision, 12*(13), 1–15. https://doi.org/10.1167/12.13.7

Wong, Y. K., Folstein, J. R., & Gauthier, I. (2012b). The nature of experience determines object representations in the visual system. *Journal of Experimental Psychology: General, 141*(4), 682–698. https://doi.org/10.1037/a0027822

Wong, A. C.-N., Zhiyi, Q., McGugin, R. N. W., & Gauthier, I. (2011c). Interference in character processing reflects common perceptual expertise across writing systems. *Journal of Vision, 11*(1), 15. https://doi.org/10.1167/11.1.15

Xu, Y. (2005). Revisiting the role of the fusiform face area in visual expertise. *Cerebral Cortex, 15*(8), 1234–1242. https://doi.org/10.1093/cercor/bhi006

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configural information in face perception. *Perception, 16*, 747–759.