In press at: *Psychonomic Bulletin and Review*

Are the facial gender and facial age variants of the composite face illusion products of a common mechanism?

Katie L.H. Gray[1*], Yvonne Guillemin[2], Zarus Cenac[2], Sophie Gibbons[1], Tim Vestner[3], Richard Cook[3]

[1]School of Psychology and Clinical Language Sciences,
University of Reading, Reading, U.K.

[2]Department of Psychology, City,
University of London, London, U.K.

[3]Department of Psychological Sciences,
Birkbeck, University of London, London, U.K.

*Corresponding author:
k.l.h.gray@reading.ac.uk
School of Psychology and Clinical Language Sciences,
University of Reading,
Reading, U.K.

Abstract

When the upper half of one face ('target region') is spatially aligned with the lower half of another ('distractor region'), the two halves appear to fuse together perceptually, changing observers' subjective perception of the target region. This 'composite face illusion' is regarded as a key hallmark of holistic face processing. Importantly, distractor regions bias observers' subjective perception of target regions in systematic, predictable ways. For example, male and female distractor regions make target regions appear masculine and feminine; young and old distractor regions make target regions appear younger and older. In the present study, we first describe a novel psychophysical paradigm that yields precise reliable estimates of these perceptual biases. Next, we use this novel procedure to establish a clear relationship between observers' susceptibility to the age and gender biases induced by the composite face illusion. This relationship is seen in a lab-based sample (N = 100) and replicated in an independent sample tested online (N = 121). Our findings suggest that age and gender variants of the composite illusion may be different measures of a common structural binding process, with an origin early in the face processing stream.

Key words:

Composite face illusion; Facial gender; Facial age; Individual differences; Psychophysics

Introduction

Upright faces are thought to engage holistic processing whereby local features are integrated into a unified whole for the purposes of accurate and efficient interpretation (Farah, Wilson, Drain, & Tanaka, 1998; McKone & Yovel, 2009; Piepers & Robbins, 2013). In the absence of canonical first-order facial information (e.g., when judging inverted or feature-scrambled faces), observers may be forced to base perceptual decisions on a piecemeal analysis of local features. The extent to which individuals process faces holistically is thought to determine their ability to recognise and interpret faces (DeGutis, Cohan, & Nakayama, 2014; Richler, Cheung, & Gauthier, 2011).

The composite face illusion is a key hallmark of holistic face processing. When the upper half of one face (the "target region") is spatially aligned with the lower half of another (a task-irrelevant "distractor region"), the two halves appear to fuse together perceptually, changing observers' subjective perception of the target region (Young, Hellawell, & Hay, 1987). The perceptual fusion observed is greatly reduced where one half is offset horizontally ('misaligned') or where aligned arrangements are presented upside-down. This feature of the illusion suggests that the human visual system integrates distal facial information only in the presence of an intact faciotopy (Murphy, Gray, & Cook, 2017; Rossion, 2013).

The composite face illusion is not merely an interference effect where distractor regions hinder perceptual decisions about a target. Importantly, distractor regions bias observers' subjective perception of target regions in systematic, predictable ways (Rossion, 2013). For example, male and female distractor regions make target regions appear masculine and feminine (Baudouin & Humphreys, 2006); young and old distractor regions make target regions appear younger and older (Hole & George, 2011), and happy and unhappy distractor regions make target regions seem happier and sorrowful (Calder, Young, Keane, & Dean, 2000).

It remains uncertain whether the different perceptual biases induced by composite face arrangements are different measures of a common structural binding process, or whether these biases should be thought of as independent illusory effects. Leading theoretical models posit that observers form a structural representation early in the face processing stream (Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000), that forms a common basis for judgements about various facial attributes (e.g., identity, expression, age, and gender). This initial modeling of face structure may generate the different biases induced by the composite face illusion. Alternatively, different perceptual biases may be products of attribute-specific holistic processing that occurs later in the face processing stream; for example, the holistic processing of facial age (Hole & George, 2011) or facial gender (DeGutis, Chatterjee, Mercado, & Nakayama, 2012), *per se*.

We sought to distinguish these possibilities by comparing individuals' susceptibility to the age and gender biases induced by the composite face illusion. If these biases are different measures of the same structural binding process, individuals' susceptibility to one bias should predict their susceptibility to the other. However, if these illusory biases arise from attribute-specific holistic processing, we might expect little or no association between observers' susceptibility to these biases.

Measuring the composite face illusion using psychophysics

We measured individuals' susceptibility to the age and gender biases induced by the composite face illusion using a novel psychophysical procedure (Figure 1). In our paradigm, the manifestation of the illusion is inferred from shifts in observers' psychometric functions. Examining how different viewing conditions modulate psychometric functions has helped vision scientists document the behavior of other illusions in this field, notably facial aftereffects (Leopold, O'Toole, Vetter, & Blanz, 2001; Webster & MacLeod, 2011). This approach also offers several advantages to those seeking to study the composite face illusion. In particular, researchers can dissociate the extent to which a distractor biases observers' perception in a particular direction – the critical measure of the composite face illusion – from observers' ability to detect and interpret the physical differences between target regions (i.e., the amount of internal noise associated with judgements about target regions). Crucially, the modeling of psychometric functions yields separate estimates of these independent parameters – the point of subjective equality (PSE) and function slope, respectively.

In both variants of our task, observers were asked to make binary categorisation judgements about upper face halves (encompassing the eyes; target regions) drawn from morph continua, while ignoring lower face halves (encompassing the mouth; distractor regions) that were not subject to morphing (i.e., presented at 100% intensity). In the gender variant of the task, target regions were drawn from a continuum blending an average male face with an average female face. Observers judged whether target regions depicted a male or female face, either in the presence of a distractor region cropped from the average male face, or a distractor region cropped from the average female face. In the age variant of the task, target regions were drawn from a continuum blending an average child face with an average adult face. Observers judged whether target regions depicted a child or adult face, either in the presence of a distractor region cropped from the average child face, or a distractor region cropped from the average adult face.

Figure-1

In both variants, participants' responses were used to construct psychometric functions that modeled the relationship between response probability and the strength of the signal present in the target region (% female or % adult). Under our psychophysical approach, observers' susceptibility to the illusion is inferred from the extent to which their psychometric functions diverge in different distractor conditions, inferred from the difference in PSE. To date, it has proved extremely difficult to measure individuals' susceptibility to the composite face illusion in a reliable way (Richler & Gauthier, 2014). Importantly, however, estimates of test-retest reliability obtained using our psychophysical paradigm range from $r = .602$ to $r = .775$ (see supplementary material).

Replicating the normative properties of the illusion

As described above, the composite face illusion manifests disproportionately when distractor and target regions are spatially aligned, and arrangements are presented upright. Little, if any, illusory distortion is seen when target and distractor regions are misaligned, or when aligned arrangements are presented upside-down (Murphy et al., 2017; Rossion, 2013). Although the focus of the present paper is inter-observer variability in composite illusion susceptibility, we first sought to confirm that our psychophysical paradigm replicated these key features of the illusion.

Two samples of typical observers completed the gender (N = 19, $M_{\text{age}} = 26$; $SD_{\text{age}} = 7.32$; 5 males) and age (N = 19, $M_{\text{age}} = 25$; $SD_{\text{age}} = 2.11$; 7 males) variants of our task. All participants had normal or corrected-to-normal vision and were tested in person, under controlled lab conditions. For all experiments described, ethical clearance was granted by the local ethics committee and the studies were conducted in line with the ethical guidelines laid down in the 6th (2008) Declaration of Helsinki.

Materials & procedure

The average male and female faces, and the average child and adult faces used to construct the morph continua were composites of eight faces sourced from the Radboud Face Database (Langner et al., 2010). Each continuum comprised seven levels that varied stimulus intensity from 20% to 80% in increments of 10%. Morphing was achieved through Morpheus Photo Morpher Version 3.11 (Morpheus Software, Inc). Facial composites subtended ~6° vertically when viewed at 58 cm. In the misaligned condition, target and distractor halves were offset horizontally by ~3°. A thin grey line (~4 pixels) was inserted in between the target and distractor to help participants distinguish the to-be-judged regions (Rossion & Retter, 2015).

Each trial began with a fixation point, followed by a composite arrangement presented for 1200ms. Participants registered their categorization decision with a key-press response. Participants judged the target under three configuration conditions: aligned distractors within an upright arrangement, misaligned distractors within an upright arrangement, and aligned distractors within an inverted arrangement. For each condition we constructed a psychometric function from 140 categorization decisions (7 target levels × 20 presentations). In total observers therefore completed 840 trials (140 trials × 3 configuration conditions × 2 levels of distractor). Experimental programs were written in MATLAB (The MathWorks, Inc) using Psychtoolbox (Brainard, 1997; Pelli, 1997). Psychometric functions were modeled by fitting cumulative Guassian functions using the Palamedes toolbox (Prins & Kingdom, 2009). Our measure of function slope was the reciprocal of the standard deviation of the symmetric Gaussian distribution underlying each cumulative Gaussian function. Goodness of fit was evaluated using *p*Dev statistics.

Results & Discussion

Gender variant. The distribution of PSEs was analysed using ANOVA with Distractor (female, male) and Configuration (aligned, misaligned, inverted) as within-subjects factors (Figure 2a). The analysis revealed a significant main effect of Distractor [$F(1,18) = 5.47$, $p = .031$; $\eta_p^2 = .23$], a non-significant main effect of Configuration [$F(2, 36) = .49$, $p = .62$; $\eta_p^2 = .03$], and a significant Distractor × Configuration interaction [$F(2, 36) = 5.78$, $p < .01$; $\eta_p^2 =$

.24]. Planned comparisons revealed a significant PSE shift when composites were upright and aligned [$t(18) = 3.12$, $p < .01$]. Observers were more likely to judge the target to be male-like in the presence of the male distractor [$M = .57$; $SD = .07$] than in the presence of the female distractor [$M = .53$; $SD = .07$]. We failed to observe significant PSE shifts when the distractor regions were misaligned [$t(18) = .85$, $p = .41$] or inverted [$t(18) = .29$, $p = .78$].

Age variant. The distribution of PSEs was analysed using ANOVA with Distractor (child, adult) and Configuration (aligned, misaligned, inverted) as within-subjects factors (Figure 2b). The analysis revealed a significant main effect of Distractor [$F(1,18) = 25.68$, $p < .001$; $\eta_p^2 = .59$], a main effect of Configuration [$F(2, 36) = 3.74$, $p = .03$; $\eta_p^2 = .17$], and a significant Distractor × Configuration interaction [$F(2, 36) = 3.63$, $p = .04$; $\eta_p^2 = .17$]. Planned comparisons revealed a significant PSE shift when composites were upright and aligned [$t(18) = 4.14$, $p = .001$]. Observers were more likely to judge the target to be adult-like in the presence of the adult distractor [$M = .56$; $SD = .06$] than in the presence of the child distractor [$M = .51$; $SD = .07$]. We observed non-significant PSE shifts when the distractor regions were misaligned [$t(18) = 1.28$, $p = .22$] or inverted [$t(18) = .55$, $p = .59$].

Figure-2

Consistent with the documented normative properties of the composite face illusion (Murphy et al., 2017; Rossion, 2013), our paradigm produced marked shifts only when distractors were spatially aligned and arrangements are presented upright; little or no modulation was seen when distractor regions were misaligned or when composite arrangements were shown upside-down. This selective modulation indicates that PSE shifts observed in upright-aligned arrangements were attributable to the illusion, not response bias.

Comparing individuals' susceptibility to the age and gender composite illusions

Next, we sought to examine whether individuals' susceptibility to the age and gender composite illusions is related. One hundred typical observers ($M_{age} = 23$; $SD_{age} = 4.84$; 36 males) completed both the gender and age variants of our task. All participants were tested in person, under controlled lab conditions. Eight of these observers were replacements for participants for whom we were unable to model psychometric functions in one or more conditions (i.e., there was no systematic relationship between stimulus intensity and their pattern of responding).

We modeled eight functions for each observer. In addition to the four aligned conditions (aligned child distractor, aligned adult distractor, aligned male distractor, aligned female distractor), we also modeled four functions describing observers' categorization decisions when distractor regions were misaligned (misaligned child distractor, misaligned adult distractor, misaligned male distractor, misaligned female distractor). Each psychometric function was estimated from 140 categorization decisions (7 target levels × 20 presentations). Over the course of two testing sessions, each observer completed 1120 trials (140 trials × 2 levels of distractor × 2 alignment conditions × 2 variants). Half the participants completed the age variant first; half completed the gender variant first.

Results and discussion

Gender variant. The distribution of PSEs was analysed using ANOVA with Distractor (female, male) and Configuration (aligned, misaligned) as within-subjects factors. The analysis revealed a significant main effect of Distractor [$F(1,99) = 132.52$, $p < .001$; $\eta_p^2 = .57$], a non-significant main effect of Configuration [$F(1, 99) = 2.03$, $p = .16$; $\eta_p^2 = .02$], and a significant Distractor × Configuration interaction [$F(1, 99) = 94.47$, $p < .001$; $\eta_p^2 = .49$]. Planned comparisons revealed a significant PSE shift when composites were aligned [$t(99) = 11.75$, $p < .001$]. Observers were more likely to judge the target to be male-like in the presence of the male distractor [$M = .57$; $SD = .08$] than in the presence of the female distractor [$M = .49$; $SD = .08$]. PSE shifts were significantly reduced in the misaligned compared to the aligned condition [$t(99) = 9.01$, $p < .001$] (see Figure 3a).

Figure-3

Age variant. The distribution of PSEs was analysed using ANOVA with Distractor (child, adult) and Configuration (aligned, misaligned) as within-subjects factors. The analysis revealed a significant main effect of Distractor [$F(1,99) = 145.19$, $p < .001$; $\eta_p^2 = .60$], a non-significant main effect of Configuration [$F(1, 99) = .01$, $p = .92$; $\eta_p^2 < .001$], and a significant Distractor × Configuration interaction [$F(1, 99) = 128.11$, $p < .001$; $\eta_p^2 = .56$]. Planned comparisons revealed a significant PSE shift when composites were upright and aligned

[$t(99) = 13.43$, $p < .001$]. Observers were more likely to judge the target to be child-like in the presence of the child distractor [$M = .58$; $SD = .12$] than in the presence of the adult distractor [$M = .46$; $SD = .11$]. PSE shifts were significantly reduced in the misaligned compared to the aligned condition [$t(99) = 10.52$, $p < .001$] (see Figure 3a).

The PSE shifts seen in the aligned conditions of the gender ($M = .08$, $SD = .07$) and age ($M = .12$, $SD = .09$) tasks correlated significantly [$r = .43$, $p < .001$, $N = 100$; bootstrapped 95% CIs = .17, .62][1] (Figure 3b). In contrast, the PSE shifts seen in the misaligned conditions of the gender ($M = .02$, $SD = .03$) and age ($M = .03$, $SD = .06$) tasks did not correlate [$r = .11$, $p = .29$, $N = 100$] (Figure 3c). The strength of the correlation seen between the aligned variants of the age and gender task, was significantly greater than the correlation seen between the misaligned variants [$z = 2.62$, $p < .01$].

Figure-3

Replication in an online sample

To verify the reliability of the relationship observed, we sought to replicate the correlation in a second sample, tested online. One hundred and thirty-seven typical observers were recruited via prolific (https://prolific.ac/) and completed both the gender and age variants of our task online. Half the participants completed the age variant first; half completed the gender variant first. We were unable to model functions for 16 of these observers, leading to a final sample of 121 ($M_{age} = 28$; $SD_{age} = 9.35$; 61 males). Online versions of the age and gender tasks were programmed in Unity and made available through Unity WebGL (https://unity.com). The stimuli and experimental procedure were identical to that employed in the lab-based study.

Results and discussion

Gender variant. The distribution of PSEs was analysed using ANOVA with Distractor (female, male) and Configuration (aligned, misaligned) as within-subjects factors. The analysis revealed a significant main effect of Distractor [$F(1,120) = 113.52$, $p < .001$; $\eta_p^2 = .49$], a non-significant main effect of Configuration [$F(1, 120) = .10$, $p = .75$; $\eta_p^2 < .01$], and a significant Distractor × Configuration interaction [$F(1, 120) = 112.86$, $p < .001$; $\eta_p^2 = .49$]. Planned comparisons revealed a significant PSE shift when composites were aligned [$t(120) = 11.29$, $p < .001$]. Observers were more likely to judge the target to be male-like in the presence of the male distractor [*M* = .58; *SD* = .07] than in the presence of the female distractor [*M* = .50; *SD* = .08]. PSE shifts were significantly reduced in the misaligned compared to the aligned condition [*t*(120) = 10.62, *p* < .001] (see Figure 4a).

Figure-4

Age variant. The distribution of PSEs was analysed using ANOVA with Distractor (child, adult) and Configuration (aligned, misaligned) as within-subjects factors. The analysis revealed a significant main effect of Distractor [$F(1,120) = 160.58$, $p < .001$; $\eta_p^2 = .57$], a non-significant main effect of Configuration [$F(1, 120) = .01$, $p = .91$; $\eta_p^2 < .01$], and a significant Distractor × Configuration interaction [$F(1, 120) = 151.79$, $p < .001$; $\eta_p^2 = .56$]. Planned comparisons revealed a significant PSE shift when composites were upright and aligned [*t*(120) = 14.24, *p* < .001]. Observers were more likely to judge the target to be child-like in the presence of the child distractor [*M* = .57; *SD* =.09] than in the presence of the adult distractor [*M* = .45; *SD* = .10]. PSE shifts were significantly reduced in the misaligned compared to the aligned condition [*t*(120) = 12.32, *p* < .001] (see Figure 4a).

As in the lab-based study, the PSE shifts seen in the aligned conditions of the gender (*M* = .09, *SD* = .09) and age (*M* = .12, *SD* = .09) tasks correlated significantly [*r* = .38, *p* < .001, *N* = 121; bootstrapped 95% CIs = .22, .57] (Figure 4b). In contrast, the PSE shifts seen in the misaligned conditions of the gender (*M* = .01, *SD* = .03) and age (*M* = .02, *SD* = .05) tasks did not correlate [*r* = .03, *p* = .74, *N* = 120] (Figure 4c). The strength of the correlation seen between the aligned variants of the age and gender task, was significantly greater than the correlation seen between the misaligned variants [*z* = 2.95, *p* < .01].

General Discussion

The correlation seen between individuals' susceptibility to the age and gender biases induced by the composite face illusion, suggests that these effects are different measures of a common structural binding process. Leading theoretical models hypothesise that structural descriptions, derived early in the face processing stream, form a common basis for judgements about various facial attributes such as identity, expression, age, and gender (Bruce & Young, 1986; Haxby et al., 2000). The locus of the structural binding process responsible for the age and gender biases is therefore likely to be early in this processing stream, before the engagement of attribute specific processing. This conclusion accords with evidence that the composite face illusion modulates the N170, an EEG measure of early face encoding (Jacques & Rossion, 2009).

What kind of structural encoding might generate the composite face illusion? In our day-to-day encounters with faces, we are exposed to naturally occurring sources of feature co-variation; male eyes co-occur with male mouths; adult eyes co-occur with adult mouths, and so on. The perceptual models that observers develop in these environments will likely reflect these structural contingencies (Gray, Murphy, Marsh, & Cook, 2017). When presented with contrived composite face arrangements that violate these naturally occurring contingencies, the visual system appears to impose the natural 'whole-face' solution that best fits each contrived composite arrangement. As a result of this modeling process, observers' subjective perception of the target region is biased in the direction defined by the distractor region.

Currently, many authors use matching paradigms to measure the composite face illusion. In these tasks, observers view pairs of target regions, either presented sequentially or simultaneously, and are asked to judge whether they are identical or not (Hole, 1994). Under this approach, individual differences in illusion susceptibility are inferred from observers' ability (or inability) to discriminate target regions in the presence of distractor regions. While these approaches have proved useful in revealing the normative properties of the illusion, they tell us little about the nature of the illusory biases experienced by observers. For example, where the manifestation of the illusion causes an observer to mistakenly judge identical target regions to be different, it is not clear *how* their two subjective percepts differed. In this vein, we have previously noted that composite effects attributed to the biasing of identity perception, may instead reflect the biased perception of expression (Gray et al., 2017). The new psychophysical approach presented here offers a complementary tool with which researchers can study the extent to which a given distractor region biases perception of a target in a particular direction.

Open practices statement

None of the Experiments described were preregistered.

References

Baudouin, J. Y., & Humphreys, G. W. (2006). Configural information in gender categorisation. *Perception, 35*(4), 531-540.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433-436.

Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305-327.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 527-551.

DeGutis, J., Chatterjee, G., Mercado, R. J., & Nakayama, K. (2012). Face gender recognition in developmental prosopagnosia: evidence for holistic processing and use of configural information. *Visual Cognition, 20*(10), 1242-1253.

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

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

Gray, K. L. H., Murphy, J., Marsh, J. E., & Cook, R. (2017). Modulation of the composite face effect by unintended emotion cues. *Royal Society Open Science, 4*(4), 160867.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-233.

Hole, G. (1994). Configurational factors in the perception of unfamiliar faces. *Perception, 23*(1), 65-74.

Hole, G., & George, P. (2011). Evidence for holistic processing of facial age. *Visual Cognition, 19*(5), 585-615.

Jacques, C., & Rossion, B. (2009). The initial representation of individual faces in the right occipito-temporal cortex is holistic: electrophysiological evidence from the composite face illusion. *Journal of Vision, 9*(6), 1-16.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24*(8), 1377-1388.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4*(1), 89-94.

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin & Review, 16*(5), 778-797.

Murphy, J., Gray, K. L., & Cook, R. (2017). The composite face illusion. *Psychonomic Bulletin & Review, 24*(2), 245-261.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*(4), 437-442.

Piepers, D. W., & Robbins, R. A. (2013). A review and clarification of the terms "holistic," "configural," and "relational" in the face perception literature. *Frontiers in Psychology, 3*(559), 1-11.

Prins, N., & Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analyzing psychophysical data. *http://www.palamedestoolbox.org*.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science, 22*(4), 464-471.

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

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

Rossion, B., & Retter, T. L. (2015). Holistic face perception: Mind the gap! *Visual Cognition, 23*(3), 379-398.

Webster, M. A., & MacLeod, D. I. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B- Biological Sciences, 366*(1571), 1702-1725.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 42*(11), 1166-1178.

Footnotes:

[1]Individuals' susceptibility to the age and gender illusions also correlated significantly when analysed using Spearman's rank correlation. This was true of the lab-based [r_s = .32, p = .001] and online samples [r_s = .48, p = .001].

Figure 1

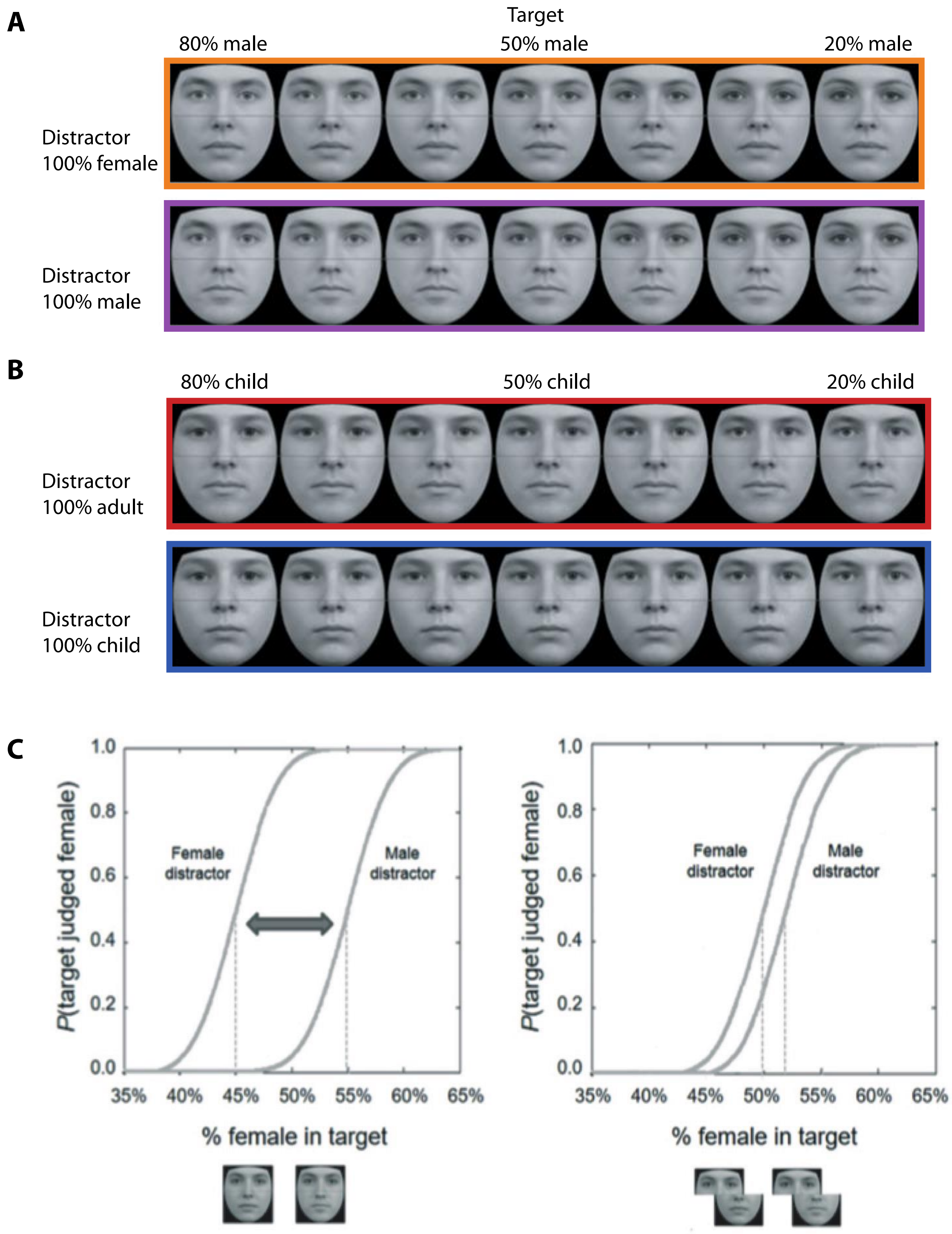

Figure 1: Stimuli used in the gender (A) and age (B) variants of the task. When shown upright and aligned, male and female distractors make target regions appear more masculine and feminine; child and adult distractors make target regions appear younger and older, respectively. Modulation of observers' perception is inferred from shifts in their psychometric functions (C).

Figure 2

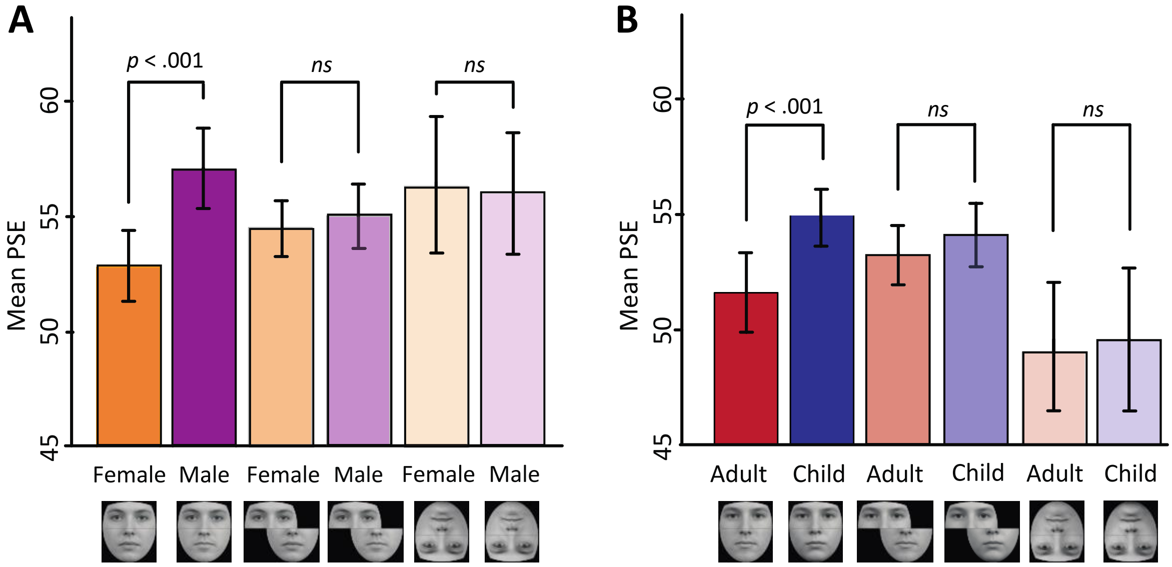

Figure 2: Results from two samples of typical observers (both N = 19) who complete the gender (A) and age (B) variants of the task. Substantial shifts were seen when aligned composite arrangements were shown upright. Little or no modulation was seen when distractor regions were misaligned or when composite arrangements were shown upside-down. Error bars denote ± SEM.

Figure 3

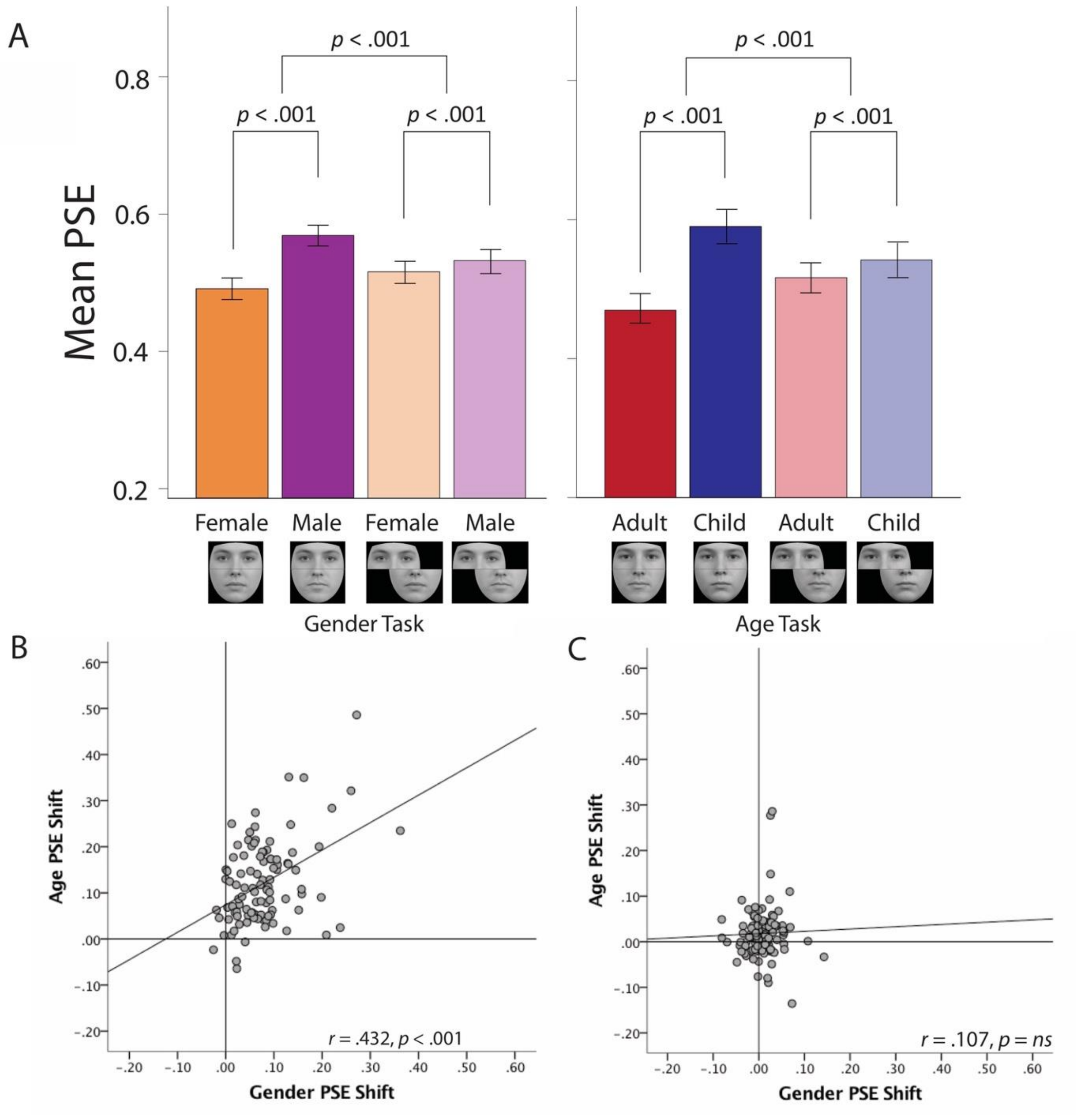

Figure 3: Results from the lab-based sample. As expected, substantial function shifts were induced when distractor regions were aligned in both versions of the task, whereas function shifts were greatly reduced when the distractor regions were misaligned (A). Observers' susceptibility to the age and gender versions of the composite face illusion correlated closely when distractors were aligned (B), but not when they were misaligned (C). Error bars = 95% CIs.

Figure 4

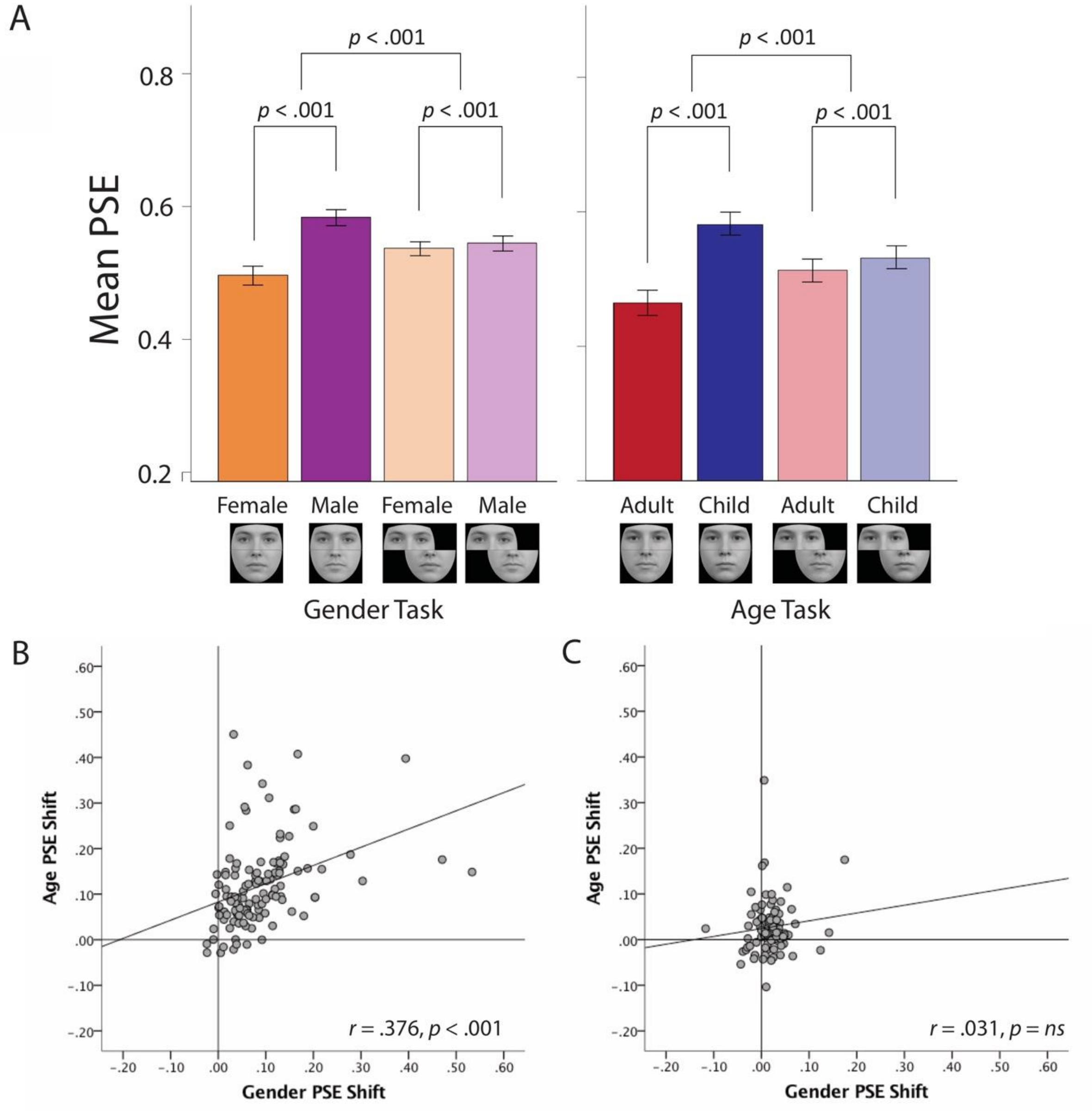

Figure 4: Results from the online sample. Again, substantial function shifts were induced when distractor regions were aligned in both versions of the task, whereas function shifts were greatly reduced when the distractor regions were misaligned (A). Observers' susceptibility to the age and gender versions of the composite face illusion correlated closely when distractors were aligned (B), but not when they were misaligned (C). Error bars = 95% CIs.

Supplementary Material

Establishing test-retest reliability

Twenty-eight typical observers (M_{age} = 23; SD_{age} = 5.89; 9 males) completed both versions of our task twice, under controlled lab conditions. Participants had normal or corrected-to-normal vision. For the gender task, we were unable to model psychometric functions for two participants' data (i.e., there was no systematic relationship between stimulus intensity and their pattern of responding), giving a final sample of 26 for this version of the task.

Testing took place over two hour-long testing sessions. In each session, we sought to model four functions for each observer: aligned child distractor, aligned adult distractor, aligned male distractor, aligned female distractor. Within each session, the tasks were blocked and completed in a counter-balanced order. Each psychometric function was estimated from 140 categorization decisions (7 target levels × 20 presentations). In total, each observer therefore completed 1120 trials (2 facial attributes × 2 levels of distractor × 140 trials per function × 2 testing sessions).

Results and discussion

Gender variant. Observers were more likely to judge the target to be male-like in the presence of the male distractor than in the presence of the female distractor in both the first session (Male distractor: [M = .58; SD =.08]; Female distractor: [M = .49; SD = .06]; [$t(25)$ = 6.59, p < .001], and the second (Male distractor: [M = .62; SD =.12]; Female distractor: [M = .53; SD =.10]; [$t(25)$ = 6.44, p < .001]). The test-retest reliability of this version of the task was high [r = .76, p < .001] (Figure S1a).

Age variant. Observers were more likely to judge the target to be child-like in the presence of the child distractor than in the presence of the adult distractor in both the first session (Child distractor: [M = .64; SD =.17]; Adult distractor: [M = .49; SD = .14]; [$t(27)$ = 5.53, p < .001]), and the second (Child distractor: [M = .65; SD =.08]; Adult distractor: [M = .51; SD =.10]; [$t(27)$ = 6.28, p < .001]). The test-retest reliability of this version of the task was also good [r = .60, p = .001] (Figure S1b).

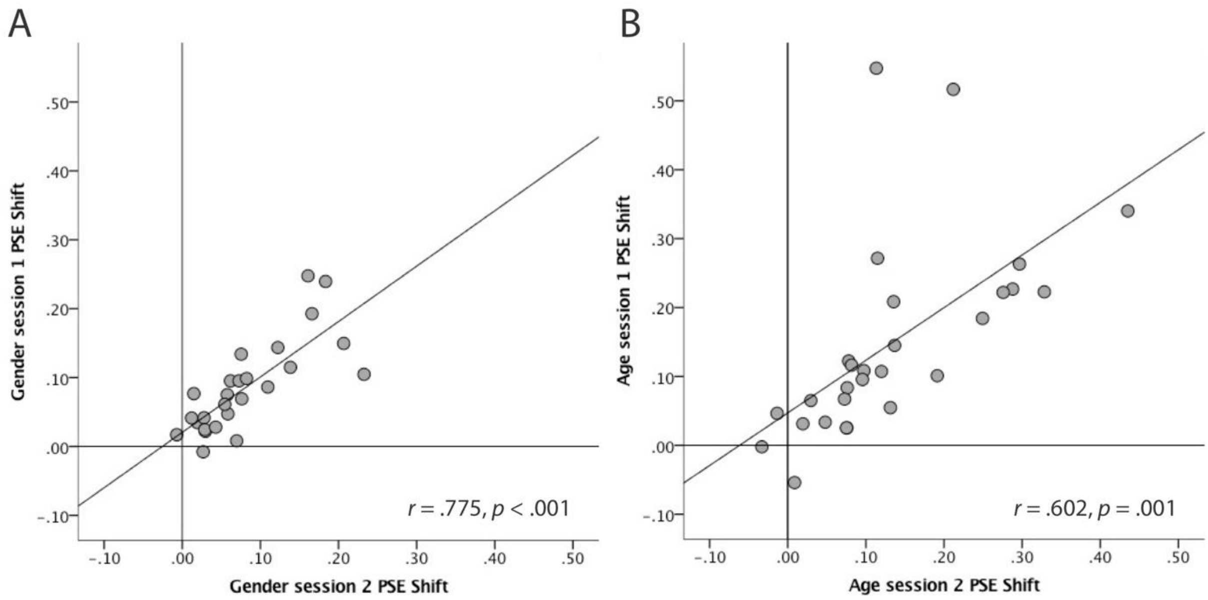

Figure S1. Scatterplots of the test-retest data (session 1 plotted against session 2) for the Gender task (A) and Age task (B).