



# The relation between holistic processing as measured by three composite tasks and face processing: A latent variable modeling approach

Paulo Ventura<sup>1</sup> · Vince Ngan<sup>2</sup> · Alexandre Pereira<sup>3</sup> · Francisco Cruz<sup>1</sup> · José C. Guerreiro<sup>1</sup> · Vivienne Rosário<sup>1</sup> · João Delgado<sup>1</sup> · Bruno Faustino<sup>1</sup> · Marta Barros<sup>1</sup> · Miguel Domingues<sup>1</sup> · Alan Wong<sup>2</sup>

Accepted: 15 July 2022 / Published online: 1 August 2022  
© The Psychonomic Society, Inc. 2022

## Abstract

We investigated the relationship between holistic processing and face processing using a latent variables approach. Three versions of the composite paradigm were used to measure holistic processing: Vanderbilt Holistic Face Processing Test, a sequential composite matching task, and a simultaneous composite matching task. Three tasks were used to measure face perception and face memory abilities respectively. We had three pairs of tasks such that within each pair (of memory and perception task), the stimuli involved, the requirement for matching across viewpoints, etc., are the same, such that the only difference is whether perception or memory is taxed. There were no significant correlations between the different versions of the composite task. We discovered no evidence to support a distinction between face perception and face memory, suggesting the existence of a general face processing factor. Finally, there was no evidence that holistic processing (as captured by either of the three composite tasks) is predictive of better face processing per se, casting doubts on the role of holistic processing in differentiating different levels of efficiency in face processing.

**Keywords** Holistic processing · Face processing · Face perception · Face memory · Holistic processing and face processing

## Introduction

Holistic processing, the tendency to process features in an integrated fashion and not as isolated features, is considered crucial to differentiate visually similar objects like faces (Diamond & Carey, 1986; Gauthier & Tarr, 2002; McKone et al., 2012; Richler & Gauthier, 2014). In the original composite study (A. W. Young et al., 1987) participants were presented with a single facial composite, and the degree of illusion-induced interference was inferred from their response latencies (Murphy, Gray, & Cook, 2017). When asked to name the identity of a face half (e.g., upper half), whilst

disregarding the remaining task-irrelevant half (e.g., lower half), observers were disproportionately slower in the upright-aligned condition. The application of this naming paradigm is limited, however, by the need to use familiar faces. The composite task (for a discussion, see Fitousi, 2015; Richler & Gauthier, 2014; Rossion, 2013), in which participants perform a same-different task involving decisions about two subsequently or concurrently presented stimuli based on one half of each stimulus while ignoring the other half, is a widely used paradigm to study holistic processing involving unfamiliar faces. The composite effect describes how an irrelevant half of a face (e.g., bottom half) influences the processing of the task-relevant half of the face (e.g., the top half). Any interference in performance from the irrelevant on the relevant part implies automatic and obligatory holistic processing of all parts of the stimuli in the composite task.

✉ Paulo Ventura  
paulo.ventura@gmail.com

<sup>1</sup> Faculdade de Psicologia, Universidade de Lisboa, Lisbon, Portugal

<sup>2</sup> Department of Psychology, The Chinese University of Hong Kong, Shatin, NT, Hong Kong

<sup>3</sup> CICANT, Centre for Research in Applied Communication, Culture, and New Technologies, Universidade Lusófona de Humanidades e Tecnologias, Lisbon, Portugal

## Association between holistic processing and face processing

Surprisingly, the proposed role of holistic processing in face processing was not tested until 2010, with mixed results. About 10 years ago this association was tested using the

composite task and face recognition. Konar et al. (2010) and Verhallen et al. (2017) found no association between holistic processing and face-recognition abilities measured by the Cambridge Face Memory Test (CFMT). Richler, Cheung, and Gauthier (2011a) on the other hand, discovered a small but significant correlation between holistic processing and CFMT performance.

Two different composite task designs were used in these studies: (i) The perceptual integration design. In this design, the irrelevant half of the face is always different, while the target half may be the same or different. Where only the same-response trials are considered, holistic processing is operationalized as an alignment effect: worse performance when the two halves of the face are aligned than when misaligned through a lateral shift (e.g., Goffaux & Rossion, 2006; Hole, 1994; Young et al., 1987). (ii) The selective attention design: this design orthogonally manipulates alignment and congruency (Gauthier & Bukach, 2007). Congruent and incongruent trials are included in both same- and different-response trials and holistic processing is signaled by a significant interaction between alignment and congruency across response-type (DeGutis et al., 2013; Richler, Cheung, & Gauthier, 2011a; for a meta-analysis, see Richler & Gauthier, 2014). The perceptual integration design was used by Konar et al. (2010) and Verhallen et al. (2017), while the selective attention design was used by Richler, Cheung, and Gauthier (2011a). These differences may arise because the perceptual integration design confounds manipulations of interest with complex response biases, resulting in poor construct validity (Richler & Gauthier, 2014).

However, the distinctions between the two designs of the composite task are not the end of the story. There was no evidence for such a relationship with CFMT using a different and more reliable selective attention measure of holistic processing in a composite task (Vanderbilt Holistic Face Processing Test (VHFPT-F); Richler et al., 2014). Recently, Boutet and Meinhardt-Injac (2021) found performance on the selective attention composite task to be a significant predictor of performance on a face-matching task (GFMT: Glasgow Face Memory Test, Burton et al., 2010), but not the face recognition task, CFMT. A factor that might contribute to these discrepancies is whether the composite task has, or has not, repetition of face parts and the face-processing test has, or has not, repetition of faces. A learning component in both tasks may give rise spuriously to a significant correlation. VHFPT-F has no learning component that might explain the lack of correlation with CFMT, a test with a strong learning component.

We have discussed the composite task and its relationship with face processing (recognition). But other tasks have been proposed to measure holistic processing: inversion (a large cost in performance for upside-down faces) and part-whole effects (poorer recognition of parts in isolation than in whole faces)

Rezlescu et al. (2017) examined the relationships between the alignment effect in the perceptual integration composite task, inversion, and part-whole effects, and whether these effects are predictive of face processing, more specifically face perception evaluated by the Cambridge Face Perception Test (CFPT; Duchaine & Nakayama, 2006). Intriguingly, the three holistic processing measures did not show a high overlap, with a significant and moderate correlation found only between inversion and part-whole measures ( $r = .28$ ). These measures also differed greatly in terms of their relationships with face perception ability, with correlations ranging from moderate ( $r = .42$  for the inversion effect) to weak ( $r = .25$  for the part-whole effect) and non-existent ( $r = .04$  for the composite effect).

The lack of correlation between the different measures of holistic processing implies that they probably reflect different mechanisms. Li et al. (2017) looked at whether the part-whole effect and composite face effect involved shared or distinct neural substrates and found evidence for distinct mechanisms. They found that the part-whole effect and composite face effect showed hemispheric dissociation in the fusiform face area (FFA), with the part-whole effect correlating with face selectivity in the left FFA and the composite face effect correlating with face selectivity in the right FFA. Furthermore, the association between the part-whole effect and face selectivity was driven primarily by the FFA response to faces, whereas the association between the composite face effect and face selectivity was caused by a suppressed response to objects in the right FFA.

The finding of Li et al. (2017) that the composite face effect is related to object processing is also consistent with Fitousi's object-based attention account of the composite face effect. Fitousi (2015) surmised that the composite face effect is not unique to faces but reflects a general attentional strategy akin to object-based attention. In object-based attention, all constituent features or parts are activated once an object has been selected. Fitousi (2015) proposes that the reduction in the congruence composite effect with misalignment is due to a disruption of objecthood. Analysis of theoretical ex-Gaussian parameters of response time (RT) distributions (Fitousi, 2020) reveals that the composite effect is generated by pure changes in the exponential component of the ex-Gaussian distribution, suggesting the involvement of attentional and working memory processes in the composite face effect.

The main goal of our study is to evaluate the relationship between holistic processing and face processing. Given the evidence reviewed above that different holistic tasks most probably reflect different mechanisms, we will next explain our approach to choosing holistic processing measures. We will then present the reasoning that led to the choice of the face-processing tasks.

## Measuring holistic processing

Rezlescu et al. (2017) observed a significant and moderate correlation only between inversion and part-whole effects, while Li et al. (2017) found no association in the neural substrates of the part-whole effect and composite face effect. The three tasks most probably reflect distinct mechanisms. Indeed, the inversion effect is an indirect measure of holistic processing (Piepers & Robbins, 2012). Holistic processing in the composite task is measured as interference from the task-irrelevant part, which is attenuated by misalignment, while holistic processing in the part-whole effect is reflected by an advantage making judgments of the whole compared to individual parts (Richler et al., 2012). Given these differences between different holistic processing measures, it is probably not very surprising that these tasks do not correlate strongly with each other.

The part-whole task and the composite task indeed reflect different mechanisms. One of the first accounts of holistic effects proposes representations in which internal parts are not differentiated/not independent representations, thus a global face template (Farah et al., 1998). However, this face template hypothesis does represent some problems. For instance, when novel objects are interleaved with faces, they are processed more holistically, as measured by a composite task, following an aligned face that is processed holistically, then a misaligned face that is not (Richler et al., 2009). This is challenging to explain with the help of a face template. Indeed, what effect would a unified perceptual representation of an aligned face have on the processing of a subsequent novel object that does not share the same configuration of features? The global face template hypothesis was also challenged by Chua, Richler, and Gauthier (2015), demonstrating that holistic effects can arise from experience with only diagnostic parts (Chua, Richler & Gauthier, 2015). These results gave rise to the learned attention account of holistic effects, which holds that congruence effects are the result of a history of learning to attend to the parts of objects: experience individuating objects leads to difficulty ignoring specific parts that have previously proven diagnostic. With faces, we cannot dissociate learning that many parts of a face contain diagnostic information from learning that these parts appear together (Gauthier, 2020).

We adopted this definition of holistic processing construct, hereafter HP-LA (holistic processing – learned attention), given evidence against the idea that internal parts are not differentiated/not independent representations (Farah et al., 1998), and the evidence in favor of the idea that it is difficult to ignore specific parts previously shown to be diagnostic for objects. The learned attention account of holistic effects is well evaluated by the selective attention composite effect. We thus chose to use the same generic composite task, but with three different versions, which may allow us to extract a latent variable that better reflects the underlying HP-LA

processing construct. Indeed, we used VHFPT-F (Richler et al., 2014) and a composite task with no repetition of face parts, including both sequential (thus with a memory component – encode facial stimuli, store them in working memory, retrieve them, and discern information about relevant facial features) and simultaneous (thus with a perceptual component – perceive facial stimuli and discern information about relevant facial features) versions. Including no repetition of face composites avoid correlations that are driven solely by a repetition of information in the composite task and in the memory task (CFMT) (cf. Richler et al., 2014). This design also allows us to evaluate specific correlations between perception-based and memory-based composite tasks, as well as face perception and face memory because we have measures of both perception and memory.

## Measuring face processing

When considering face processing, several individual differences studies have shown evidence for a processing factor common to face memory and face perception tests. Thus, several studies have found significant associations between different measures of face-identity processing, although using relatively few measures, for example, Boutet and Meinhardt-Injac (2021), who used the CFMT and an unfamiliar face matching test, GFMT; McCaffery et al. (2018), who also used a difficult recognition task, the Before They Were Famous Test; and Verhallen et al. (2017), who also used a Mooney Faces Test: Participants in the test are shown series of black and white distorted photographs, presented in such a way that would require them to perform closure. Moreover, these associations were typically larger than associations between identity-processing and other cognitive abilities. These findings have been interpreted as evidence for a general face-processing factor (f) (Verhallen et al., 2017).

Contrarily, several other individual differences studies by Sommer and colleagues have shown a dissociation of face processing into face memory and perception. For example, Wilhelm et al. (2010) used a latent variables approach and found a distinction between latent variables representing face perception and face recognition. In agreement with this, Fysh et al. (2020) used four face matching tasks and two face memory tasks and found evidence of at best modest generalization from one type of test to another. The distinction between face perception and face memory is also consistent with functional and neuroanatomical models of face cognition. Functional models (e.g., Calder & Young, 2005) distinguish between a structural encoding stage, which represents processes of face perception, and the stages of Face Recognition Unit activation, which mediate processes of face memory. Neuroanatomical models (e.g., Gobbini & Haxby, 2007) propose different brain areas underlying face perception and face memory.

## Present study

In the present study, we ask whether HP-LA and face processing are related. Given the apparent distinction between face perception and face memory, we evaluated both face perception and face memory together with HP-LA processing (composite task) each with three tasks. For memory and perception, we had three pairs of tasks such that within each pair (of memory and perception task), the stimuli involved, the requirement for matching across viewpoints, etc., are the same, such that the only difference is whether perception or memory is taxed. In this way, we will not have to be concerned that an eventual higher correlation between HP-LA processing and either perception or memory is attributable to the stimulus or task format overlap. Face perception was measured by indicators requiring perceptual comparisons of face stimuli without any reliance on memory processes. Face memory was assessed through measures that required the learning and recognition of face stimuli. The pairs of tasks were: (i) CFMT and CFPT, which use the same faces but in a memory versus perception task; (ii) Vanderbilt Face Matching Test (Sunday et al., 2015): we divided the test such that approximately half the trials were retained for the memory task, and approximately half were adapted for a perception task; (iii) Kent Face Matching Task (Fysh & Bindemann, 2018): we used 40 trials (the number of trials in the short version) for the perception task and adapted 40 trials for the memory task.

## Latent variable approach

Latent variable modeling, which allows researchers to estimate latent variables such as underlying abilities, through the common variance from several indicators (for instance, many different tasks), may be particularly productive (Tomarken & Waller, 2005; see Richler et al., 2019) in evaluating the relations between face HP-LA and face processing (cf. Russell et al., 1998, for the advantages of latent variable modeling). In the present study, we adopted this latent variable modeling approach and compared the relations between HP-LA processing and face processing. It is indeed crucial to use a range of tasks in the study of individual differences (e.g., Hildebrandt et al., 2013). When only two tasks are contrasted, for example, one holistic task and one face-processing task, as in previous findings showing discrepancies in the relations between holistic processing and face processing, (Konar et al., 2010; Rezliescu et al., 2017; Richler, Cheung, & Gauthier, 2011a), it is possible to make hasty interpretations of a type of pattern between them (Sunday et al., 2019); for example, the previous affirmations about whether holistic processing is related to face processing reviewed above.

Sommer's group has used a latent variable approach to face perception and memory but (i) tasks in each type were not obviously related (e.g., Wilhelm et al., 2010). In our study, and as mentioned earlier, we used three pairs of tasks for memory and perception, and the stimuli involved, the requirement for matching across viewpoints, and so on were all the same in each pair. (ii) Also importantly, previous work using a latent variable approach evaluated face perception and face memory but did not evaluate the hypothetical relations between holistic processing and face processing. In our study and considering the learned attention account of holistic processing (Gauthier, 2020), we included VHFPT-F (Richler et al., 2014) and a composite task with no repetition of face parts, including both sequential (thus with a memory component) and simultaneous (thus with a perceptual component) versions.

It is expected that the use of three indicators for HP-LA processing, face perception, and face memory would allow extraction of individual differences specific to the three constructs, allowing for a better examination of the relationship between HP-LA and face processing.

## Method

### Participants

A total group of 115 participants accepted our invitation for the study and completed all the tasks. A commonly accepted ratio for cases to items in SEM is 10 (Nunnally, 1978) to achieve an adequate level of statistical power, with a floor of 5 (Gorusch, 1983; Tinsley & Tinsley, 1987), but only if the total sample size approaches 300 or above (Devellis, 2017). Our proposed model is simple and does not include higher-order effects. We included a total of nine tasks and therefore nine indicator variables in the model, which means the minimum sample size required for adequate power, according the first rule-of-thumb, is 90. Therefore, the sample size of 115 was likely of sufficient power. Participants were psychology students from the University of Lisbon, all native speakers, and skilled readers of Portuguese, with normal or corrected-to-normal vision and hearing, who received a course credit. The study was conducted according to the guidelines of the Declaration of Helsinki and was approved by the Ethics Committee of the Department of Psychology of the University of Lisbon. All participants provided written informed consent. All tasks were performed online, and none had a time limit. Three tasks were run in JavaScript (CFMT, CFPT, VHFPT-F) and the other six tasks were run in Qualtrics (Qualtrics, Provo, UT, USA, <https://www.qualtrics.com>).

Please see Fig. 1 for depictions of the nine tasks

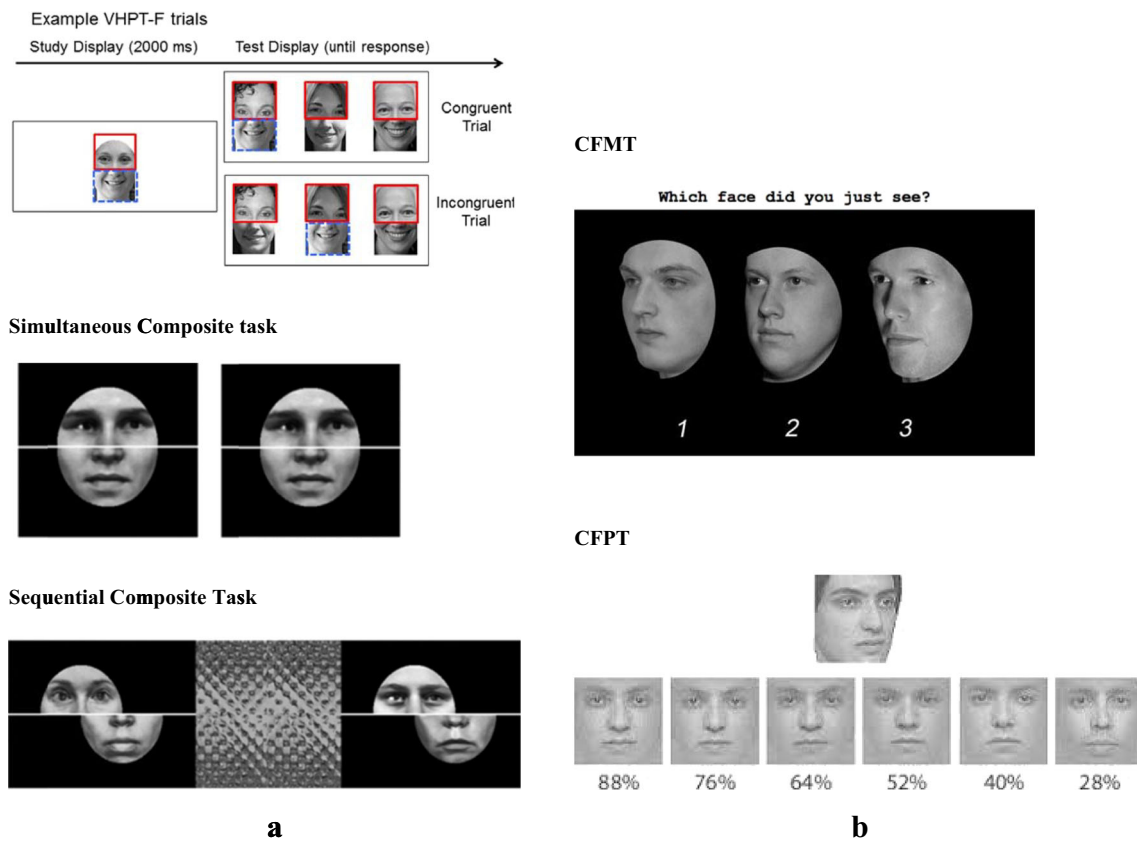


Fig. 1 Depiction of the tasks

### Sequential composite task

For the face composite task, 92 grayscale front-view images with neutral expressions from the MPI face database (Troje & Bühlhoff, 1996) were cropped to remove hair and ears. A total of 184 different face composites were created (with no repetition of face parts) for the trials that were same-congruent-aligned, same-incongruent-aligned, different-congruent-aligned, different-incongruent-aligned, same-congruent-misaligned, same-incongruent-misaligned, different-congruent-misaligned, different-incongruent-misaligned.

Aligned composites were used in a single aligned face stimulus block (92 trials). Another block with misaligned composite face stimuli (92 trials) was run (i.e., presentation of stimulus alignment was blocked). The order of the blocks was chosen at random. Face images were divided along a horizontal line at the bridge of the nose to create composite images. Participants completed eight practice trials with different stimuli using the same procedure as in the experimental trials. Throughout the experiment, the 1 and 2 keys on the keyboard were used by participants. Face stimuli were displayed in the center of the computer screen. First, there was a 1-s study stimulus. Then, for 1 s, a mask was shown, followed by the test stimulus for 1 s. The response screen

remained until a response was received. The sequential face composite task required participants to indicate whether the top two halves of the sequentially presented face stimuli were the same or different while ignoring the bottom halves.

### Simultaneous composite task

A total of 184 face composites (different from those in the sequential composite task), from 92 grayscale front-view images with neutral expressions from the MPI face database (Troje & Bühlhoff, 1996) (with no repetition of face parts) were made for the trials that were same-congruent-aligned, same-incongruent-aligned, different-congruent-aligned, different-incongruent-aligned, same-congruent-misaligned, same-incongruent-misaligned, different-congruent-misaligned, different-incongruent-misaligned.

In this task, two face composites were shown side by side, and participants had to decide whether the tops of both composites were the same or different by pressing one of two keys on a computer keyboard. The stimuli remained on the screen until a response was received. Eight practice trials were followed by two blocks, one aligned and one misaligned (each with 92 trials). The order of the blocks was randomized.

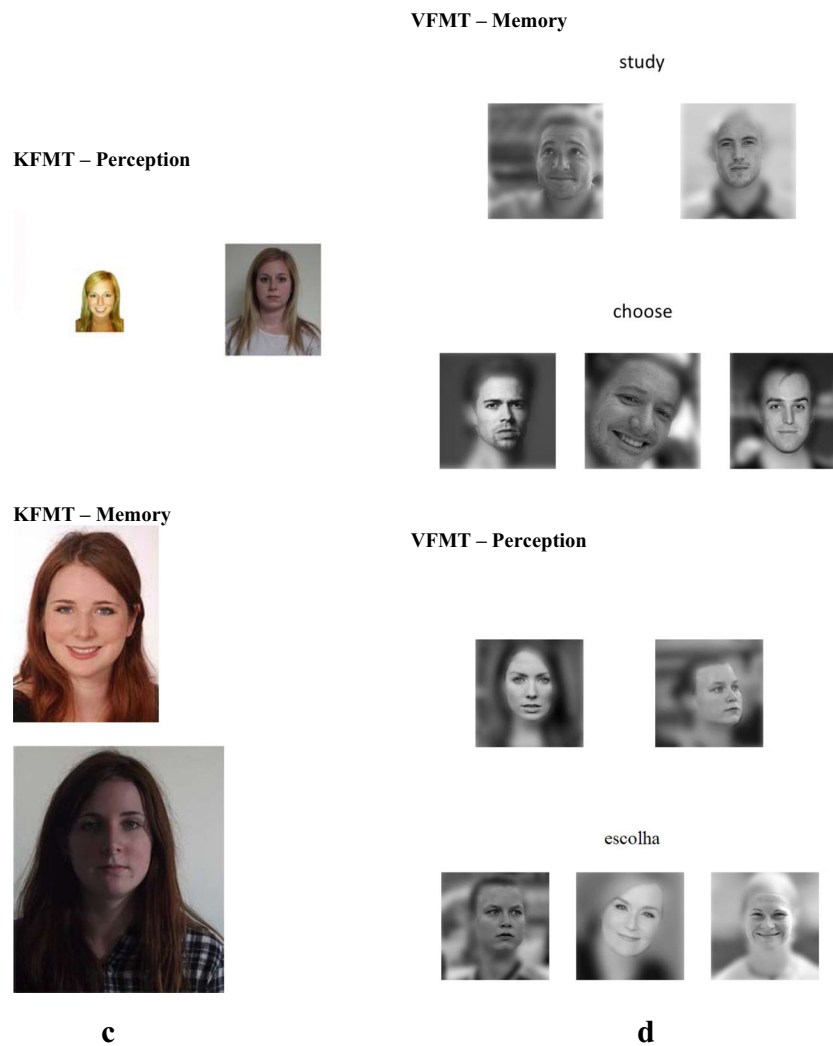


Fig. 1 (continued)

### VHFPT-F

We used the Vanderbilt Holistic Face Processing Test (VHFPT-F) (Richler et al., 2014).

A study composite face was shown for 2 s in each trial. A red outline box surrounded the target face part. Participants were instructed to concentrate only on the target area of the face and to ignore the rest of the face. For 1 s, the study face was replaced by a blank screen. Three composite faces were then displayed on the screen, one on the left, one in the center, and one on the right. A red box denoted the target area on each of the three faces. Participants were instructed to indicate which of the three composite faces contained the target segment with the same identity (but different image) as the study composite while ignoring the distractor segments.

The correct target part was paired with either the same distractor parts (congruent trial) or different distractor parts (incongruent trial). There is no manipulation of alignment in this task.

By pressing one of three buttons on the keyboard, participants were asked to indicate which face contained the same target part as the study face. The test display was shown until the participants responded. Within set, the location of the correct response was counterbalanced.

There were 180 trials in total (90 congruent, 90 incongruent), with 20 trials for each target segment condition. At the start of the experiment, there were three very simple introductory trials (two incongruent and one congruent). Muppet faces that were presented in color were used to create composites for the introductory trials. There were nine possible types of face parts that were shown in the following order: lower two-thirds of the face, upper two-thirds of the face, lower half, upper half, lowest third, uppermost third, eyes, mouth, and nose. The participants were informed of the target part to which they were to pay attention at the start of each of these nine blocks.

## CFMT and CFPT

CFMT (Duchaine & Nakayama, 2006): In this test, participants must memorize six unfamiliar male faces. They are then shown triads of faces and asked to separate the target face from two distractors in (i) 18 trials during the encoding phase, (ii) 30 trials under novel lighting and viewpoint conditions, and (iii) 24 trials with additional visual noise. There is no time limit on the task. In the first two phases, the faces were the exact same ones as the ones shown before. Before the second and third phases, participants review the six faces for 20 s.

CFPT (Duchaine & Nakayama, 2006): The same faces as CFMT are used in this test. The CFPT requires participants to order a series of faces for similarity to a target face on each trial, where the comparison stimuli include the target face morphed into several different faces to varying degrees.

## VFMT and VFMT-perception

VFMT: In the original VFMT (Sunday et al., 2019), two novel faces were presented for a 4-s study period. Following that, subjects were shown one face that matched the identity of one of the two study faces, as well as two distractor faces. The target face images used in both the study and the test were different versions of the same person. Subjects were instructed to use the keyboard to select the face that matched the identity of one of the study faces (1, 2, and 3). There was no time limit for responding, and male and female items were mixed. Three practice items with cartoon faces were provided. We included 54 trials of the original VFMT.

VFMT-perception: In the adaptation of the VFMT to a perception test, participants saw two novel faces at the top of the screen and three faces at the bottom of the screen. One of those three faces matched the identity of one of the two faces at the top of the screen. The target face images and perception test images were different images of the same person. Subjects were instructed to use the keyboard to select the face that matched the identity of one of the two faces at the top of the screen (1, 2, and 3). There was no time limit for responding, and male and female items were mixed. There were three practice items with cartoon faces. We adapted 56 different trials of the original VFMT.

## KFMT and KFMT-memory

KFMT: The KFMT (Fysh & Bindemann, 2018) includes a high-resolution portrait and a student ID photograph in each pair. The portrait images were cropped to show only the target's head and shoulders and placed on the right side of a blank white canvas. Student ID photos were placed to the left of the digital photos. As a result, each image pair in the KFMT is made up of an optimized target photograph taken under controlled conditions, like a passport photograph, and an ambient

photograph (student ID photograph) in which targets are depicted in a variety of poses and with various facial expressions.

We chose 40 image pairs of the KFMT, 20 of which depicted the same identity and the remaining 20 of which depicted different individuals (40 trials is the number of trials in the short version of KFMT). Fysh and Bindeman (2018) paired target images based on their visual similarity in terms of hair color, face, and brow shape to create these mismatch trials. Six practice trials were held. Participants gave their responses by pressing one of two keys on a computer keyboard. The trials were carried out at the participant's own pace.

KFMT-memory: In this adaptation of KFMT to a memory test, participants first saw a student ID photograph for 4 s, then saw an optimized photograph and had to decide whether it depicted the previous identity seen. Observers had no time limit and had to respond using one of two keys on a computer keyboard. We adapted 40 (the same number of trials as in the short version of KFMT) different KFMT image pairs, 20 of which depicted the same identity and the remaining 20 of which depicted different individuals. Six practice trials were held.

## Variables

Accuracy data of the face memory, face perception, and face HP-LA tasks from 115 participants were collected. For face HP-LA tasks, the dependent measures are the residuals of sensitivity  $A'$ , calculated by regressing the congruency effect of aligned trials on the congruency effect of misaligned trials, with a positive, larger value indicating an increase in face HP. Residuals were chosen over difference scores due to the higher reliability of residuals over difference scores in past studies (DeGutis et al., 2013). The resultant HP scores for composite tasks in all models depicted in this study were therefore derived from an aggregate measure, which was calculated from regressing the aligned congruency effect (the accuracy of all aligned congruent trials minus that of all aligned incongruent trials) on misaligned congruency (the accuracy of all misaligned congruent trials minus that of all misaligned incongruent trials). For face memory and face perception tasks, the raw scores were used except for CFPT, in which the scores were calculated only from upright face trials with the formula of  $144$  (maximum number of possible errors that could be made) – score (error made) of upright trials (the CFPT has an upright and an inverted presentation; we used only the upright presentation), as the raw scores were an inverse measure, and that upright face trials should reflect normal face processing more closely and lead to higher construct validity.

## Factor analyses and structural equation modeling

Model development and assessment were conducted with factor analyses and structural equation modeling (SEM), on R version 3.6.2, using the lavaan package version 0.6.7.

## Results

### Data preparation

Data filtering was conducted to remove outliers for which scores were in excess of 3 standard deviations (SD) from the mean of the group in each of the tasks. Data for two participants in the simultaneous composite task, three participants in the sequential composite task, and two participants from the VHFPT were discarded as the mean accuracy from all conditions was below 3 SD in relation to the whole group of participants. The data from CFPT for five participants were discarded as they achieved a score below 3 SD in relation to all participants.

Similarly, participants were removed if they had any of their scores removed in any of the tasks. This resulted in a total of nine participants being removed (7.8% of 115 participants), which is below the commonly accepted threshold of data missingness of 10% (Bennett, 2001).

### Descriptive statistics

The descriptive statistics of the measures used are shown in Table 1. In addition, the existence of expected HP-LA effects in the face HP measures was ascertained with inferential statistical tests. The interaction between alignment and congruency was significant in the simultaneous composite paradigm,  $F(1,105) = 55.316, p < .001, \eta^2 p = .345$ ; and significant alignment by congruency interaction was also detected in the sequential composite paradigm,  $F(1,105) = 68.754, p < .001, \eta^2 p = .396$ . Manipulation by misaligning faces did indeed produce observable disruption of HP, and therefore the residuals reflect the extent of face HP in each individual. For VHFPT, the congruency effect was also significant,  $t(105) = 18.97, p < .001, d = 1.842$ , showing HP in the form of the increased accuracy in congruent trials and decreased accuracy in incongruent trials. The group-level effects sizes reflected by the interaction effects in the two composite tasks and the congruency effect in VHFPT were fairly large in magnitude, suggesting that they were adequate in producing stable individual differences. The partial-eta squared ( $\eta^2 p$ ) of both interaction effects were larger than .3, while according to Cohen (1988), an effect with a partial-eta squared larger than .14 can be considered large. Cohen's  $d$  for congruency effect from VHFPT was 1.842, which was larger than the recommended threshold for Cohen's  $d$  that reflects a large effect from Cohen (1988), which was 0.8.

### Correlation and factor analyses

The current study examines the relationship between face memory, face perception and face HP-LA. As face HP-LA is theorized to be a hallmark for face recognition expertise,

we postulate that face memory and face perception are predicted by face HP-LA.

First, we looked at the correlation matrix (Table 2) of the nine measures for a basic evaluation of the factor structure. The correlations revealed that indicators for face memory and indicators for face perception have significant moderate cross-correlations within their respective constructs. Moreover, there were moderate zero-order correlations between the indicators of the face memory and face perception ( $r_s = .342 - .533$ ), comparable to the correlations between indicators within the same proposed constructs ( $r_s = .347 - .531$ ). With factor analysis we compared whether a 1-factor model is sufficient to describe the data structure and reflect the underlying general face processing, instead of a two-factor model separating the indicators into two latent variables. The Scree test and the chi-square goodness-of-fit test were conducted for the six indicators for face memory and face perception. The Scree test suggested a one-factor solution, as the acceleration factor (AF) was at 2, suggesting that the elbow of the scree plot appeared at 2 factors and only 1 factor should be extracted. The chi-square goodness-of-fit was also insignificant at 1 factor,  $\chi^2(9) = 2.96, p = .966$ , meaning that a 1-factor solution for the six indicators was adequate.

The indicators for the proposed face HP latent factor had very low cross-correlations, suggesting that they should be treated as reflecting three separate instead of one common theoretical construct.

### Structural equation modeling

In view of the correlation and factor analysis results, we generated a working model for structural equation modeling analysis (Fig. 2) so that the three indicators of face memory and the three indicators of face perception are now represented by one single latent variable that we termed “face processing.” We also separated the three HP indicators so that they point to different latent variables representing potentially different aspects of face HP. Results of analyses using the original, theory-driven model with separate face memory and perception constructs, as well as a common construct for all three HP-LA processing measures, are described in the Appendix in the Online Supplementary Material. Table 3 presents reliability estimates, composite reliability, average variance explained and loadings of indicators for face memory, face perception and face HP.

The goodness-of-fit indices of the current SEM model were further assessed. The model is identified with a positive degree of freedom, and the chi-square statistics of the likelihood-ratio test (LRT) of the specified model are not significant,  $\chi^2(24) = 18.767, p = .764$ , suggesting a very good fit of the model to the data structure. Indeed, CFI and TFI both exceed the recommended cut-off at 0.9 (Awang, 2012; Hair, 2010) at 1 and 1.041. RMSEA is also well below the cut-off of 0.07 (Hair, 2010) at 0 [95% confidence interval: 0–0.056]. In fact, the current model nearly fits perfectly with the data, which is implied from a CFI of



**Table 1** Descriptive statistics of indicators for face memory, face perception and face holistic processing (HP)

	Mean	SD	Range	Possible range
Face memory				
VFMT	31.84	4.674	22–44	0–54
KFMT-memory	28.226	3.594	20–35	0–40
CFMT	57.632	9.089	34–72	0–72
VFMT-perception	40.132	5.206	25–48	0–54
KFMT	29.415	3.852	18–38	0–40
CFPT	114.906	14.726	62–136	0–144
Face HP (accuracy)				
VHFPT – Congruent trials	0.779	0.061	0.61–0.92	0–1
VHFPT – Incongruent trials	0.613	0.087	0.32–0.81	0–1
Simultaneous Composite Task – Aligned congruent trials	0.983	0.024	0.86–1	0–1
Simultaneous Composite Task – Aligned incongruent trials	0.938	0.045	0.78–0.99	0–1
Simultaneous Composite Task – Misaligned congruent trials	0.98	0.021	0.91–1	0–1
Simultaneous Composite Task – Misaligned incongruent trials	0.967	0.023	0.9–0.99	0–1
Sequential Composite Task – Aligned congruent trials	0.968	0.033	0.85–1	0–1
Sequential Composite Task – Aligned incongruent trials	0.922	0.069	0.54–1	0–1
Sequential Composite Task – Misaligned congruent trials	0.954	0.037	0.8–1	0–1
Sequential Composite Task – Misaligned incongruent trials	0.96	0.04	0.75–1	0–1

1 and an RMSEA of 0. Critically, the Guttman's Lambda-2 reliability of the HP measures were all above 0.6, which meant at least 60% of the variances originated from true scores, and therefore the lack of intercorrelation of HP measures was unlikely to be due to unstable measurement.

Such near-perfect fit indices are not common, but do not necessarily suggest that the current model has common pitfalls with similar models, which might output such fit indices if they are underpowered, under-identified, or just identified. As the standardized regression weights between the face HP factors and face-processing factor indicated, the strength of the paths between the factors was low, and the collective variance explained from face HP to face processing is merely .07, which indicated low correlations between the factors. As the latent factors do not correlate with each other strongly, the model does not

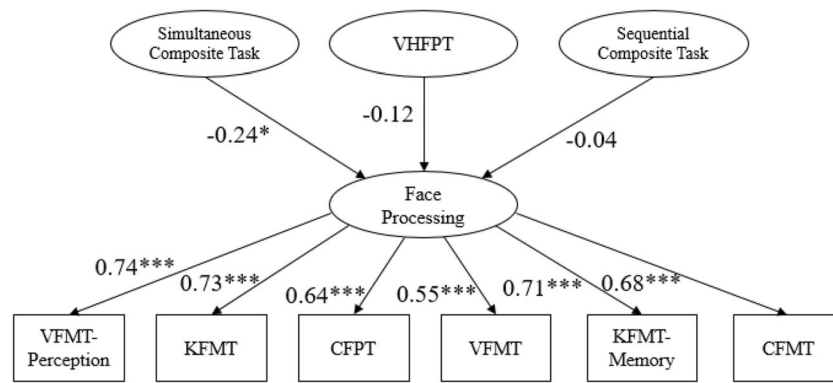
impose much restriction upon the alternative hypothesis of the Likelihood Ratio Test (LRT), therefore the current model would be able to fit the data more easily than more restricted models, and the chi-square test and relative fitness indices would not indicate much deviation from perfect fitness. Still, this does not diminish the utility of the current model to answer theoretical questions of our interest, which is whether there is a relationship between face HP and general face processing.

### Comparison of the final model and theory-driven model

While we reached a one-factor solution for face processing, it might still be of theoretical relevance to attempt to compare a

**Table 2** Zero-order correlations between the 9 indicator measures used in the present study. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

		VFMT	KFMT-memory	CFMT	VFMT-perception	KFMT	CFPT	VHFPT	Simultaneous composite task	Sequential composite task
VFMT	Pearson's r	—								
	p-value	—								
KFMT-memory	Pearson's r	0.347***	—							
	p-value	<.001	—							
CFMT	Pearson's r	0.416***	0.452***	—						
	p-value	<.001	<.001	—						
VFMT-perception	Pearson's r	0.408***	0.523***	0.533***	—					
	p-value	<.001	<.001	<.001	—					
KFMT	Pearson's r	0.397***	0.533***	0.477***	0.531***	—				
	p-value	<.001	<.001	<.001	<.001	—				
CFPT	Pearson's r	0.342***	0.498***	0.42***	0.449***	0.446***	—			
	p-value	<.001	<.001	<.001	<.001	<.001	—			
VHFPT	Pearson's r	-0.101	0.043	-0.038	-0.093	-0.129	-0.122	—		
	p-value	0.304	0.66	0.702	0.341	0.189	0.213	—		
Simultaneous composite task	Pearson's r	-0.064	-0.14	-0.134	-0.086	-0.307**	-0.18	-0.061	—	
	p-value	0.517	0.153	0.17	0.378	0.001	0.065	0.532	—	
Sequential composite task	Pearson's r	-0.097	-0.009	-0.077	-0.007	0.037	-0.009	0.03	-0.07	—
	p-value	0.324	0.929	0.435	0.944	0.708	0.924	0.763	0.473	—



**Fig. 2** The complete SEM model with a general face processing factor and three face holistic processing (HP) factors, along with the associated standardized path coefficients. All specified paths were shown in the model, which included no correlated residuals or correlated factors. Indicators were depicted with rectangles, while latent factors were

depicted with ellipses. Significant path coefficients are indicated with \* at the .05 level, \*\* at the .01 level, and \*\*\* at the .001 level. Face HP latent factors are equivalent to their single indicators, and therefore indicators are omitted

two-factor model with our final data-driven one-factor model. To compare our final model with a more theoretically cogent model where face perception and face memory are reflected by two separate constructs (Fig. 3), we computed the goodness-of-fit of our final model, as well as a model in which face processing is broken down according to our a priori categorization of face memory and face perception. A LRT for nested model was conducted to compare the goodnesses-of-fit of the two models. The difference in  $\chi^2(4) = 6.51, p = 0.164$ , indicated an insignificant increase in fit. Hence the smaller current model was able to sufficiently explain the current data structure. It would therefore be appropriate not to overly interpret the implications brought by the path relationship of the more complex model, based on the principle of parsimony for model selection. Still, the path between Simultaneous Composite Task and Face Perception is significant, while that between Simultaneous Composite Task and Face Memory is not, hinting at differential contributions of HP processes reflected by the Simultaneous Composite Task to different face processes.

The relationship between face HP and general face processing is partially supported, as the standardized regression

coefficient between the simultaneous composite task and the general face processing factor was small to moderate (-0.24) and significant at the .05 level, while the other two composite tasks did not achieve significance at the .05 level, with their paths to the face-processing factor. However, the significant correlation is negative. Thus, we found no evidence of a positive correlation between face processing and face-composite effect, casting doubts on the role of HP-LA processing in face recognition.

In sum, the results suggest the existence of a general face-processing factor over-arching the face memory and face-perception measures. Face processing seems not to be related to processes captured by composite tasks.

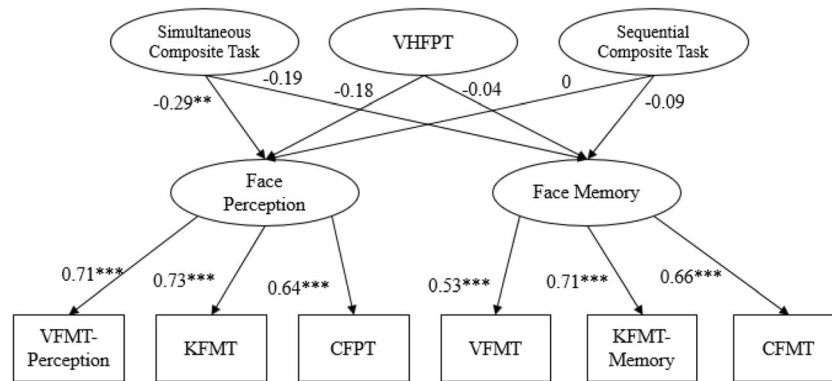
### General discussion

The main goal of the present study was to evaluate whether face holistic processing is related to face processing. We first discuss separately the results concerning face HP-LA and face processing, followed by an evaluation of their relation.

**Table 3** Reliability estimates, composite reliability (CR), average variance explained (AVE), and loadings of indicators for face memory, face perception, and face holistic processing (HP). For reliability, odd-even split-half reliabilities are presented unbolded, Cronbach's alphas are

presented bolded, while Guttman's  $\lambda_2$  are presented in italics. CRs, AVEs, and loadings are derived from simple measurement models of the face processing construct

	Reliability	CR	AVE	Loading
Face processing	<b>0.729</b>	0.833	0.43	
VFMT	0.443 ( <b>0.469</b> )			0.564
KFMT-memory	0.471 ( <b>0.447</b> )			0.614
CFMT	<b>0.721</b>			0.736
VFMT-perception	0.633 ( <b>0.719</b> )			0.732
KFMT	0.549 ( <b>0.524</b> )			0.726
CFPT	0.648 ( <b>0.639</b> )			0.614
VHFPT	<i>0.63</i>			0.381
Simultaneous Composite Task (Residual)	<i>0.616</i>			-0.164
Sequential Composite Task (Residual)	<i>0.634</i>			-0.187



**Fig. 3** Theory-driven model in which face processing is re-separated into face perception and face memory factors, along with the associated standardized path coefficients. All specified paths were shown in the model, which included no correlated residuals or correlated factors.

Indicators were depicted with rectangles, while latent factors were depicted with ellipses. Significant path coefficients are indicated with \*\* at the .01 level and \*\*\* at the .001 level

## Face holistic processing

Different effects (e.g., composite, part-whole, and inversion) have been commonly regarded as “holistic processing” effects, which implies that they all represent the same or highly overlapping holistic processing mechanisms (Piepers & Robbins, 2012). There are, however, significant differences between the mechanistic accounts proposed for the various holistic processing effects (cf. Richler et al., 2012). As we discussed in the *Introduction*, in the study of Rezlescu et al. (2017), the three holistic processing measures did not show a high overlap, with a significant and moderate correlation found only between inversion and part-whole effects. Additionally, Li et al. (2017) found distinct neural substrates for the part-whole effect and composite effect

In our study, we focused on the composite paradigm, given negative evidence for the idea that a face template might explain part-whole effects, but with three different versions. We used the VHFPT-F (Richler et al., 2014) and a sequential composite matching task and a simultaneous composite matching task both with no repetition of face parts. Composite effects were found in all three tasks, in the form of an alignment  $\times$  congruency interaction or a congruency effect (VHFPT-F). However, we discovered very small and non-significant correlations between the three composite tasks, despite the similar paradigms used. There were several differences in the details of the three tasks. For example, the format of VHFPT-F was different from the other two tasks, with nine instead of two possible relevant parts, and with three test faces instead of one test face in each trial. In addition, the study and target face parts were different images of the same identity in VHFPT-F, but for the sequential and the simultaneous matching tasks, the study and target face parts were identical. These differences in methodological details may contribute to introducing differences in what is being measured. In future studies comparing the three versions of the

composite task, study and target face parts in simultaneous and sequential matching tasks might be different images of the same identity

Our finding of separate processes measured by different versions of the composite task might also reflect that these tasks draw on different aspects or sub-processes of HP-LA face processing. A first possibility is that the perception-based composite task can be easily completed successfully by general visual/perceptual (not face-specific) mechanisms because the stimuli are presented side by side for an indefinite time. This interpretation is, however, challenged by the observation of significant alignment  $\times$  congruency interactions in both composite tasks. Also, when sequential and simultaneous composites tasks are considered together, the three-way interaction is not significant,  $F < 1$ . A second possibility is that HP-LA processing can be seen as more dependent on perceptual aspects of the face versus more dependent on working memory-related processes. This interpretation does not receive support because we found evidence of a negative correlation between simultaneous composite task and the perception-based/matching KFMT task face processing. Also, there was no pattern of correlations between memory-based holistic processes (sequential composite task) and face memory processes. Thus, the small overlap between the different versions of the composite effect does not reflect meaningful differences corresponding to different perception and memory requirements. A further possibility concerns the difference between early holistic processes versus experience-dependent holistic processes. Demonstrations of the importance of experience in establishing holistic processing have been called into question. Zhao et al. (2016) found face-like holistic processing for nonface, novel stimuli in the absence of perceptual expertise. Without any training, line patterns with salient Gestalt information (i.e., connectedness, closure, and continuity between parts) were processed as holistically as faces. The dual-route holistic processing account (Zhao et al., 2016) proposes

two paths to holistic processing: a stimulus-based route and an experience-based route. Thus, while there is no evidence of functional overlap in the processing of faces and novel gestalt stimuli in a task tapping experience-based contributions to holistic processing, that is a n-2 part-matching task with interpolated faces and novel gestalt stimuli (Curby et al., 2019), Curby and Moerel (2019) found evidence of reciprocal interference between the holistic processing of faces and novel gestalt stimuli in a task targeting early holistic mechanisms in which faces and Gestalt line patterns were superimposed. Thus, a future avenue of work might be the evaluation of the contribution of early versus experience-dependent processes for different composite tasks: for example, the simultaneous composite task may depend more on early processes than the sequential composite task or VHFPT.

In sum, Rezlescu et al. (2017) showed that there is no common set of mechanisms measured by the three putative holistic tests included in their study: inversion, part-whole, and composite task. In Rezlescu et al. (2017), the holistic tasks corresponded to different operational/mechanistic definitions of HP and had very low relations. Our results show additionally that different versions of the same composite task may introduce differences in what is being measured, and these differences should be fully investigated. Additionally, different composite tasks may draw on different aspects/subprocesses of holistic processing, such as early versus experience-dependent holistic processes, which also deserve future investigation.

## Face processing

With regard to face processing, and as discussed in the *Introduction*, individual difference studies have produced mixed results. Indeed, some studies have shown evidence for a face-processing factor common to face memory tests and face perception tests. Some studies have found significant associations between different measures of face-identity processing (Boutet & Meinhardt-Injac, 2021; McCaffery et al., 2018; Verhallen et al., 2017; Wilmer, 2017). Some other studies, using a latent variables approach (Fysh et al., 2020; Wilhelm et al., 2010), found differences between face perception and face memory constructs. Research on the normal development of face cognition suggests that face perception matures at the same rate as other object perception, whereas face memory appears to develop more slowly, over the first 10 years or more of life, and with a longer developmental trajectory than memory for other classes of objects (Weigelt et al., 2014).

We used a latent variable modeling approach, incorporating three pairs of face perception and face memory tasks such that within each pair (of memory and perception task), the stimuli involved, the requirement for matching across viewpoints, etc., are the same, such that the only difference is whether perception or memory is taxed. We did not find

evidence with our current data set for a distinction between face perception and face memory. Correlation analysis revealed that indicators for face memory and indicators for face perception have significantly moderate cross-correlations within their respective constructs. Moreover, there were moderate zero-order correlations between the indicators of face memory and face perception, comparable to the correlations between indicators within the same proposed constructs. With factor analysis, we confirmed that a one-factor model is thus sufficient to describe the data structure and reflect the underlying general face processing, instead of a two-factor model separating the indicators into two latent variables. Thus, we found evidence for an over-arching single latent variable, “face processing,” which seemed to represent performance in the six-face perception and memory tasks well. Our results lend further support for the existence of a general face factor *f* (McCaffery et al., 2018; Verhallen et al., 2017; Wilmer et al., 2014). While there is previous evidence that face perception and face memory performance may depend on overlapping cognitive mechanisms (cf. the *Introduction* and, e.g., Fysh & Bindemann, 2018; Robertson et al., 2017), other work suggests that the underlying processes can also operate independently but especially so for average and high performers (e.g., Bate et al., 2018; Fysh, 2018). In a recent study, Fysh et al. (2020) used several perception and memory tasks and found that among both high (top 5%) and low (low 5%) performers, there was limited evidence for generalization between tests. This means that the top-performing individuals in, for example, the CFMT, might show little evidence of superior ability in the other tasks and that it is possible for one person to be deficient in one measure of face matching but not in another measure of face matching or face memory. An interesting avenue of future research is to evaluate whether average individuals will show more homogeneous performance profiles across the subprocesses and tests of face cognition.

## Relation between holistic processing and face processing

We hypothesized that individual differences in face processing would be predicted by face HP-LA, considering the benefits of holistic processing for the individuation of highly similar objects, an assumption that is also common in the literature (e.g., Richler et al., 2012). Previous studies evaluating this relationship used one indicator of HP and one indicator of face processing (CFMT) with contrasting results: for example, while Konar et al. (2010) found no association between HP and CFMT, Richler, Cheung, and Gauthier (2011a), on the other hand, found a small but significant correlation between holistic processing and CFMT performance. When only two tasks are contrasted, it is possible to make hasty conclusions about the pattern found between them (Sunday et al., 2019). Rezlescu et al. (2017) used three different HP tasks but only a

face perception task (CFPT), and the relationships between holistic effects and face perception ability varied greatly, with correlations ranging from moderate (inversion effect) to weak (the part-whole effect) to nonexistent (the composite effect – perceptual integration design).

Our results in a latent variable approach with three pairs of perception-memory tasks and three versions of the composite task showed evidence of a correlation between HP-LA captured by the simultaneous composite task and general face processing, but it was a small and negative correlation. Thus, HP does not seem to be predictive of better face processing per se. Considering holistic processing theories, and as discussed in the *Introduction*, the template account of holistic processing faces several challenges (Chua, Richler, & Gauthier, 2015; Richler et al., 2009). We adopted the learned attention account of holistic processing, according to which the composite face effect arises due to an overlearned pattern of attention when presented with an upright aligned face. But the fact that HP does not predict face processing may suggest that HP-LA is highly saturated in most individuals due to strong experience with faces (Gauthier, 2020). This means that HP-LA may be a hallmark of a certain type of information entering face computations, but variations within the normal range are not necessarily related to efficiency of face processing. A certain level of experience with faces could be more than enough for high levels of holistic processing.

Another possibility worth examining is whether the composite face effect truly reflects holistic processing. The first question is whether information from the top and bottom parts are processed sequentially (serially), concurrently (parallel), or integrated such that evidence from the two sections coalesces into a single channel (coactive). Fitousi (2015) and Cheng et al. (2018) discovered scant evidence for a coactive system and, hence, for true integration of information. These findings cast some doubt on the holistic concept of information integration (Fitousi, 2015).

Another related possibility is that the composite face task is not a pure measure of holistic processing but rather a manifestation of more general abilities, such as working memory and a broader attentional strategy similar to object-based attention (Fitousi, 2015). When an object is detected, all of its constituent features or parts are activated. The decrease in the congruence composite effect associated with misalignment could be explained by a similar disruption of objecthood (Fitousi, 2015). Analysis of theoretical ex-Gaussian parameters of RT distributions (Fitousi, 2020) reveals that the composite effect is generated by pure changes in the exponential component of the ex-Gaussian distribution, suggesting the involvement of attentional and working memory processes in the composite face effect.

One should be careful in drawing general conclusions regarding holistic processing and the relation between holistic processing and face processing as they may be related to the

specific tasks used. Nevertheless, we used two operationalizations of one of the most popular tasks used to measure holistic processing (composite task) and we used a very reliable measure of holistic processing (VHPT-F).

Finally, we acknowledge that the initial theory-driven model and also the model with separate face HP constructs could be somewhat affected by power issues according to some form of understanding of SEM power calculation. In our original formulation of power analyses, we considered Nunnally's (1978) case-to-item ratio, and also generally Bentler and Chou's (1987) recommendations for a 5:1 ratio of case-to-free parameters (which was 21 in the final model, which therefore called for about 105 participants). We acknowledge that other recommendations call for a higher case-to-free-parameter, and case-to-latent-variable ratio, as well as other formulations that take into account more precise parameters than the more rudimentary rules of thumb would suggest (<https://www.danielsoper.com/statcalc/default.aspx>).

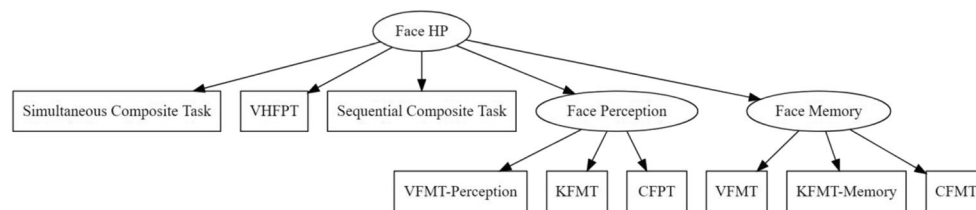
We calculated a post hoc estimate of the achieved power of our final model at  $\alpha = .05$  with the *semPower* R module, using the Goodness-of-Fit index (GFI = .964), degrees of freedom (24), as well as the number of observed variables (9) and number of observations (106). This gives rise to an achieved power of .662. We acknowledge that it fell short of the usually targeted power of .8, however, we think that it still carries a certain degree of robustness against type-II error. Meanwhile, the model with two factors in the Appendix had an achieved power of .472.

In conclusion, we found evidence for an overarching latent variable, “face processing,” that seemed to accurately represent performance in the face perception and face memory tasks. Our discovery of small and non-significant correlations between different versions of the composite task suggests that (i) differences in methodological details may contribute to introducing differences in what is being measured in otherwise very similar composite tasks, and/or (ii) different composite tasks may reflect differential contributions of holistic sub-processes, such as early versus experience dependent. We found no evidence that HP-LA (as captured by three composite tasks) is predictive of better face processing per se, casting doubts on the role of holistic processing in predicting differences in face processing or in the interpretation of the concept of holistic processing

## Appendix

### Data fitting of the initial theory-driven model

This theory-driven model presents face holistic processing, face memory, and face perception as latent factors, each derived from three indicators. The relationship between the



**Fig. 4** The proposed initial model. Indicators were depicted with rectangles, while latent factors were depicted with ellipses

latent factors is examined through linear structural equation modeling (SEM). The two face processing factors would in theory reflect two distinct components of face processing, one memory-based and one perception-based. They would in turn be potentially related to face holistic processing as part of the mechanism for better face expertise. The predicted relationship between the three latent variables is shown in Appendix Fig. 4.

Composite reliabilities (CR) and Average Variance Explained (AVE) of the proposed constructs were computed, as an assessment of internal consistency of each proposed construct. Of note, the split-half reliability and Cronbach's alpha for CFPT are calculated by the correlation between up and inverted blocks, as only aggregated block data for each condition was available. For CFMT, the Cronbach's alpha was calculated as the test was conducted for three test conditions, and an alpha is computed across the three raw scores for each participant.

**Table 4** Reliability estimates, composite reliability (CR), average variance explained (AVE), and loadings of indicators for face memory, face perception, and face HP. For reliability, odd-even split-half reliabilities are presented unbolded, Cronbach's alphas are presented bolded,

	Reliability	CR	AVE	Loading
Face memory	<b>0.592</b>	0.675	0.412	
VFMT	0.45 ( <b>0.469</b> )			0.564
KFMT-memory	0.465 ( <b>0.447</b> )			0.614
CFMT	<b>0.721</b>			0.736
Face perception	<b>0.528</b>	0.733	0.48	
VFMT-perception	0.633 ( <b>0.719</b> )			0.732
KFMT	0.551 ( <b>0.524</b> )			0.726
CFPT	0.648 ( <b>0.639</b> )			0.614
Face HP	<b>0</b> [ $< 0$ ]	0 [ $< 0$ ]	0.069	
VHFPT	0.63			0.382
Simultaneous Composite Task (Residual)	0.616			-0.161
Sequential Composite Task (Residual)	0.634			-0.185

#### Appendix Table 4

Cronbach's alphas for face memory and face perception were lower than 0.7, a threshold which was suggested by Nunnally (1978). The composite reliabilities of face memory approached the recommended value of 0.7 by Hair (2010), while that of the face perception construct reaches and exceeds the threshold, and therefore is acceptable. The measurement of construct validity, AVEs, approached the threshold of 0.5 (Fornell & Larcker, 1981) for face memory and face perception. Face HP did not have a non-zero positive reliability for Cronbach's alpha, composite reliability and it had a low AVE. Also, the loadings of the indicators were low and possess different signs, which further pointed to the need of revising the measurement model for the construct as the three indicators likely reflected different latent constructs instead of one.

while Guttman's  $\lambda^2$  are presented in italics. CRs, AVEs, and loadings are derived from simple measurement models (three indicators for one latent variable) of each proposed construct

**Code availability** EPrime code is available upon reasonable request.

**Authors' contributions** Given it's optional, we prefer not to detail each author contribution.

**Funding** This work was supported by the Research Center for Psychological Science (CICPSI), Faculdade de Psicologia, Universidade de Lisboa, Lisboa, Portugal

**Data availability** All the data for the study can be found on the Open Science Framework at: <https://osf.io/5kh9p/>

## Declarations

**Competing interests** The authors have no relevant financial or non-financial interests to disclose.

**Ethics approval** The study was conducted according to the guidelines of the Declaration of Helsinki and was approved by the Ethics Committee of the Department of Psychology of the University of Lisbon.

**Consent to participate** All participants provided written informed consent.

**Consent for publication** Not applicable.

## References

- Awang, Z. (2012). *Structural equation modeling using AMOS graphic*. Penerbit Universiti Teknologi MARA.
- Bate, S., Frowd, C., Bennetts, R., Hasshim, N., Murray, E., Bobak, A. K., Wills, H., & Richards, S. (2018). Applied screening tests for the detection of superior face recognition. *Cognitive Research*, 3, 22. <https://doi.org/10.1186/s41235-018-0116-5>
- Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health*, 25(5), 464–469. <https://doi.org/10.1111/j.1467-842X.2001.tb00294.x>
- Bentler, P. M., & Chou, C. P. (1987). Practical Issues in Structural Modeling. *Sociological Methods & Research*, 16, 78–117. <https://doi.org/10.1177/0049124187016001004>
- Boutet, I., & Meinhardt-Injac, B. (2021). Measurement of individual differences in face-identity processing abilities in older adults. *Cognitive Research: Principles and Implications*, 6, 48. <https://doi.org/10.1186/s41235-021-00310-4>
- Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow face matching test. *Behavior Research Methods*, 42, 286–291. <https://doi.org/10.3758/BRM.42.1.286>
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews. Neuroscience*, 6(8), 641–651. <https://doi.org/10.1038/nrn1724>
- Cheng, X. J., McCarthy, C. J., Wang, T. S., Palmeri, T. J., & Little, D. R. (2018). Composite faces are not (necessarily) processed coactively: A test using systems factorial technology and logical-rule models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(6), 833–862.
- 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, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge Academic.
- 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>
- Curby, K. M., Huang, M., & Moerel, D. (2019). Multiple paths to holistic processing: Holistic processing of Gestalt stimuli do not overlap with holistic face processing in the same manner as do objects of expertise. *Attention, Perception & Psychophysics*, 81(3), 716–726. <https://doi.org/10.3758/s13414-018-01643-x>
- DeGutis, J., Mercado, R. J., Wilmer, J., & Rosenblatt, A. (2013). Individual differences in holistic processing predict the own-race advantage in recognition memory. *PLoS One*, 8(4), e58253. <https://doi.org/10.1371/journal.pone.0058253>
- DeVellis, R. F. (2017). *Scale Development: Theory and Applications* (4th ed.). Sage.
- 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>
- Duchaine, B., & Nakayama, K. (2006). The Cambridge face memory test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, 44(4), 576–585. <https://doi.org/10.1016/j.neuropsychologia.2005.07.001>
- 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>
- Fitousi, D. (2015). Composite faces are not processed holistically: Evidence from the Garner and redundant target paradigms. *Attention, Perception, & Psychophysics*, 77(6), 2037–2060. <https://doi.org/10.3758/s13414-015-0887-4>
- Fitousi, D. (2020). Decomposing the composite face effect: Evidence for non-holistic processing based on the ex-Gaussian distribution. *Quarterly Journal of Experimental Psychology*, 73, 819–840.
- Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, 18(3), 382–388. <https://doi.org/10.2307/3150980>
- Fysh, M. C. (2018). Individual differences in the detection, matching and memory of faces. *Cognitive Research*, 3, 20. <https://doi.org/10.1186/s41235-018-0111-x>
- Fysh, M. C., & Bindemann, M. (2018). The Kent Face Matching Test. *British Journal of Psychology*, 109, 219–231. <https://doi.org/10.1111/bjop.12260>
- Fysh, M. C., Stacchi, L., & Ramon, M. (2020). Differences between and within individuals, and subprocesses of face cognition: Implications for theory, research and personnel selection. *Royal Society Open Science*, 7, 200233. <https://doi.org/10.1098/rsos.200233>
- Gauthier, I. (2020). What we could learn about holistic face processing only from nonface objects. *Current Directions in Psychological Science*, 29(4), 419–425. <https://doi.org/10.1177/0963721420920620>
- 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., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal*

- of *Experimental Psychology: Human Perception and Performance*, 28(2), 431–446.
- Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, 45(1), 32–41. <https://doi.org/10.1016/j.neuropsychologia.2006.04.015>
- Goffaux, V., & Rossion, B. (2006). Faces are "spatial"—holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, 32(4), 1023–1039. <https://doi.org/10.1037/0096-1523.32.4.1023>
- Gorusch, R. L. (1983). *Factor Analysis* (2nd ed.). Lawrence Erlbaum Associates.
- Hair, J. F. (2010). *Multivariate data analysis: A global perspective*. Pearson Education.
- Hildebrandt, A., Wilhelm, O., Herzmann, G., & Sommer, W. (2013). Face and object cognition across adult age. *Psychology and Aging*, 28(1), 243–248. <https://doi.org/10.1037/a0031490>
- Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception*, 23(1), 65–74. <https://doi.org/10.1068/p230065>
- Konar, Y., Bennett, P. J., & Sekuler, A. B. (2010). Holistic processing is not correlated with face identification accuracy. *Psychological Science*, 21(1), 38–43. <https://doi.org/10.1177/0956797609356508>
- Li, J., Huang, L., Song, Y., & Liu, J. (2017). Dissociated neural basis of two behavioral hallmarks of holistic face processing: The whole-part effect and composite-face effect. *Neuropsychologia*, 102, 52–60. <https://doi.org/10.1016/j.neuropsychologia.2017.05.026>
- McCaffery, J. M., Robertson, D. J., Young, A. W., & Burton, A. M. (2018). Individual differences in face identity processing. *Cognitive Research: Principles and Implications*, 3, 21. <https://doi.org/10.1186/s41235-018-0112-9>
- McKone, E., Crookes, K., Jeffery, L., & Dilks, D. D. (2012). A critical review of the development of face recognition: Experience is less important than previously believed. *Cognitive Neuropsychology*, 29(1–2), 174–212. <https://doi.org/10.1080/02643294.2012.660138>
- 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>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
- Piepers, D. W., & Robbins, R. A. (2012). A review and clarification of the terms "holistic", "configural", and "relational" in the face perception literature. *Frontiers in Psychology*, 3, 559. <https://doi.org/10.3389/fpsyg.2012.00559>
- 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. <https://doi.org/10.1037/xhp0000400>
- 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., Bukach, C. M., & Gauthier, I. (2009). Context influences holistic processing of nonface objects in the composite task. *Attention, Perception & Psychophysics*, 71(3), 530–540. <https://doi.org/10.3758/APP.71.3.530>
- Richler, J. J., Cheung, O. S., & Gauthier, I. (2011a). Beliefs alter holistic face processing ... if response bias is not taken into account. *Journal of Vision*, 11(13), 17. <https://doi.org/10.1167/11.13.17>
- Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology*, 3, Article 553. <https://doi.org/10.3389/fpsyg.2012.00553>
- 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), 15 1–14. <https://doi.org/10.1167/14.11.10>
- Richler, J. J., Tomarken, A. J., Sunday, M. A., Vickery, T. J., Ryan, K. F., Floyd, J. R., ... Gauthier, I. (2019). Individual differences in object recognition. *Psychological Review*, 126(2). <https://doi.org/10.1037/rev0000129>
- Robertson, D. J., Jenkins, R., & Burton, A. M. (2017). Face detection dissociates from face identification. *Visual Cognition*, 25, 740–748. <https://doi.org/10.1080/13506285.2017.1327465>
- 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>
- Russell, D. W., Kahn, J. H., Spoth, R., & Altmaier, E. M. (1998). Analyzing data from experimental studies: A latent variable structural equation modeling approach. *Journal of Counseling Psychology*, 45(1), 18–29. <https://doi.org/10.1037/0022-0167.45.1.18>
- Sunday, M., Richler, J., & Gauthier, I. (2015). The Vanderbilt face matching test (VFMT 1.0). *Journal of Vision*, 15(12), 168–168.
- Sunday, M. A., Patel, P. A., Dodd, M. D., & Gauthier, I. (2019). Gender and hometown population density interact to predict face recognition ability. *Vision Research*, 163, 14–23. <https://doi.org/10.1016/j.visres.2019.08.006>
- Tinsley, H. E., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, 34, 414–424. <https://doi.org/10.1037/0022-0167.34.4.414>
- Tomarken, A. J., & Waller, N. G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology*, 1, 31–65. <https://doi.org/10.1146/annurev.clinpsy.1.102803.144239>
- Troje, N. F., & Bühlhoff, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research*, 36(12), 1761–1771. [https://doi.org/10.1016/0042-6989\(95\)00230-8](https://doi.org/10.1016/0042-6989(95)00230-8)
- 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. <https://doi.org/10.1016/j.visres.2016.12.014>
- Weigelt, S., Koldewyn, K., Dilks, D. D., Balas, B., McKone, E., & Kanwisher, N. (2014). Domain-specific development of face memory but not face perception. *Developmental Science*, 17(1), 47–58. <https://doi.org/10.1111/desc.12089>
- Wilhelm, O., Herzmann, G., Kunina, O., Danthiir, V., Schacht, A., & Sommer, W. (2010). Individual differences in perceiving and recognizing faces - One element of social cognition. *Journal of Personality and Social Psychology*, 99(3), 530–548. <https://doi.org/10.1037/a0019972>
- Wilmer, J. B. (2017). Individual differences in face recognition: A decade of discovery. *Current Directions in Psychological Science*, 26(3), 225–230. <https://doi.org/10.1177/0963721417710693>
- Wilmer, J. B., Germine, L. T., & Nakayama, K. (2014). Face recognition: A model specific ability. *Frontiers in Human Neuroscience*, 8(769). <https://doi.org/10.3389/fnhum.2014.00769>
- Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, 16(6), 747–759.
- Zhao, M., Bühlhoff, H. H., & Bühlhoff, I. (2016). 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>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.