The relationships between reading fluency and different measures of holistic word processing

Paulo Ventura[1] · Helen W.-Y. Tse[2] · José C. Guerreiro[1] · João Delgado[1] · Miguel F. Ferreira[1] ·
António Farinha-Fernandes[1] · Bruno Faustino[1,3] · Alexandre Banha[4] · Alan C.-N. Wong[2]

Abstract

Recently, paradigms in the face recognition literature have been adopted to reveal holistic processing in word recognition. It is unknown, however, whether different measures of holistic word processing share similar underlying mechanisms, and whether fluent word reading relies on holistic word processing. We measured holistic processing effects in three paradigms (composite, configural sensitivity, part–whole) as well as in reading fluency (3DM task: reading aloud high- and low-frequency words and pseudowords). Bin scores were used to combine accuracy and response time variables in the quest for a more comprehensive, reliable, and valid measure of holistic processing. Weak correlations were found between the different holistic processing measures, with only a significant correlation between the configural sensitivity effect and part–whole effect ($r = .32$) and a trend of a positive correlation between the word composite effect and configural sensitivity effect ($r = .21$). Of the three holistic processing measures, only one (part–whole effect) correlated with a lexical access measure of 3DM ($r = .23$). We also performed a principal component analysis (PCA) of performance in the three lists of 3DM, with the second most probably reflecting lexical access processes. There was a tendency for a positive correlation between part–whole bin measure and Component 2 of PCA. We also found a positive correlation between composite aligned in accuracy and Component 2 of PCA.

Our results show that different measures of holistic word processing reflect predominantly different mechanisms, and that differences among normal readers in word reading do not seem to depend highly on holistic processing.

Keywords Perceptual expertise · Individual differences · Holistic processing · Word recognition · Face recognition

---

Public significance

- Recently, holistic processing has been considered for visual word recognition. Words are made of very similar, and confusing, limited sets of letters and thus, like for other objects of visual expertise, holistic processing is believed to promote individuation of these visually similar words.
- Holistic measures in face recognition literature have been extended to the holistic processing of words. We measured holistic processing effects in three paradigms (composite, configural sensitivity, part–whole) as well as in a reading fluency task.
- Different holistic measures don´t seem to reflect similar mechanisms
- Relations between holistic measures and fluency tasks were not strong

✉ Paulo Ventura
paulo.ventura@gmail.com

[1] Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisbon, Portugal

[2] Department of Psychology, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

[3] HEI-Lab, Lusófona University, Lisboa, Portugal

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

Introduction

Holistic processing, or the obligatory attention to all parts of an object., is believed to promote individuation of visually similar objects, such as faces (Richler et al., 2012). As increased attention has been devoted to visual word recognition under a perceptual expertise framework (e.g., Ventura, 2014; Wong & Gauthier, 2007), there have been discussions on whether holistic processing can also be a marker of perceptual expertise in word recognition (Gauthier et al., 2010; Wong & Gauthier, 2007). On the one hand, Farah and colleagues (Farah, 1991, 1992; Farah et al., 1998; Tanaka & Farah, 1993) postulated that word perception (part-based) and face perception (holistic processing) are the two extremes of a continuum of visual object recognition. It has long been shown that words are not processed simply at the whole word level with the representations of letter identities bypassed. For example, Pelli et al. (2003) provided evidence that identifying a word implies feature detection and word holistic processing is not just about supra letter features (e.g., Paap et al., 1984). On the other hand, word recognition has been suggested to involve similarly challenging discrimination as face recognition, as typical reading experiences involve rapid identification of a large number of similarly looking words formed by arranging a fixed number of alphabetic symbols (Kleinschmidt & Cohen, 2006; Wong et al., 2011).

Similar to faces and other nonface categories of expertise, behavioral (Chen et al., 2013; Ventura et al., 2017; Wong et al., 2011; Wong et al., 2012) and brain (Cai et al., 2020), research has suggested the involvement of holistic representations in visual word recognition. In addition, recent studies begin to show a relationship between measures of holistic word processing and word recognition ability (Conway et al., 2017; Ventura et al., 2020; Wong et al., 2019). Two important issues remain unclear. First, do different measures of holistic processing of words share similar underlying mechanisms? Second, could holistic word processing have some sort of functional significance for word recognition, as suggested by the correlation between them? In this study, the relationships between different measures of holistic processing of words, as well as their relationships with word reading ability, were examined. As many paradigms used for measuring holistic processing has been adopted from the face recognition literature (Tanaka & Farah, 1993; Yin, 1969; Young et al., 1987), we will first discuss the precise tasks developed in the domain of face recognition and then their use in the realm of word recognition.

Paradigms measuring holistic processing

Three experimental paradigms have come to be widely regarded as standard measures of holistic processing in the face recognition literature (Murphy et al., 2017; Rezlescu et al., 2017; Richler et al., 2012.). These include the composite task, the configural sensitivity task, and the part–whole task. The first two tasks were adopted in the study of holistic word processing in recent years, while the part–whole task has been linked to the classical paradigm showing the part–whole effect.

Composite task (Fig. 1) typically requires subjects to judge if the target parts of two objects (e.g., left half of two words, top half of two faces) are the same or different, while ignoring the irrelevant parts (e.g., right half of the words, bottom half of two faces). Despite the instruction to focus on the target part only, performance is often affected by the irrelevant parts. Performance is better when the irrelevant parts would lead to consistent response as the target part (i.e., 'same' for both target and distractor parts, or 'different' for both) than when the target and distractor parts would lead to inconsistent responses (i.e., 'same' for target parts and 'different' for distractor parts, or vice versa). Typically, this congruency effect is reduced by misaligning the target and distractor parts and thus disrupting the overall configuration. Composite effects have been described for fluent readers with English words (Wong et al., 2011), Chinese characters (Chen et al., 2016; Wong et al., 2012), and Portuguese words (Ventura et al., 2017). The difficulty in limiting our attention to one object part apparently reflects obligatory attention to all parts of an object. For faces, it has been suggested that such a perceptual strategy may have become automatized with experience and/or due to a history of learned attention to diagnostic parts (Chua et al., 2014; Richler et al., 2012; Richler, Wong, & Gauthier, 2011b; see also Richler & Gauthier, 2014, for a review; but see Rossion, 2013, and Gauthier & Bukach, 2007, for different designs and interpretations).

Configural sensitivity task (Fig. 2), often called the *inversion paradigm*, requires an observer to judge whether a pair of objects differ in terms of small configural relations (e.g., jittering of letters in a word, or distance between the nose and mouth in a face). Performance is typically higher for upright than inverted objects, and such an inversion effect is larger for stimuli differing in configural relations than in the shape of individual components (Diamond & Carey, 1986; Leder & Bruce, 1998; Le Grand et al., 2001; Mondloch et al., 2002). A common interpretation of the effect for faces is that representations about individual parts and their configuration are used to recognize faces, with configural information particularly dominant for upright faces (Richler et al., 2012). Configural sensitivity thus reflects the explicit representation of spatial relations between features (e.g., Diamond & Carey, 1986; Leder & Bruce, 2000; Rhodes et al., 1993; Searcy & Bartlett, 1996).

For words, many studies have shown an association between one's expertise with a writing system and the inversion effect found in lexical decision or reading aloud, across alphabetic and nonalphabetic languages (Björnström et al., 2014; Conway et al., 2017; Kao et al., 2010; Koriat & Norman,
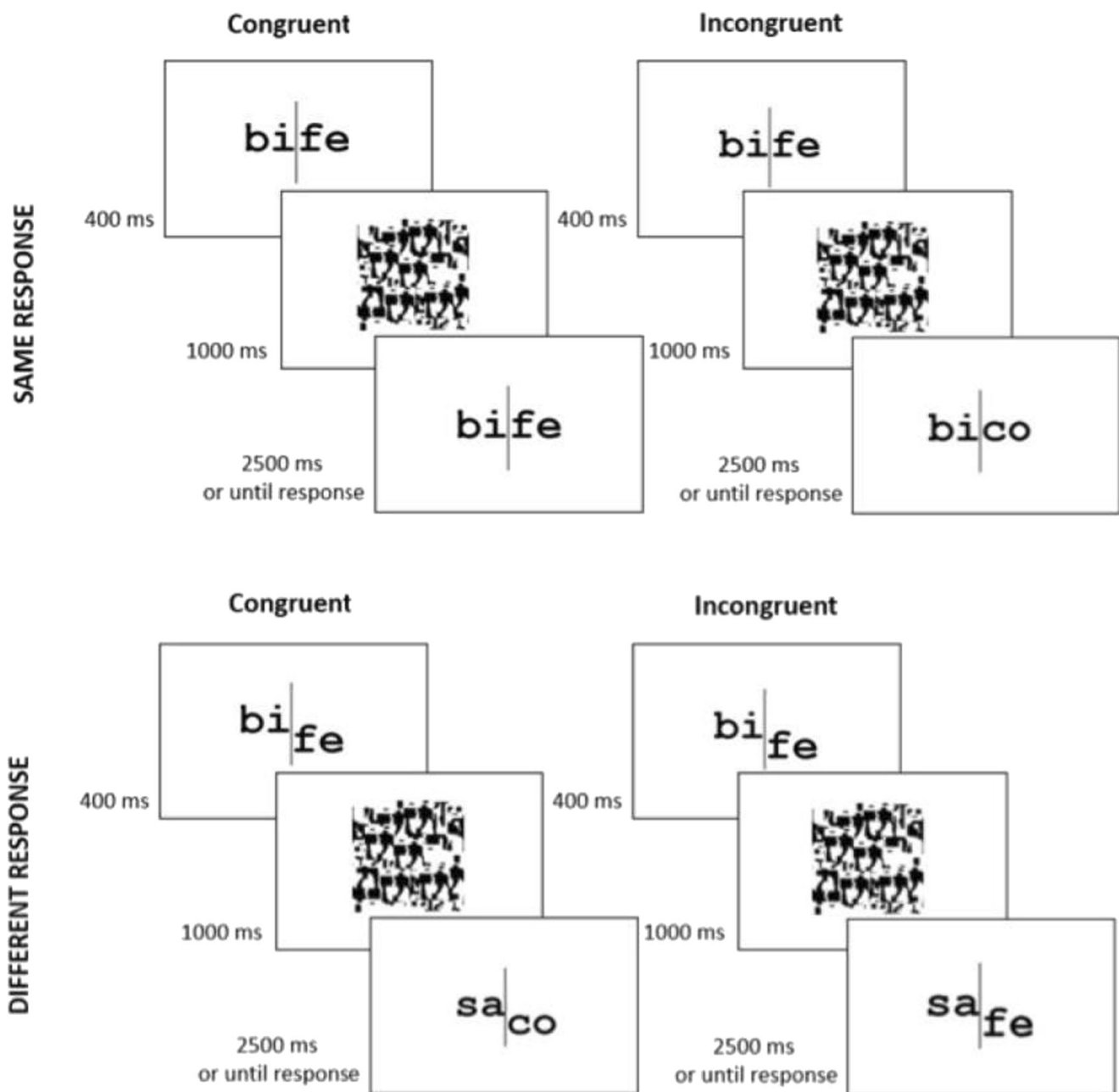
Fig. 1 The composite task with Portuguese words. In this example, participants should judge if the left half of two sequentially presented words are the same or different. An example trial is shown for each of the Congruency × Response conditions

1985). In these studies, however, the stimuli in the different trials involved different words and thus confound featural and configural changes. Wong et al. (2019) evaluated the processing of featural and configural information separately and found a larger inversion effect in fluent readers for words involving configural changes than component changes, similar to what has been found for faces (Le Grand et al., 2001; Rakover, 2013). Therefore, it has been suggested that representations involving configural information are particularly dominant compared with componential information for upright words (Wong et al., 2019).

Finally, in *part–whole task* (Fig. 3), it is easier to perceive a feature (e.g., a letter, eyes) when it is presented on a whole object (e.g., a word or face) than when it is presented separately, indicating facilitation of the whole on part processing (Tanaka & Farah, 1993; Tanaka & Simonyi, 2016). The common interpretation is that object parts are not represented in isolation but instead integrated as a larger chunk or unit possibly covering the whole object. The part–whole task has been implemented slightly differently in the face and word recognition literature. For faces, the effect has been typically shown in terms of the better performance for parts shown in a whole
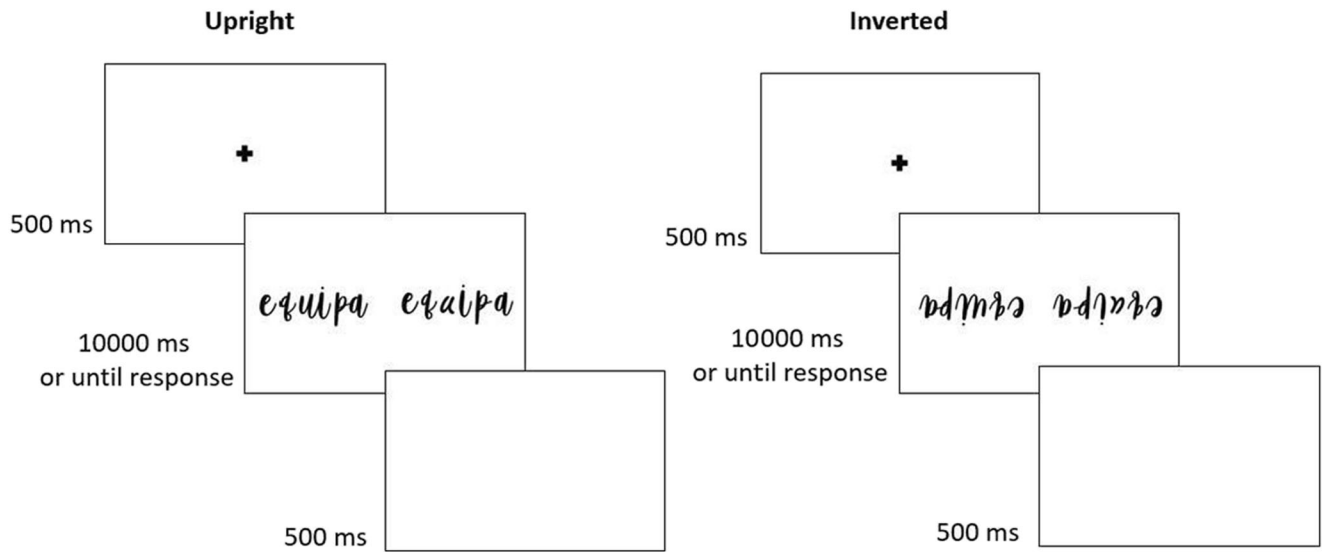
A

B

Fig. 2 The configural sensitivity task with Portuguese words. **a** Two versions of a word with different spatial relationship between letters as used in a different trial. **b** Sequence of events in an upright trial and an inverted trial
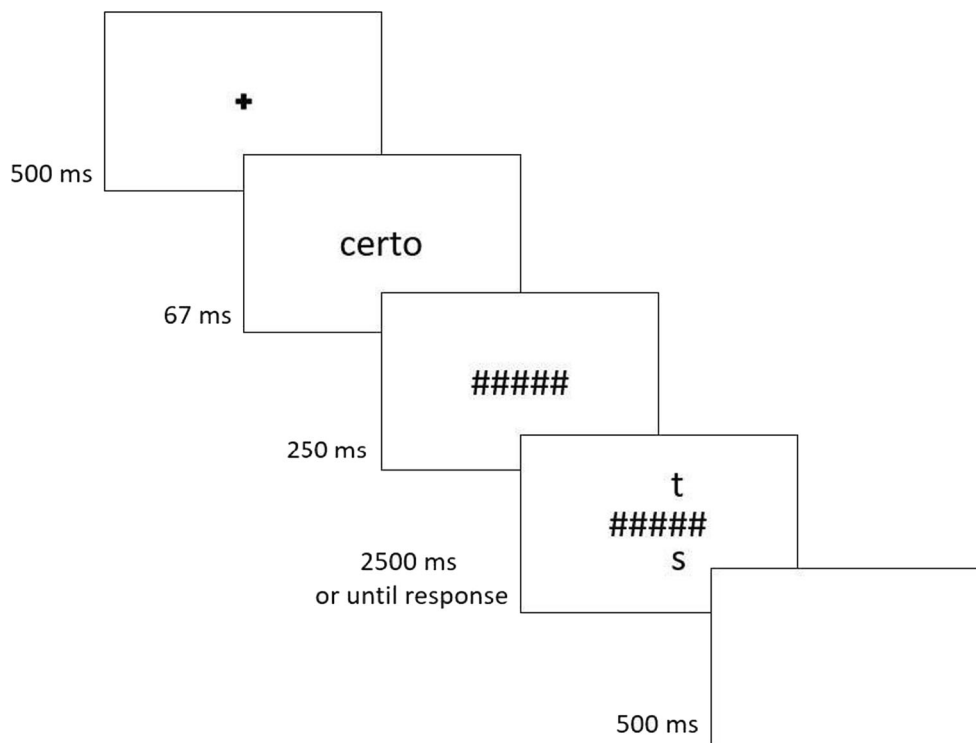

Fig. 3 The part–whole task with Portuguese words

face than when presented in isolation (Tanaka & Farah, 1993), although a similar advantage has also been shown for parts shown in an intact face than in a new face configuration (Tanaka & Sengco, 1997). A similar task for words can be found in the classic Reicher–Wheeler paradigm (Reicher, 1969; Wheeler, 1970). Observers viewed a briefly presented letter string that is subsequently masked and had to identify the letter at a specific location in the string. Accuracy is typically higher when the letter is within a word (e.g., wine) than a pseudoword (e.g., kwar) or a nonword (e.g., yegb). The advantage of the word context has been interpreted as a top-down influence of whole- word representations at the orthographic level on letter identification level (McClelland & Rumelhart, 1981). Therefore, the part–whole effect for words and faces involves similar proposed mechanisms of whole-object level representations facilitating individual part processing.

Relationships between holistic processing measures and word reading

While holistic processing of words has received increasing attention, two important questions arise. The first question concerns whether the holistic processing effects found in different paradigms share similar underlying mechanisms. Referring all these effects as "holistic processing" effects implies that these effects indicate the same or highly overlapping holistic processing mechanisms. However, examination of the mechanistic accounts proposed for the different holistic processing effects suggests both similarities and differences. For example, configural sensitivity and the part–whole effect could involve similar mechanisms, as both involve the processing of letters and word parts assisted by higher-level processing. In the case of configural sensitivity, multi-letter or whole-word representations can contain information about word shape and spatial relationships between word parts that would be useful for the extraction of letter identities from extrinsic features like font (Gauthier et al., 2006; Sanocki, 1988). In the case of the part–whole effect, whole-word orthographical representations may facilitate, via feedback, mechanisms earlier in letter processing (McClelland & Rumelhart, 1981). While the higher-level representations for the two effects may be of different nature (visual vs. lexical), their top-down effects on letter processing could constitute an overlap. For the composite effect, some interpret it as reflecting holistic visual representations (e.g., Young et al., 1987), and it would be natural to link this to the multi-letter or whole-word representations underlying configural sensitivity. To others who regard the composite effect as reflecting compulsory attention to all parts of an object (Richler et al., 2011a, b), such compulsory attention does not necessarily mean enhanced representations at the whole object level or enhanced sensitivity to the configural relationships between parts. Inspirations can

also be drawn from recent studies in the face recognition literature. Rezlescu et al. (2017) examined the relationships between composite, inversion, and part–whole effects. Intriguingly, the three holistic processing measures did not show a high overlap, with a significant and moderate correlation found only between the inversion and part–whole effects ($r = .28$).

The second question concerns whether holistic word processing is associated with word recognition. Previous studies have shown that the experience of discriminating between highly similar novel objects resulted in larger holistic processing, as compared with the experience of coarser classification of the objects (Wong et al., 2009). An intriguing possibility is that holistic processing may be associated with or even contributes to efficient individuation of objects as required in word recognition. Empirically there is evidence for such an association. For example, Ventura et al. (2020) found that among normal readers, the magnitude of the word composite effect correlated with performance in lexical decision, where participants judged if a string of letters corresponded to a real or invented word (across participants, the larger the composite effect, the smaller the frequency effect, $r = -.25$). Responses were faster for high- than low-frequency words, and this frequency effect was reduced for people with more experience with a particular language, as they performed well with both high- and low-frequency words (e.g., Yap et al., 2012).

In addition, Wong et al. (2019) found a positive correlation between configural sensitivity and word recognition fluency. Despite these findings, there remains to be seen a study where different holistic processing effects are considered together. This kind of study is also lacking in the face recognition literature, with the only exception being Rezlescu et al. (2017), which examined the relationship between three measures of holistic processing with face recognition ability evaluated by Cambridge Face Perception Test (CFPT; Duchaine et al., 2007), with correlations ranging from moderate ($r = .42$ for the inversion effect), to weak ($r = .25$ for the part–whole effect) and nonexistent ($r = .04$ for the composite effect). Another set of studies adopting the Vanderbilt Holistic Processing Test with faces (VHPT-F), a highly reliable composite measure that assesses individual variations in holistic processing, did not find any correlation between holistic face processing and face recognition ability (Richler et al., 2014, 2015; Verhallen et al., 2017). These results cast doubt on the role of holistic processing on efficient individuation of objects.

The current study

Up to now, these three different holistic word processing measures have never been used in the same study. In the present study we, first, evaluated the extent to which different word holistic tasks reflect a common underlying holistic phenomenon. The composite effect, configural sensitivity, and the

part–whole task were measured in the same participants. Second, we examined the potential link between different holistic processing effects and reading ability measured in a fluency task called the 3DM test (e.g., Fernandes et al., 2017). The 3DM reading test is a time-limited reading-aloud task composed of three lists of high- and low-frequency words and pseudowords. The number of correctly read items reflects performance. This reading fluency test has high test–retest reliability (Pacheco et al., 2014) and does not have the issue of ceiling performance in college students. In addition, its scores correlate with silent reading measures and reading comprehension measures in college students and discriminate between good and bad adult college readers (Fernandes et al., 2017), making it a good measure of reading ability.

The 3DM reading test also shows effects considered to reflect fast and efficient access to lexical orthographic representations. One is the word-frequency effect, reflecting efficient access to the lexicon, with significantly more high-frequency words read in 30 s than low-frequency words (Fernandes et al., 2017; Lima & Castro, 2010). Another one is the lexicality effect, with better performance for high-frequency and low-frequency words than for pseudowords (Fernandes et al., 2017), reflecting access to lexical orthographic representations. We thus used two different indexes reflecting higher efficiency in activating the correct lexical representations. (i) frequency effect: [high frequency] − [low frequency]; and (ii) lexicality effect: [(high frequency) + (low frequency)]/2 − [pseudowords].

The overall score in 3DM reflects multiple components including verbal speed, general processing speed, and sublexical, phonological components. In order to extract the component more directly related to orthographical lexical access, we also did a principal component analysis (PCA) of performance in the three lists. As there are as many principal components as there are variables in the data, principal components are constructed in such a manner that the first principal component accounts for the largest possible variance in the data set, thus reflecting what is common to the three lists: verbal speed, general processing speed and sublexical, phonological processes. The second principal component is calculated in the same way, with the condition that it is uncorrelated with (i.e., perpendicular to) the first principal component and that it accounts for the next highest variance. Thus, this component should reflect what is left after partialling out verbal speed, general processing speed, and sublexical, phonological processes: that is, lexical access processes. In order to confirm this logic, we performed correlation analyses between the two PCA components and: (i) the three lists in 3DM, (ii) the sum of low frequency and pseudoword lists, and two difference indexes reflecting higher efficiency in activating the correct lexical representations., (iii) [high frequency] − [low frequency]; and (iv) [high frequency] − ([low frequency] + [pseudowords]).

Method

Participants

A convenience sample of 81 Portuguese-reading students studying psychology in a university was recruited on a voluntary basis in exchange for course credit in one semester. This sample size would allow us to detect a significant correlation of .366 at α = .05 with a power of .94 according to G*Power (Version 3.1). This target correlation (.366) was determined by the geometric mean of the correlation coefficients found between holistic processing and recognition performance for faces (inversion: r = .42, part–whole: r = .25; Rezlescu et al., 2017), and that between configural sensitivity and recognition performance for words (r = −.469, Wong et al., 2019). Data from 10 additional participants were excluded either due to difficulties in recording the fluency task or for refusing that the fluency task was recorded. Data from another six additional participants were discarded because their performance in one or more tasks was too low (percent correct below .60). All participants had normal or corrected-to-normal vision and reported no difficulties in reading across development. This study followed the Declaration of Helsinki and Portuguese Deontological Regulation. Also, it was approved by the Deontological Committee of Faculdade de Psicologia of Universidade de Lisboa. Freely given informed consent to participate in the study was obtained from participants.

Material and procedure

Participants were tested in pairs. All participants completed the three holistic processing tasks and then the 3DM task, and the order of the three holistic tasks was randomized. The 3DM task was always the last task. A CRT monitor was used to present the stimuli. The stimulus presentation and the recording of response time were manipulated by E-Prime 2.0 (Schneider et al., 2012a, 2012b)

Word composite task The task was identical to that in Ventura et al.'s (2017) Experiment 1 (Fig. 1). Subjects were asked to distinguish whether the left halves (the target part) of two sequentially shown words were the same or different by pressing the "1" or "2" key (with a green or red label) as quickly and precisely as possible while neglecting the other part. After 2.5 s or upon the response, the stimuli disappeared. A blank screen then appeared for 500 ms, followed by the next trial. The task adopted the complete version of the composite paradigm (Richler & Gauthier, 2014), in which "congruent" trials refer to the same status between the correct response for the target part (same or different) and that for the irrelevant part (same or different), whereas "incongruent" trials refer to the different status between the correct responses for these parts. The target and irrelevant parts were aligned with each other in

half of the trials and misaligned in the other half. Holistic processing is typically inferred from an interaction between alignment and congruency (e.g., Richler et al., 2011b): the performance advantage in the congruent over incongruent trials is usually greater in the aligned than the misaligned condition.

Words were generated in sets of four (24 word sets) four-letter (consonant- vowel.consonant-vowel) CV.CV Portuguese words (Ventura et al., 2017; cf. the present study Appendix 1). Each set of four words (e.g., bife (steak), bico (beak), safe (get way), saco (bag)) contained two pairs of words with interchangeable initial syllables (e.g., bife and bico share the initial syllable "bi" and safe and saco share the initial syllable "sa") and two pairs of words with interchangeable final syllables (e.g., bife and safe share the final syllable "fe" and bico and saco share the final syllable "co"). All possible congruency by response pairings can be created by re-pairing the words in a set, thus ensuring each word to appear equally frequently in all four Congruency × Same/Different conditions.

Thus, frequency, syllabic structure, or other word-specific factors was not an issue for the composite task.

Although frequency is not an issue in our composite task, we nevertheless estimated the average frequency of the words used. We employed the standardized Zipf scale measure (obtained from P-PAL; Soares et al., 2018), which was computed by adding 3 to the log10 of the per-million-word frequency (see van Heuven et al., 2014, for details). Zipf scale is considered as a simpler and more intuitive way to capture the word frequency distribution. While words with a Zipf value below 4 are perceived as low-frequency words (with frequencies of 1 per million words or below), words with a Zipf value ranging from 4 to 10 are perceived as high-frequency words (with frequencies of 10 per million words or higher). The words used in the composite task had a mean Zipf value of 4.4.

Each word was divided into two halves by a vertical line. The two halves of the words were interchanged to create the four conditions in the complete design. Each word was shown in courier font with a 20-point size, (3.44° × 1.04° at a viewing distance of 90 cm on a 17-in. screen). This resulted in a vertical span of 1.66° instead of 1.04° for the words in the misaligned condition. There were two aligned and two misaligned blocks of trials with 96 trials each (six sets of words). At the start of the session, the experimenter showed four examples on paper and gave feedback on the correct response. Participants then completed 16 practice trials with different stimuli on the computer and with no feedback.

Configural sensitivity task This task was similar to the one used in Wong et al. (Wong et al., 2019; Fig. 2). Firstly, a 500-ms fixation cross shown at the center of the screen. Two words then appeared side by side from a viewing distance of 90 cm. The same identity and font always applied to the two

words of each trial. In half the trials, however, these words differed in the spatial relationships (vertical jittering). To indicate whether these words were visually indistinguishable, subjects had to press the green labelled "1" key or the red labelled "2" key, respectively, as accurately and quickly as possible. After 10s or upon response, the word pair faded, followed by a 500-ms blank screen before next trial.

Forty Portuguese words were used (Appendix 2). The Portuguese words consisted of 5 to 6 lowercase letters, with a Zipf value of 4.91. Two versions of each word were generated, and they differed slightly in the vertical jittering of letters. The stimuli could appear in different font styles (arial or freestyle) and different orientations (upright or inverted: flipped both vertically and horizontally). Each word subtended an angle between 3.02° × .08° and 3.11° × 1.51° from a viewing distance of 90 cm. Each one of the 8 blocks contained 80 trials. The order of trials in different conditions was randomized. The block order was randomized.

Part–whole task On each trial (Fig. 3), a fixation cross for 500 ms was followed by a target letter string for 67 ms. A string of hashes of the same length was then shown immediately as a backward mask for 250ms. After that, two letters were displayed above and below a targeted position and participants had to decide as quickly and accurately which letter was presented on that position. Participants pressed the keys "1" and "2" if they thought it was the letter presented above or below the hash, respectively. After 2.5 s or upon response, another trial began with a 500 ms blank screen.

The targets were 27 words and 27 nonwords of three letters (Appendix 3; visual angle: .91° × .22° from a viewing distance of 90 cm), 40 words and 40 nonwords of four letters (visual angle: 1.15° × .3° from a viewing distance of 90 cm), and 40 words and 40 nonwords of five letters (visual angle: 1.4° × .3° from a viewing distance of 90 cm). The word targets had Zipf values of 5.2 (three letter-words), 4.84 (four letter-words), and 4.7 (five letter-words). Most of nonwords were unpronounceable, with a small number of pronounceable nonwords. They were presented in black lowercase courier-new 18-point font. The distractor letter matched with the target letter for feasibility. In the word condition the target and the distractor both made a word, and in the nonword condition the target and the distractor both made a nonword. We also computed the visual similarity between the two letters of each trial taking into consideration the status of the stimuli (word vs. nonword) and number of letters of the stimuli (3, 4, 5) by reference to letter visual similarity matrix for Latin-based alphabets of Simpson et al. (2013). The effect of number of letters was significant, $F(1, 210) = 4.97$, $p = .008$, $\eta_p^2 = .05$, indicating higher visual similarity between the two letters for longer strings (2.0, 2.15 and 2.5 for 3, 4, and 5 letter stimuli). Importantly, there was no difference between visual similarity of the letters for words vs. nonwords ($F < 1$) and no interaction between stimuli type and number of letters ($F < 1$)

Participants first received three examples on paper, with feedback from the experimenter. Next, there was practice with 16 trials. As for the experimental blocks, there were three different blocks, separately for 3-, 4- and 5-letter stimuli, each one containing an equal number of words and nonwords, presented randomly.

Performance would be compared between the word and nonword conditions, which was different from the first part–whole paradigm studies for faces (e.g., Tanaka & Farah, 1993), where whole-face and isolated-part conditions were compared. Yet apart from isolated parts, new face configurations formed by modifying spacing between parts have also been used in the control condition, showing the same effect (reviewed in Tanaka & Simonyi, 2016). Therefore, the use of isolated parts in the control condition is not a must for showing the effect, and we opted to follow the procedure of the classical paradigm showing the part–whole task (Reicher, 1969; Wheeler, 1970).

Reading fluency task—3DM In the 3DM reading fluency test participants are requested to read as many items as possible in 30 seconds. Each screen had 15 items for a total of 75 items for high- and low-frequency words and nonwords each. Difficulty level increased through the five screens for each material with respect to the number of syllables (2–4), syllabic structure (with and without consonant clusters), and phoneme–grapheme correspondence rules (regular and irregular). The pseudowords were derived from the high-frequency words, which were separated into syllables and then rearranged to form pseudowords.

The lists were presented in a fixed order. The experimenter asked the subjects to read aloud as soon as the stimuli were shown on the screen (always 15 items each time), with 30 seconds allotted per list. Also, they were required to read the items as accurately as possible. The number of correctly read items reflected their performance.

Results

The overall performance in the three holistic processing tasks and the reading fluency task are presented first, followed by the introduction of the bin scores for individual differences analyses, the reliability of the bin scores for different tasks, and the correlations between the three holistic measures and then the correlation between each holistic measure and the reading measure. Response time (RT) analyses were conducted for errorless trials in all holistic processing tasks. Trials with an RT 2.5 standard deviations above the average RT for each participant, or with an RT below 150 ms were discarded. That resulted in 3.2%, 2.2%, and 4.4% of the trials discarded for the composite, configural sensitivity, and the part–whole tasks, respectively. Accuracy was also evaluated.

For the reading fluency (3DM) task, performance was indicated by the number of items correctly pronounced within 30 seconds per list, but we were interested more specifically in frequency and lexicality effects and in the second principal component in a PCA including the three lists, a component reflecting what is left after partialling out verbal speed, general processing speed and sublexical, phonological processes: that is, lexical access processes.

Overall task performance

For the composite task, a 2 (Alignment) x 2 (Congruency) analysis of variance (ANOVA) on RT revealed a significant congruency effect, $F(1, 80) = 21.22$, $p < .001$, $\eta_p^2 = .21$, with better performance on congruent ($M = 527.6$, $SD = 104.2$) than on incongruent trials ($M = 536.7$, $SD = 108.3$). Alignment was not significant. Neither was the interaction of alignment and congruency ($Fs < 1$). The same 2×2 ANOVA on d' revealed a significant congruency effect, $F(1, 80) = 4.33$, $p < .05$, $\eta_p^2 = .05$, with better sensitivity on congruent ($M = 4.16$, $SD = .54$) than on incongruent trials ($M = 4.04$, $SD = .61$). There was a tendency for a significant alignment effect, $F(1, 80) = 3.12$, $p = .08$, $\eta_p^2 = .04$, with better sensitivity on misaligned ($M = 4.14$, $SD = .54$) than aligned trials ($M = 4.06$, $SD = .56$). There was also a tendency for a significant interaction of alignment and congruency, $F(1, 80) = 2.78$, $p = .09$, $\eta_p^2 = .03$, with a higher congruency effect for aligned ($M = .18$) than for misaligned trials ($M = .05$). In previous studies, the Alignment × Congruency effect occurred sometimes in terms of sensitivity and at other times in terms of RT. The non-significant trend in sensitivity could be due to the effect occurring in both sensitivity and RT to a different extent for different participants. Using bin scores to combine RT and accuracy may thus be even more important. In any case, for our purpose of individual difference analyses, it is more important to have a sufficient variability in the composite effect across participants than have everybody showing the composite effect to the same degree.

For the configural sensitivity task, paired-sample t-tests on RT and on accuracy both showed an inversion effect. Participants were faster for upright ($M = 1765.08$, $SD = 422.25$) than inverted trials ($M = 2127.42$, $SD = 612.66$), $t(80) = 14.40$, $p < .0001$.

Participants were also more accurate for upright ($M = .91$, $SD = .07$) than inverted trials ($M = .88$, $SD = .09$), $t(80) = 5.46$, $p < .001$.

For the part–whole task, paired-sample t-tests on RT and on accuracy both showed a context effect. Participants were faster in identifying a letter that had been presented in the context of a word ($M = 910.33$, $SD = 170.39$) than in the context of a nonword ($M = 960.11$, $SD = 200.09$), $t(80) = 5.59$, $p < .001$. Participants were also more accurate in identifying a letter in the context of a word ($M = .84$, $SD = .09$) than

in the context of a nonword ($M = .73$, $SD = .09$), $t(80) =$ 14.785, $p < .001$.

For the reading fluency (3DM) task, a one-way ANOVA showed a significant effect of list (type of item), $F(2, 160) =$ 303.0, $p < .001$, $\eta_p^2 = .79$. Similar to previous findings (Fernandes et al., 2017), participants read significantly more high-frequency words in 30 s ($M = 59.56$, $SD = 7.92$) than low-frequency words ($M = 54.22$, $SD = 8.80$), $t(80) = 7.33$, $p < .001$, thus evidence of a frequency effect. Lexicality was also significant, with higher performance for the average of high- and low-frequency words than pseudowords ($M = 56.89$, $SD = 7.71$ vs. $M = 41.48$, $SD = 7.31$), $t(80) = 22.81$, $p < .001$.

We also ran a PCA. The first principal component accounts for the largest possible variance in the data set, thus reflecting what is common to the three lists: verbal speed, general processing speed, and sublexical, phonological processes. The second principal component accounts for the next highest variance. Thus, this component should reflect what is left after partialling out verbal speed, general processing speed, and sublexical, phonological processes: that is, lexical access processes.

The first component correlated positively with high frequency list ($r = .84$, $p < .001$), low frequency list ($r = .93$, $p < .001$), nonword list ($r = .85$, $p < .001$) and low+nonword ($r = .96$, $p < .001$). Thus, the first component reflects what is common to all three lists. But this seems not to include a lexical access component, as attested by the negative correlation between the first component and [high frequency] − [low frequency]; ($r = −.23$, $p = .039$). The correlation between the first component and [high frequency] − ([low frequency] + [pseudowords]) was also not significant ($r = −.08$, $p = .458$).

The second component reflects lexical access. Indeed, it correlates positively with what reflects lexical access: high frequency list ($r = .51$, $p < .0001$), [high frequency] − [low frequency]; ($r = .64$, $p < .0001$), and (ii) [high frequency] − ([low frequency] + [pseudowords]); ($r = .93$, $p < .0001$). The second component correlates negatively with what reflects more sublexical, phonological processes: low frequency list ($r = −.02$, n.s.), nonword ($r = −.48$, $p < .0001$), low+nonword ($r = −.25$, $p = .026$).

Individual measures of holistic processing and reading fluency

The holistic processing tasks all involved RTs and accuracy measures.

Incorporating these two measures for individual differences research has always been an issue. Here a rank-ordering binning procedure was used to compute for each task in each participant a bin score that incorporated RTs and accuracy in one measure. The bin scores have been shown to result in higher reliability and validity than alternatives such as the inverse efficiency scores, or RTs and accuracy in isolation (Draheim et al.,

2016; Hughes et al., 2014). Therefore, the bin scores are used for the holistic processing tasks here. In general, results with bin scores agreed with the analyses of RTs and accuracy separately (Appendix 5), but see our comments below.

Computing the bin scores involves the following general steps: (1) For each participant, compute the average RT of all accurate trials in the base condition. (2) For each participant, subtract the average RT obtained in Step 1 from the RT of each accurate trial of the target condition. (3) Rank order the scores from Step 2 for all participants into deciles and assign a bin value ranging from 1 to 10 from the smallest to the largest decile. (4) Assign a bin value of 20 (twice as high as the value for the slowest accurate trial in the target condition) to all inaccurate trials in the target condition. (5) Sum the bin values from Steps 3 and 4 to obtain the bin score for each participant in each task. The bin score of a participant thus indicates how much worse the performance of the target condition is compared with the base condition, in comparison with other participants.

In this study, the bin scores were calculated for the three holistic tasks. For the word composite task, the aligned congruent condition was the base condition while the aligned incongruent condition was the target condition[1]. For the configural sensitivity task, the upright and inverted conditions corresponded to the base and target conditions, respectively. For the part–whole (word superiority) task, the word condition and nonword condition were regarded as the base and target conditions, respectively. For all three holistic processing tasks, a higher bin score corresponded to larger holistic processing.

Reliabilities were indicated by Guttman's λ^2 computed across different blocks in each holistic processing task (cf. Table 1). For the reading fluency (3DM) task, Guttman's λ^2 was computed across the three lists. Overall reliabilities were medium to high (all >.5), showing sufficient precision and variability for further examination of their relationships with each other. The reliabilities were comparable with those in the different holistic processing tasks used with faces in Rezlescu et al. (2017).

Relationships between different holistic processing measures

Table 2 and Fig. 4 show the Pearson-product correlations between the different holistic processing and reading fluency measures and the scatterplots, respectively.[2] The three holistic

---

[1] Here, we reported results with the composite effect calculated on the aligned trials only. The same pattern of correlations with other holistic processing measures and word reading fluency was found when the congruency effect in the misaligned trials was partialled out from the congruency effect in the aligned trials to indicate the composite effect.

[2] One bivariate outlier was identified for the correlations between the composite effect and all other measures ($p<.001$ in chi-square tests of Mahalanobis Distance) and thus data from one participant was discarded in all correlation analyses involving the composite effect. Reliability of the composite effect rose to .52 with this outlier discarded.

Measures	Guttman's λ^2
Word composite effect	0.52
Configural sensitivity effect	0.85
Part–whole (word superiority) effect	0.62
Reading fluency (3DM)	0.86

processing effects were only weakly correlated with each other (Figs. 4a–c). There was a significant correlation between the configural sensitivity effect and part–whole effect ($r = .32$, $p = .003$). The correlation between the composite and the part–whole effects was not significant ($r = .15$, $p = .19$). Finally, the correlation between the composite and the configural sensitivity task was marginally significant ($r = .21$, $p = .057$).

Considering raw data (cf. Appendix 5), we found the same pattern as the one obtained from the bin scores. The part–whole effect in RT was positively correlated with the configural sensitivity effect in RT ($r = .23$, $p = .04$), and negatively correlated with the configural sensitivity effect in accuracy ($r = -.25$, $p = .027$). There was also a positive correlation between the word composite effect in RT and configural sensitivity effect in RT ($r = .25$, $p = .027$). The correlation between the composite and the part- whole effect was not significant ($r = .11$, $p = .30$).

Relationships between holistic processing and reading fluency measures

There was little relationship between each holistic processing bin measure and reading (Table 2; Figs 4d–f), all rs < . 16

except for a small positive correlation between part–whole task bin measure and the lexicality effect ($r = .23$, $p = .04$).

Considering raw scores, and similar to what was found with bin scores, there was not much relationship between the holistic processing measures and reading fluency. All correlations were low (rs < .19), except for negative correlations between the part–whole effect in RT and the frequency effect in 3DM task performance ($r = -.22$, $p = .054$). and between the composite effect in RT and the frequency effect in 3DM task performance ($r = -.22$, $p = .048$).

Considering raw scores, we also found a tendency for a correlation between the part–whole task in RT and Component 2 of PCA: $r = .20$, $p = .076$. We also found a positive correlation between composite aligned in accuracy and Component 2 of PCA: $r = .26$, $p = .017$.

General discussion

In our study, we adopted an individual differences avenue to examine the relationships of different measures of word holistic processing (composite effect, configural sensitivity, and part–whole task) as well as their link with word reading fluency. Overall, the holistic processing measures were only weakly correlated.

## Separate mechanisms underlying different holistic processing measures

Correlations between the three holistic measures were low ($r = .14–.32$). The correlation between the word composite effect and the part–whole task was the lowest ($r = .16$), and insignificant. This is intriguing, given the apparently similar task format—both tasks involved matching of sequentially presented stimuli, and the matching was required only on a part of the

Table 2 Pearson-product correlations between the bin scores of the three holistic processing measures and the 3DM scores

		Composite effect	Configural sensitivity	Part–whole task
Composite effect	r			
	p			
Configural sensitivity	r	.21		
	p	.057		
Part–whole task	r	.16	.32**	
	p	.19	.003	
Frequency effect	r	−.05	−.02	.11
	p	.64	.85	.30
Lexicality effect	r	−.12	−.07	.23*
	p	.87	.55	.04

** Correlation is significant at the 0.01 level (two-tailed).

* Correlation is significant at the 0.05 level (two-tailed).
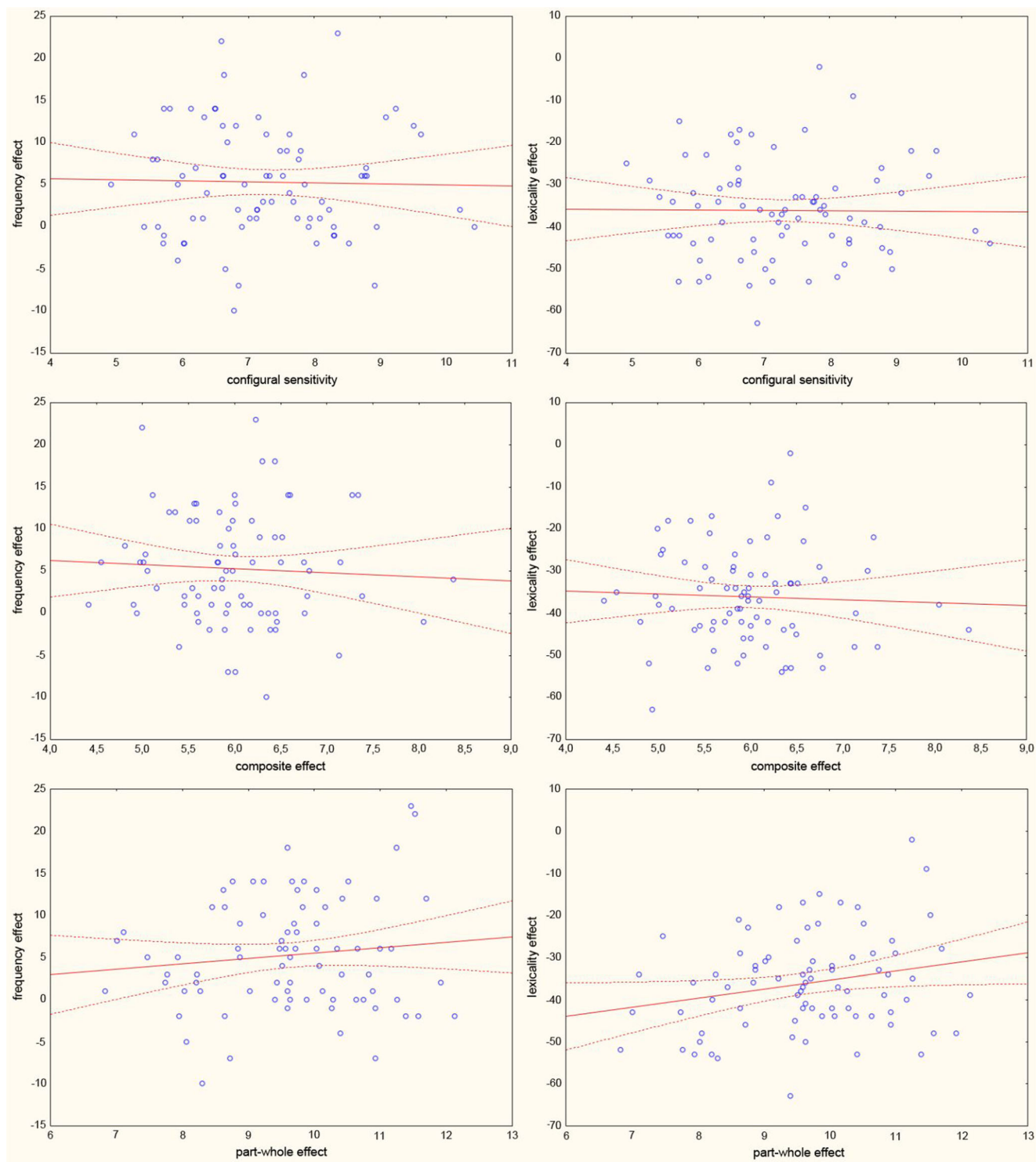
Fig. 4 Scatterplots illustrating the relationships between the bin scores of the three holistic processing measures and the 3dm scores: frequency and lexicality effects

stimulus. In contrast, the configural sensitivity task required matching of simultaneously shown stimuli that differed in the spatial configuration between different parts. Yet configural sensitivity was correlated significantly with the part–whole task ($r =$ .32), and its correlation with the composite effect ($r =$ .21) was

also marginally significant. It is therefore unlikely that the differences in task format provide a full account of the pattern of low correlations between the measures. Instead, the three measures likely tap on largely different processing characteristics defined under holistic processing.

The relationship between the configural sensitivity and part–whole effects may reflect the susceptibility of letter-level processing to the effects of high-level processes. Configural sensitivity involves the facilitation of letter processing by higher-level representations that are sensitive to the spatial relationships between letters and word parts. The part–whole task is commonly regarded as reflecting the top-down influence of the whole word orthographic representations on lower-level letter processing (McClelland & Rumelhart, 1981). Therefore, the correlation between configural sensitivity and part- whole task may reflect better communication between letter processing and higher-level visual and lexical representations.

The marginally significant correlation between configural sensitivity and the composite effect may lie in earlier, visual processing. For the composite effect, contributions by both visual and non-visual factors have been identified (Zhao et al., 2016a; Ventura et al., 2017). Focusing on the visual factors, Zhao et al. (2016b) reported composite effects for novel gestalt figures with which participants had no experience. Curby and Moerel (2019) further showed interference between the holistic processing of faces and that of the gestalt patterns. The interference was attributed to contributions of early perceptual mechanisms to the face composite effect. The configural sensitivity and composite effects may thus reflect an early visual overlap of holistic processing mechanisms.

The generally weak association between different measures of holistic word processing is akin to the result that was demonstrated previously for faces (Rezlescu et al., 2017). In both theirs and the current studies, the strongest correlation found was that between the inversion/configural sensitivity effect and the part–whole task. So together the two studies showed distinct mechanisms underlying holistic processing in the composite, configural sensitivity, and part–whole tasks. Despite this, there were differences in the specific patterns of correlations found in the two studies. In Rezlescu et al. (2017), a significant correlation was only found between the inversion and part- whole effects. However, there was no significant correlation between the inversion and the composite effect, or between the part–whole and the composite effects. Also, the level of expertise of the participants may not be comparable: Rezlescu et al.'s (2017) participants comprised of both college students and older adults participating over the Web, our participants were all students in a Portuguese university with expertise in Portuguese words more concentrated at the expert end. These differences will need to be controlled in future studies aiming at comparing the detailed correlation patterns across domains.

Holistic processing and word reading fluency

The other main goal of our study was to evaluate the potential links between different holistic processing measures and reading fluency. Ventura et al. (2020) showed that the composite effect correlated with lexical decision performance among fluent readers, although holistic processes in the word composite task are more important for low frequency words. Conway et al. (2017) found that the inversion effect with words was smaller for dyslexic adults than for control. And in Wong et al. (2019), there was a correlation between configural sensitivity and performance in a word fluency task among a group of Chinese-learning exchange students.

In the present study, we computed the correlations between bin scores for the three holistic tasks and two indexes reflecting access to the lexicon, frequency effect, and lexicality effect. We found a positive correlation between the part–whole task bin score and the lexicality effect ($r = .23$). The part–whole task is regarded as the result of the interaction between whole-word lexical representations (top-down influences) and low-level bottom-up processing at the letter level (e.g., McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). This correlation reflects, most probably, access to whole-word holistic lexical representations. Considering raw data, we found a negative correlation ($r = -.22$) between part–whole RT and frequency effect in 3DM and a negative correlation ($r = -.22$) between composite effect RT and frequency effect in 3DM. Considering raw scores, we found a tendency for a correlation between the part–whole task in RT and Component 2 of PCA: $r = .20$, $p = .076$, and a significant correlation of composite aligned accuracy and Component 2, reflecting lexical access, of PCA, but the importance of this correlation should be taken with caution since it was not found for the bin scores.

The composite effect reflects 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, which can involve inflexible attentional weighting to all parts of the object (Chua et al., 2014; Richler et al., 2012; Richler & Gauthier, 2014; Richler et al., 2011b). Recently we showed (Ventura et al., 2020) that holistic processing of visual words is associated with efficient access to the orthographic lexical representations, among adult fluent readers. In the present study, we found no significant correlation between composite bin scores and either frequency or lexicality effects in 3DM. Considering raw data, we found a significant negative correlation of composite aligned RT and lexicality effect and a significant correlation of composite aligned accuracy and Component 2, reflecting lexical access, of PCA. The importance of these correlations should be taken with some caution since they were not found for the bin scores.

Considering now configural sensitivity and its lack of correlation with 3DM, note that Wong et al. (2019; Experiment 2) did not find any significant association between sensitivity to configural information and word recognition by Chinese native readers, that is, skillful readers and experts on Chinese

words. Thus, an interesting possibility, based on the available research (e.g., Conway et al., 2017; Ventura et al., 2020; Wong et al., 2019), is that configural sensitivity, as assessed by the word- inversion effect, better predicts differences in reading performance between beginning (or less fluent) readers—that is, nonexperts.

The little correlation found between holistic processing and word reading measures may suggest that holistic processing simply develops as a by-product of fluent word reading, instead of being an important skill to acquire for optimizing reading fluency. This claim should be taken with caution. First, the current study was conducted on typical readers only. An alternative explanation is that holistic processing is required to achieve a certain level of reading fluency, as for novice readers learning a language. Once an individual's reading ability passes a particular level, holistic processing would not entail any further benefit. Second, while multiple paradigms were used to measure holistic processing, the word reading fluency and lexical access measure were obtained from the same 3DM test. It remains to be seen whether the pattern of relationship between holistic processing and reading fluency (part–whole task bin score correlating with lexicality effect; negative correlation between both part–whole RT and composite effect RT and frequency effect in 3DM) generalizes to a variety of word reading paradigms involving different task demands and contexts.

An alternative (to the idea that holistic processing simply develops as a by-product of fluent word reading and/or the idea that holistic processing is required to achieve a certain level of reading fluency) is to consider that the differences in the correlations between the different HP measures and reading result from the differences in the mechanisms they are tapping and those involved in reading at different stages of reading acquisition. As discussed above, the configural sensitivity effect may reflect an early visual processing mechanism that might be more important for beginning readers. Indeed, in Wong et al. (2019), there was a correlation between configural sensitivity and performance in a word fluency task among a group of Chinese-learning exchange students. The impairment of this sensitivity to configuration might be related to difficulties or very low performance in reading for beginners in reading. The composite effect reflects 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, which can involve inflexible attentional weighting to all parts of the object (Chua et al., 2014; Richler et al., 2012; Richler & Gauthier, 2014; Richler et al., 2011b). The anterior portion of the VWFA contains neurons tightly tuned to whole-word orthographic representations (e.g., Thesen et al., 2012; Vinckier et al., 2007), which differentiate whole words (Strother et al., 2017), This evidence suggests that this region is the neural underpinning of holistic, lexical (whole-word) orthographic representations (Bouhali

et al., 2019; Lerma-Usabiaga et al., 2018; White et al., 2019). It also agrees with evidence showing that neurons of the vOT become selective for whole representations of items of expertise (after training discrimination at the individual level on these multi-component items, e.g., Baker et al., 2002). This would imply a progressive role of holistic processing in the composite task as children progress from sublexical to lexical routes in reading. Finally, the part–whole task is commonly regarded as reflecting the top-down influence of the whole word orthographic representations on lower-level letter processing. Most probably this task would be correlated with efficient use of the lexical route in reading.

These aspects could be tested either by (i) running studies with children at different levels of reading; or (ii) running studies of dyslexic children and adults.

As discussed above, our results show clear similarities between word processing and face processing (Rezlescu et al., 2017). Considering the correlation between holistic measures and measures of lexical access there are also similarities with the findings of Rezlescu et al. (2017). Indeed, in the present study, we found only a significant correlation for part–whole bin scores effect. Rezlescu et al. also found only a significant correlation, but in their case the inversion effect. Face and visual word recognition involve different neural mechanisms with an opposite hemispheric lateralization (VWFA; e.g., Dehaene & Cohen, 2011; and fusiform face area, e.g., Kanwisher et al., 1997). It is then possible that faces and words share similar holistic processes in their own separate face and word processing systems.

Conclusion

In sum, using bin scores that combine response speed and accuracy and thus increase reliability, we found some overlap between the three common holistic word processing measures. Overall, reliabilities were medium to high (and comparable with those in the different holistic processing tasks used with faces in Rezlescu et al., 2017) and thus the low/null correlations were not a result of noisy data. Configural sensitivity and part–whole effect were the most correlated with each other, the correlation between the composite and the configural sensitivity task was marginally significant while the composite and part–whole effects were not related. The similarities may correspond to processing at earlier visual levels, while the differences may concern more with higher-level processing. Of the three holistic processing bin score measures, only one (part–whole effect) correlated with a lexical access measure of 3DM. This suggests either that holistic processing is a by-product instead of a contributor of fluent word reading, or that a certain level of holistic processing is sufficient for efficient word reading, with little further benefits for readers with a certain level of skills and beyond.

However, and as discussed above, an alternative is to consider that the differences in the correlations between the different HP measures and reading result from the differences in the mechanisms they are tapping and those involved in reading at different stages of reading acquisition. The pattern of relationships found between holistic processing and word reading bared some similarities with those found for faces, suggesting some overlap in the principles underlying holistic face and word processing.

Appendix 1

The bisyllabic four-letter Portuguese words used to create the stimuli for the different experimental conditions in the composite task

Set			Words	
1	BIFE (steak)	BICO (beak)	SAFE (get way)	SACO (bag)
2	BODE (goat)	BOGA (bogue)	RUDE (rude)	RUGA (wrinkle)
3	BULE (teapot)	BUDA (buddha)	ROLE (roll)	RODA (wheel)
4	DOSE (dose)	DOCA (dock)	BISE (encore)	BICA (spout)
5	DURE (last)	DUNA (dune)	MIRE (aim)	MINA (mine)
6	FIGO (fig)	FITA (ribbon)	NEGO (I deny)	NETA (granddaughter)
7	LISA (smooth)	LIGO (care)	PESA (weighs)	PEGO (I catch)
8	MEGA (mega)	MERO (mere)	FIGA (fig)	FIRO (I hurt)
9	PUDE (I could)	PUXA (pull)	LIDE (deal)	LIXA (sandpaper)
10	RIJA (tough)	RIME (rime)	FUME (smoke)	FUJA (run away)
11	SINA (lot)	SIGO (following)	RENA (reindeer)	REGO (gully)
12	VICE (vice)	VIGA (beam)	ROCE (rook)	ROGA (entreats)
13	BASE (base)	BAFO (breath)	PISE (tread)	PIFO (drunkenness)
14	CUME (top)	CUJA (whose)	SOME (add)	SOJA (soy)
15	DIGA (say)	DITO (said)	NEGA (denies)	NETO (grandson)
16	FASE (phase)	FARO (flair)	VISE (aim)	VIRO (turn)
17	FOLE (bellows)	FORA (outside)	PULE (jump)	PURA (pure)
18	LEVA (takes)	LEGO (Lego)	DIVA (diva)	DIGO (I say)
19	LIMA (lime)	LISO (smooth)	TEMA (theme)	TESO (stiff)
20	LUTE (fight)	LUVA (glove)	NOTE (notice)	NOVA (new)
21	REGA (irrigation)	REMO (rowing)	LIGA (alloy)	LIMO (slime)
22	TINA (tub)	TIPO (type)	CENA (scene)	CEPO (block)
23	TIVE (had)	TIRA (strip)	NOVE (nine)	NORA (daughter-in-law)
24	VIME (rattan)	VILA (town)	COME (eat)	COLA (glue)

Appendix 2

The words used in the configural sensitivity task

acabar (end)	facto (fact)	local (local)	pronto (ready)
acordo (agreements)	filme (film)	lugar (place)	pouco (little)
antes (before)	final (end)	manter (keep)	prova (proof)
cerca (fence)	forma (shape)	menos (less)	quanto (how much)
cinco (five)	frente (front)	minha (mine)	quase (almost)
coisa (thing)	grupo (group)	mundo (world)	sector (sector)
dentro (inside)	guerra (war)	nosso (ours)	semana (week)
desde (since)	homem (man)	número (number)	único (single)
equipa (team)	levar (take)	0utra (another)	várias (various)
estado (state)	líder (leader)	papel (paper)	viver (to live)

Appendix 3

The words and pseudowords used in the word superiority task

3-letter items
Words

nos (at the)	dos (of)	ser (to be)
mim (me)	era (it was)	ter (have)
vem (comes)	nem (nor)	vez (turn)
rei (king)	diz (say)	boa (good)
vou (I will)	fez (did)	paz (peace)
dor (pain)	sul (south)	vir (come over)
tom (tone)	dom (endowment)	mar (sea)
lar (home)	iva (vat)	fio (thread)
tua (your)	ama (love)	rir (laugh)

Nonwords

nbu	ron	lxr
mji	qrv	ozr
vni	aej	aoz
rqt	mid	qfa
vmp	cex	zqz
dzu	tuv	lfr
tzm	qos	hlr
liq	hve	nuo
tvh	hmb	tlr

4-letter items
Words

duas (two)	tudo (all)	cada (each)	pelo (fur)
deve (should)	bons (good)	fase (phase)	nova (new)
rede (network)	alma (soul)	nove (nine)	essa (that)
fala (speech)	lixo (trash)	meia (sock)	cujo (whose)
medo (fear)	erro (error)	dada (given)	cair (fall)
ouro (gold)	caro (expensive)	visa (visa)	dura (hard)
anda (walks)	mata (woods)	pele (skin)	vejo (watch)
vaga (vacancy)	ramo (branch)	maia (Mayan)	ouve (hears)
cria (create)	muro (wall)	caos (chaos)	bela (beautiful)
neve (snow)	seco (dry)	rota (route)	olho (eye)

Nonwords

dtnm	puxh	gjdo	mzuo
dvhq	poxn	pvsn	xhba
rorx	aljr	gzvb	ahoa
fmth	zivc	jnil	htmo
mgou	drfm	iedp	ixsr
ocmz	xamt	dusb	ndja
aont	jabf	culd	sgho
vmfr	hatv	cdis	zsxe
crif	ruaj	jtoz	dnta
nvqh	uexm	sotv	nsco

5-letter items
Words

certo (right)	tanta (so much)	venha (come)	salvo (saved)
gosto (taste)	creio (I believe)	ponta (tip)	falha (failure)
morto (dead)	pasta (folder)	fogos (fires)	longo (long)
costa (coast)	santo (holy)	vacas (cows)	posse (possession)
banho (shower)	antas (tapirs)	desde (since)	pediu (asked)
penas (feathers)	conto (tale)	novas (new)	vinho (wine)
caras (faces)	forte (strong)	estas (these)	plena (full)
natal (Christmas)	julho (July)	carro (car)	firma (firm)
carta (letter)	greve (strike)	esses (those)	justa (fair)
posso (can)	custo (cost)	preto (black)	fraca (weak)

Nonwords

ctzpb	jadxh	ihndp	jnuve
gdxav	grzeu	mhnfg	nuzht
mbgdj	iapxu	tsgan	vnjgo
cpdvz	oatfu	hlcvn	mchre
bogqr	dngjf	cegdp	toxdu
phojd	mourh	pitab	fntjo
cpsah	bsrxc	deqap	stixa
nvdob	izlpx	zqars	hrjda
pafjg	mteac	oqxez	arjsa
mortd	xgsoa	hcntz	sidxa

Appendix 4

The high-frequency words, low-frequency words and the pseudowords used in the reading fluency (3DM) task

High-frequency words

lata (tin)	ferro (iron)	circo (circus)	espelho (mirror)	escrever (write)
foca (seal)	mocho (owl)	barco (boat)	trabalho (work)	estrelas (stars) pele (skin)
pele (skin)	banho (shower)	fruta (fruit)	carnival (carnival)	depressa (quickly) bico (spout)
bico (spout)	bicho (bug)	grupo (group)	devagar (slowly)	narrador (storyteller)
fato (fact)	burro (dumb)	jardim (garden)	cigarra (cigarette)	problema (problem)
dono (owner)	milho (corn)	pasta (folder)	hospital (hospital)	lavrador (farmer ramo (branch)
ramo (branch)	sonho (dream)	clara (clear)	segredo (secret)	conversa (talk)
rogo (fire)	passa (goes by)	pedir (ask)	conhecer (to know)	procurer (find)
sono (sleep)	carro (car)	trigo (wheat)	esparto (smart)	floresta (floresta)
bolo (cake)	fundo (bottom)	jornal (newspaper)	vermelho (red)	personagem (character)
mata (woods)	palha (straw)	chover (to rain)	mensagem (message)	espantatalho (scarecrow)
belo (beautiful)	monte (lot)	pastor (shepherd)	estrada (road)	importante (important)
saco (bag)	linha (line)	jantar (diner)	presente (present)	professorar (teacher)
vila (village)	massa (pasta)	pardal (sparrow)	pergunta (question)	borboletas (butterflies)
fome (hunger)	ninho (nest)	grilo (cricket)	surpresa (surprise)	diferentes (different)

Low-frequency words

lota (fish market)	forro (lining)	cerco (siege)	espelho (mirror)	escravos (slaves)
foco (focus)	macho (male)	barca (ferry)	presilha (clamp)	espremer (squeeze)
pala (pala)	banha (lard)	frota (fleet)	cardinal (cardinal)	caruncho (woodworm)
beco (alley)	bucha (bushing)	gripe (flu)	divagar (ramble)	massagem (massage)
feto (fetus)	birra (tantrum)	marfim (ivory)	cigarro (cigarette)	grossura (thickness)
duna (dune)	milha (mile)	pasto (pasture)	marginal (marginal)	pastilha (tablet)
rama (rama)	senha (password)	cloro (chorine)	sagrado (sacred)	concurso (contest)
fuga (escape)	fossa (fossa)	podar (prune)	sonhador (dreamer)	contrato (contract)
sina (sina)	coche (coach)	prego (nail)	esperta (smart)	frisados (beaded)
bala (bullet)	fenda (slit)	farnel (farnel)	sardinha (sardine)	consumidor (consumer)
mito (myth)	malha (mesh)	chocar (shock)	consolar (consoling)	desfolhada (defoliated)
bule (teapot)	manta (blanket)	pastar (graze)	estrado (platform)	cintilante (sprakling)
soco (punch)	linho	conter (to contain)	presunto (ham)	comprimido (tablet)
vala (ditch)	posse	portal (portal)	surfista (surfer)	convocados (summoned)
fama (fame)	pinho	greve (strike)	discreto (discreet)	disfarçado (disguised)

Nonwords

lano	felha	cirta	espretal	espresa
fomo	rinho	barlo	tragunda	derralas
pefa	bacho	frugo	carsagar	escrema
bito	binho	gruco	depeval	matredor
fata	bussa	jarnal	cinalho	provertra
dole	ticho	pasco	hosmeta	concurar
raca	sorro	clata	segrelho	lablever
folo	palmo	petor	copergem	flovrassa
sogo	canha	tripo	esgate	proresdor
boco	funte	jordir	versento	perfetates
maco	panho	chodim	mentrasa	esbotante
beme	monfa	pasver	espicer	implanegem
salo	lirro	jandal	prebarra	posossollho
vita	malco	partar	pervenho	borerentas
fono	nissa	grirar	survalho	dipornara

Appendix 5

Correlation analyses with raw accuracy and corrected RT measures

Reliability of the measures

The Guttman's λ^2 was .395 in RT and .396 in accuracy for the composite effect, .85 in RT and .45 in accuracy for the configural sensitivity effect, and .66 in RT and .42 in accuracy for the part–whole effect. The Guttman's λ^2 for the 3DM task as a whole was .85.

Relationships between holistic processing and reading fluency measures

Table 3 show Pearson-product correlation coefficients between the different holistic processing and reading fluency measures and the scatterplots, respectively

The three holistic processing effects were mildly correlated with each other's similar to, though less significant than the results obtained from the bin scores. There was a positive correlation between the word composite effect in RT and configural sensitivity effect in RT ($r = .25$, $p = .027$). The part–whole effect in RT was positively correlated with the configural sensitivity effect in RT ($r = .23$, $p = .04$), and negatively correlated with the configural sensitivity effect in accuracy ($r = −.25$, $p = .053$). The correlation between the composite and the part–whole effect was not significant (rs = .11 and −.01, respectively).

Similar to what was found with the bin scores, there was not much relationship between the holistic processing measures and reading fluency (cf. Fig. 5). All correlations were low (rs < .19), except for negative correlations between the part–whole effect in RT and the frequency effect in 3DM task performance ($r = −.22$, $p = .054$). and between the composite effect in RT and the frequency effect in 3DM task performance ($r = −.22$, $p = .048$).

Table 3 Pearson-product correlations between the RT and accuracy of the three HP measures and the 3DM scores

		Composite effect (RT)	Composite effect (ACC)	Configural sensitivity (RT)	Configural sensitivity (ACC)	Part– whole task (RT)	Part– whole task (ACC)
Composite effect (RT)	r						
	p						
Composite effect (ACC)	r	.21*					
	p	.05					
Configural sensitivity (RT)	r	.25*	−0.14				
	p	.027	0.18				
Configural sensitivity(ACC)	r	.008	.13	-0.21			
	p	.95	0.18	0.053			
Part-whole task (RT)	r	.11	0.076	.23*	−.25*		
	p	.33	0.47	0.04	0.027		
Part-whole task (ACC)	r	−.01	0.012	−0.06	−0.061	0.039	
	p	.96	0.9	0.60	0.59	0.73	
Frequency effect (3 DM)	r	−.22*	0.01	−0.16	−0.02	−.22*	0.07
	p	.048	0.92	0.15	0.81	0.054	0.55
Lexicality Effect(3DM) .48	r	−.19	.03	−.11	−.11	−.05	.08
	p	.08	.77	.34	.34	.66	.48

** Correlation is significant at the 0.01 level (two-tailed).

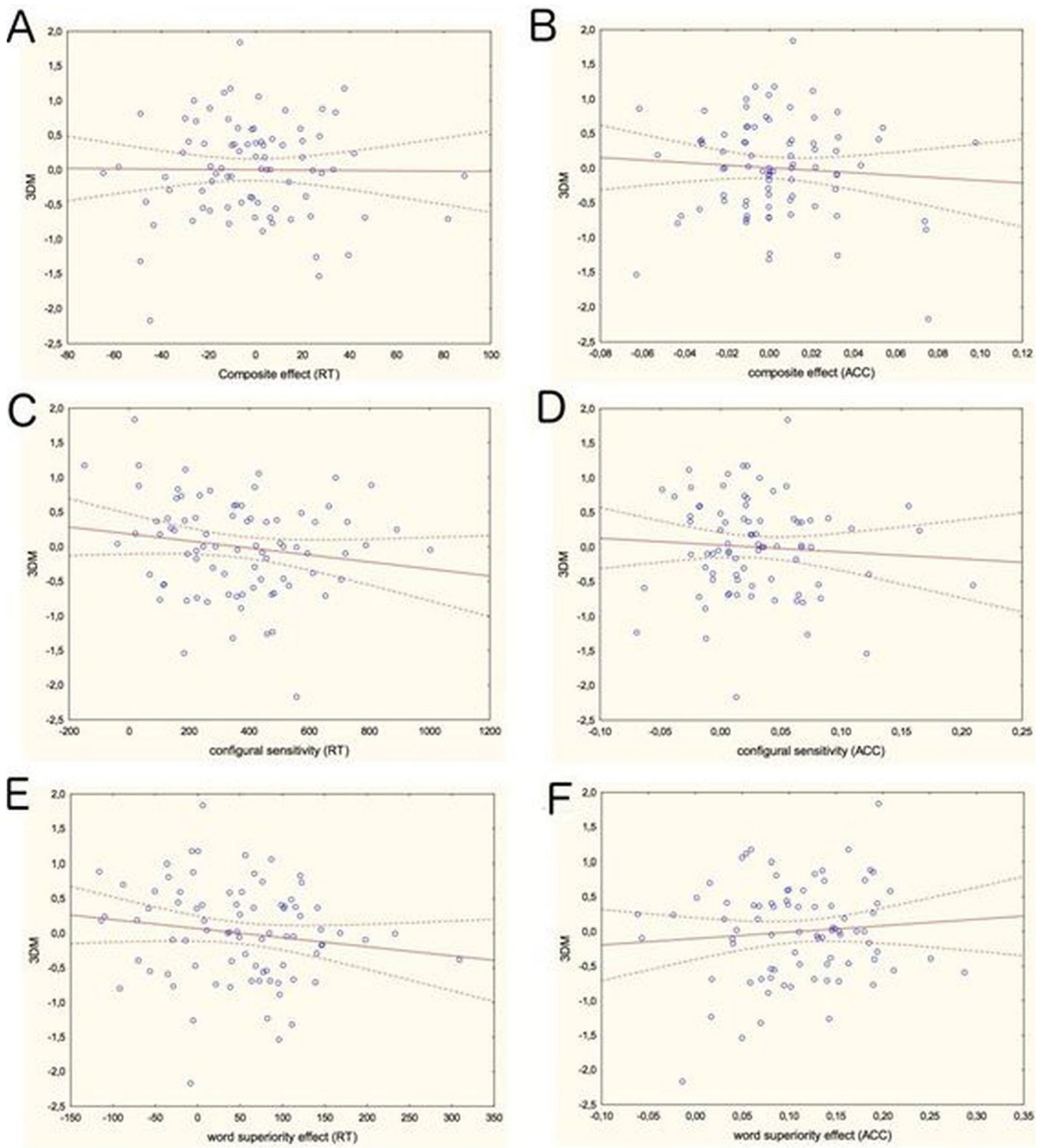* Correlation is significant at the 0.05 level (two-tailed).

Fig. 5

Declarations

Conflicts of interest/Competing interests The authors have no relevant financial or nonfinancial interests to disclose

Ethics approval This study followed the Declaration of Helsinki and Portuguese Deontological Regulation. Also, it was approved by the Deontological Committee of Faculdade de Psicologia of Universidade de Lisboa.

Consent to participate Freely given, informed consent to participate in the study was obtained from participants.

Consent for publication All participants were informed that their data were anonymous.

References

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

Björnström, L. E., Hills, C., Hanif, H., & Barton, J. J. (2014). Visual word expertise: A study of inversion and the word-length effect, with perceptual transforms. *Perception, 43*(5), 438–450.

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,* Article 201904184. https://doi.org/10.1073/pnas.1904184116

Cai, Y., Mao, Y., Ku, Y., & Chen, J. (2020). Holistic integration in the processing of Chinese characters as revealed by electroencephalography frequency tagging. *Perception, 49*(6), 658–671. https://doi.org/10.1177/0301006620929197

Chen, C., Abbasi, N. U. H., Song, S., Chen, J., & Li, H. (2016). The limited impact of exposure duration on holistic word processing. *Frontiers in Psychology, 7*, 646. https://doi.org/10.3389/fpsyg.2016.00646

Chen, H., Bukach, C. M., & Wong, A. C. N. (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.

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

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*, 1–8.

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

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.

Draheim, C., Hicks, K. L., & Engle, R. W. (2016). Combining reaction time and accuracy: The relationship between working memory capacity and task switching as a case example. *Perspectives on Psychological Science, 11*(1), 133–155.

Duchaine, B., Germine, L., & Nakayama, K. (2007). Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. *Cognitive Neuropsychology, 24*(4), 419–430.

Farah, M. J. (1991). Cognitive neuropsychology: Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology, 8*(1), 1–19.

Farah, M. J. (1992). Agnosia. *Current Opinion in Neurobiology, 2*, 162–164.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*(3), 482–498.

Fernandes, T., Araújo, S., Sucena, A., Reis, A., & Castro, S. L. (2017). The 1-min screening test for reading problems in college students: Psychometric properties of the 1-min TIL. *Dyslexia, 23*(1), 66–87.

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., Wong, A. C., Hayward, W. G., & Cheung, O. S. (2006). Font tuning associated with expertise in letter perception. *Perception, 35*(4), 541–559. https://doi.org/10.1068/p5313

Gauthier, I., Wong, A. C., & Palmeri, T. J. (2010). Manipulating visual experience: comment on Op de Beeck and Baker. *Trends in Cognitive Sciences, 14*(6), 235–237. https://doi.org/10.1016/j.tics.2010.03.009

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior Research Methods, 46*(3), 702–721.

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 : the official journal of the Society for Neuroscience, 17(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

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.

Kao, C. H., Chen, D. Y., & Chen, C. C. (2010). The inversion effect in visual word form processing. *Cortex, 46*(2), 217–230.

Koriat, A., & Norman, J. (1985). Reading rotated words. *Journal of Experimental Psychology: Human Perception and Performance, 11*(4), 490–508.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 51*(3), 449–473. https://doi.org/10.1080/713755777

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

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature, 410*(6831), 890.

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

Lima, C. F., & Castro, S. L. (2010). Reading strategies in orthographies of intermediate depth areflexible: Modulation of length effects in Portuguese. *European Journal of Cognitive Psychology, 22*(2), 190–215. https://doi.org/10.1080/09541440902750145

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*(5), 375–407.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*(5), 553–566. https://doi.org/10.1068/p3339

Murphy, J., Gray, K. L., & Cook, R. (2017). The composite face illusion. *Psychonomic Bulletin & Review, 24*(2), 245–261. https://doi.org/10.3758/s13423-016-1131-5

Paap, K. R., Newsome, S. L., & Noel, R. W. (1984). Word shape's in poor shape for the race to the lexical. *Journal of Experimental Psychology: Human Perception and Performance, 10*(3), 413–428.

Pacheco, A., Reis, A., Araújo, S., Inácio, F., Petersson, K. M., & Faísca, L. (2014). Dyslexia heteroge-neity: Cognitive profiling of Portuguese children with dyslexia. *Reading and Writing, 27*, 1529–1545. https://doi.org/10.1007/s11145-014-9504-5

Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature, 423*(6941), 752–756. https://doi.org/10.1038/nature01516

Rakover, S. S. (2013). Explaining the face-inversion effect: The face–scheme incompatibility (FSI) model. *Psychonomic Bulletin & Review, 20*(4), 665–692.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81*(2), 275–280.

Rezlescu, C., Susilo, T., Wilmer, J. B., & Caramazza, A. (2017). The inversion, part–whole, and composite effects reflect distinct perceptual mechanisms with varied relationships to face recognition. *Journal of Experimental Psychology: Human Perception and Performance, 43*(12), 1961–1973.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces?. *Cognition, 47*(1), 25–57. https://doi.org/10.1016/0010-0277(93)90061-y

Richler, J. J., Floyd, R. J., & Gauthier, I. (2014). The Vanderbilt Holistic Face Processing Test: A short and reliable measure of holistic face processing. *Journal of Vision, 14*(11), 1–14.

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

Richler, J. J., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011a). Inverted faces are (eventually) processed holistically. *Vision Research, 51*(3), 333–342.

Richler, J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology, 3*, 553.

Richler, J. J., Wong, Y. K., & Gauthier, I. (2011b). 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

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

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89*(1), 60–94.

Sanocki, T. (1988). Font regularity constraints on the process of letter recognition. Journal of experimental psychology. *Human Perception and Performance, 14*(3), 472–480. https://doi.org/10.1037//0096-1523.14.3.472

Schneider, W., Eschman, A., and Zuccolotto, A. (2012a). *E-Prime user's guide.* Psychology Software Tools, Inc.

Schneider, W., Eschman, A., and Zuccolotto, A. (2012b). *E-Prime reference guide.* Psychology Software Tools, Inc.

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. Journal of experimental psychology. Human Perception and Performance, 22(4), 904–915. https://doi.org/10.1037//0096-1523.22.4.904

Simpson, I. C., Mousikou, P., Montoya, J. M., & Defior, S. (2013). A letter visual-similarity matrix for Latin-based alphabets. *Behavior Research Methods, 45*, 431–439. https://doi.org/10.3758/s13428-012-0271-4

Soares, A. P., Iriarte, Á., de Almeida, J. J., Simões, A., Costa, A., Machado, J., & Perea, M. (2018). Procura-PALavras (P-PAL): A web-based interface for a European Portuguese lexical database. *Behavior Research Methods, 50*(4), 1461–1481.

Strother, L., Zhou, Z., Coros, A. K., & Vilis, T. (2017). An fMRI study of visual hemifield integration and cerebral lateralization. *Neuropsychologia, 100*, 35–43. https://doi.org/10.1016/j.neuropsychologia.2017.04.003

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology Section A, 46*(2), 225–245.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25*, 583–592. https://doi.org/10.3758/Bf03211301

Tanaka, J. W., & Simonyi, D. (2016). The "parts and wholes" of face recognition: A review of the literature. *Quarterly Journal of Experimental Psychology, 69*(10), 1876–1889. https://doi.org/10.1080/17470218.2016.1146780

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

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology, 67*(6), 1176–1190.

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., Almeida, V. B., Casqueiro, I., & Wong, A. C. N. (2017). The word composite effect depends on abstract lexical representations but not surface features like case and font. *Frontiers in Psychology, 8*, 1036. https://doi.org/10.3389/fpsyg.2017.01036

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

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

Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Bargary, G., & Mollon, J. D. (2017). General and specific factors in the processing of faces. *Vision Research, 141*, 217–227.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1*(1), 59–85.

White, A. L., Palmer, J., Boynton, G. M., & Yeatman, J. D. (2019). Parallel spatial channels converge at a bottleneck in anterior word-selective cortex. *Proceedings of the National Academy of Sciences, 116*(20), 10087–10096. https://doi.org/10.1073/pnas.1822137116

Wong, A. C. N., Bukach, C. M., Hsiao, J., Greenspon, E., Ahern, E., Duan, Y., & Lui, K. F. (2012). Holistic processing as a hallmark of perceptual expertise for nonface categories including Chinese characters. *Journal of Vision, 12*(7), 1–15. https://doi.org/10.1167/12.7.1

Wong, A. C. N., Bukach, C. M., Yuen, C., Yang, L., Leung, S., & Greenspon, E. (2011). Holistic processing of words modulated by reading experience. *PLOS ONE, 6*(6), e20753.

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

Wong, A. C., 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., Wong, Y. K., Lui, K. F., Ng, T. Y., & Ngan, V. S. (2019). Sensitivity to configural information and expertise in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 45*(1), 82–99.

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English Lexical Project. *Journal of Experimental Psychology: Human Perception and Performance, 38*(1), 53–79.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141–145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*(6), 747–759.

Zhao, M., Bülthoff, H. H., & Bülthoff, I. (2016a). Beyond faces and expertise: Facelike holistic processing of nonface objects in the absence of expertise. *Psychological Science, 27*(2), 213–222. https://doi.org/10.1177/0956797615617779

Zhao, M., Bülthoff, H. H., & Bülthoff, I. (2016b). A shape-based account for holistic face processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(4), 584.