JDDM
Journal of Dynamic Decision Making

Open Access
Original Research

Evidence for the dynamic human ability to judge another's sex from ambiguous or unfamiliar signals

Justin Gaetano
School of Psychology, University of New England, Armidale, Australia

Humans make decisions about social information efficiently, despite – or perhaps because of – the sheer scale of data available. Of these various signals, sex cues are vitally important, yet understanding whether participants perceive them as static or dynamic is unknown. The present study addressed the related question of how expertise impinges on sex judgements. Participants (80 Caucasian, 80 Asian) were asked to target female and male exemplars from a set of own- or other-race hand images. Data show: (1) that the own-race sex categorisation advantage observed previously using face stimuli can occur in relation to hands, and (2) sensitivity of Asian participants, but not Caucasian participants, is dynamic relative to how many fe/males there are in a set. Implications of these findings are discussed as further evidence that there exists a pan-stimulus sex processor, and as fresh evidence that human sex perception can change probabilistically.

Keywords: cross-cultural judgement, perceptual decision making, dynamic sex discrimination, other-race effect, own-race advantage, prior target probability

Sex is one of an exclusive set of categories by which a person may classify another person automatically. Auditory (Junger et al., 2013; Li, Logan, & Pastore, 1991), olfactory (Hacker, Brooks, & van der Zwan, 2013; Kovács et al., 2004), and of course visual (Kozlowski & Cutting, 1977; Yamaguchi, Hirukawa, & Kanazawa, 1995) information about others can lead to judgements of sex. Focussing on vision, some behavioural correlates of sex perception have been demonstrated. For instance, *male bias* – the systematic tendency to judge perceptually noisy or androgynous stimuli as male – can arise from a diverse range of visual sex cues including whole-body data (motion cues: Troje, Sadr, Geyer, & Nakayama, 2006; amorphous drawings: Brielmann, Gaetano, & Stolarova, 2015; Wenzlaff, Briken, & Dekker, 2018), and silhouette, static representations of the face (Davidenko, 2007) and hand (Gaetano, van der Zwan, Blair, & Brooks, 2014; Gaetano et al., 2016). Perceptual ambiguity is indeed a key predictor of male biased responding, however studies of child and adult participants imply that the viewer's expertise might also interact with male bias (White, Hock, Jubran, Heck & Bhatt, 2018; Wild et al., 2000; cf. Bayet et al., 2015; Tsang et al., 2018). Understanding how sex perception works not just under noisy conditions, but more generally, as a dynamic function of the perceiver's experience is the present objective.

Perceptual experience has shown to change social judgements over extended periods of time. A class of phenomena that demonstrate this point are other-race effects, which refer to participants' differential processing of stimuli that bear a less familiar resemblance race-wise (Meissner & Brigham, 2001; O'Toole et al., 1994). Other-race effects can have a powerful impact on eyewitness testimonies (Behrman & Davey, 2001; Pezdek, Blandon-Gitlin, & Moore, 2003), forensic line-up identifications (Smith, Lindsay, Pryke, & Dysart, 2001; Wells & Olson, 2001), and even visual sex discrimination (O'Toole et al., 1996 cf. Zhao & Hayward, 2010).

In this work, the nomenclature reserved for cases that demonstrate heightened *sensitivity* for own-race cues are known as *own-race advantages* (ORAs). In O'Toole and colleagues' (1996) sex categorisation study, for example, Caucasian and Asian judges categorised Caucasian and Asian faces individually as 'female' or 'male'. Overall, Caucasians and Asians were equally proficient at the task, with both groups achieving higher-than-chance sensitivity (O'Toole et al., 1996). Particularly significant in the current context, it was found that *both groups* were more astute of *own-race* faces than other-race faces. Therefore, ORA is not defined by the judge's race per se. Plausibly then, ORA could be the result of relative expertise for own-race faces that develops over many years of experience.

If sensitivity to sex cues depends on long term development per se, then such findings should not arise exclusively from face stimuli. Evidence of ORA for non-face stimuli would support this theory. Another potential source of support is the hypothetical *own-sex advantage*, by which a person would show heightened sensitivity judging people who are the same sex as them. So far, this theory has been explored within the face perception research domain exclusively. Current evidence suggests that *women* have an enhanced capacity to judge faces compared to men, particularly when the faces depict *women* (Herlitz & Lovén, 2013; Lewin & Herlitz, 2002; Rehnman & Herlitz, 2007). In other words, unlike the ORA, the own-sex advantage is apparently specific to female perceptual development. Whether this extends to non-face stimuli remains to be tested. The more immediate question is whether women develop an enhanced ability to judge social cues per se, irrespective of whether those belong to other men or women. The current study asks simply whether the general female advantage extends beyond face judgement scenarios, leaving the specific question of an own-sex advantage open to future research.

Corresponding author: Justin Gaetano, School of Psychology, Building S006, University of New England, Armidale, NSW 2351, Australia, e-mail: jgaetan2@une.edu.au

In summary, the current focus is to investigate ORA over own-sex advantage, because the former is a stronger, more prevalent class of phenomena with important ramifications (e.g. Meissner & Brigham, 2001), and it has at least been demonstrated in a sex judgement study before (i.e. O'Toole et al., 1996). However, it remains to be tested whether ORA in sex categorisation is a genuine, expertise-driven effect, or an artefact of stimulus labelling. In O'Toole et al.'s (1996) study, participants were told the race of faces they would be shown at the start of each viewing sequence, leaving open the possibility that knowledge of the race categories might systematically affect outcomes. Furthermore, facial features such as eye shape and colour obviously do differ by race, and sex signals can confound judgements of emotion from faces (e.g. Taylor, 2017). On those grounds, testing ORA using a less accessible set of cues may therefore yield different outcomes.

Finally, empirical accounts of ORA seem to illustrate how categorical judgements of others are based upon population-based norms (Jaquet, Rhodes, & Hayward, 2007; Valentine, 1991), yet it is currently unknown whether the norms extend beyond just face-based norms. The sole study of ORA in sex categorisation (O'Toole et al., 1996) is, like other cross-race perception studies, focussed on perceptions of own- and other-race *faces*. While the face might be the primary target in social development, it is certainly not the only sexually dimorphic feature that participants seem attentive to (for a review, see Gaetano et al., 2012). The visual system may in fact develop expertise with regard to hands, and inherent in those, the dynamic (albeit non-verbal) cues that hands contribute toward communication (e.g. Cook & Tanenhaus, 2009; Goldinmeadow, Wein, & Chang, 1992). Thus, the present research asks whether sex categorisation ORA *can* arise without race priming and can generalise to perceptions of hand stimuli. If so, then the case could be made that sex processing has a *common*, expertise-dependent basis.

Of course, perceptions of sex can also be influenced by higher order information (Bailey, LaFrance, & Dovidio, 2018; Freeman & Ambady, 2011). Thus far, top-down sex perception has almost exclusively been tested in relation to stereotypes. In those terms, stereotypically feminine or masculine emotion (Hess, Adams, Grammer, & Kleck, 2009); or stereotypes associated with Asian or African appearance (Johnson, Freeman, & Pauker, 2012); may facilitate respectively 'female' or 'male' judgements of otherwise androgynous *faces*. In the absence of morphological signals, it is also apparently easier to judge 'sad'- or 'angry'-primed *body motions* as female or male (Johnson, McKay, & Pollick, 2011). Such higher order face and body signal effects demonstrate the need for any comprehensive model of sex processing to take into account that sex perception is part of a dynamic person processing system (Freeman & Ambady, 2011). In light of the lower-level focus of the present study, the influence of stereotyped expressions and facial features should be minimised. Whilst evidence from the face perception domain is divergent about whether expertise really does drive the ORA (e.g. Zhao, Hayward, & Bülthoff, 2014), use of stimuli other than faces might, in future studies, at least control for face-based stereotypes.

Thus, the present study seeks to infer how experience might shape perceptions of sex beyond solely face-based processing accounts. Of course, *expertise* is a long term form of experience that has been defined and studied by way of participant race (O'Toole et al., 1996), sex (Lewin & Herlitz, 2002), and age (Wild et al., 2000); and as based on perceptual (Jaquet et al., 2007; Jaquet, Rhodes, & Hayward, 2008) and neuroimaging measures (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; McGugin, Newton, Gore, & Gauthier, 2014).

Less extensively investigated is the role of changing experience in the *short term*. In many studies of sex discrimination, the prior probability of male and female stimulus presentation are static and equal. In such studies, participants engage in binary tasks, in which they must choose between two responses – 'target sex' or 'not target sex' (e.g. Gaetano et al., 2014), or 'female' or 'male' (e.g. O'Toole et al., 1996) – on each trial. As it happens, large human populations (e.g. all citizens of a city, state, or nation) are roughly composed of 50% female and 50% male individuals (Central Intelligence Agency, 2014), so it seems reasonable to construct perceptual tasks with equal numbers of female and male stimuli. However, systematic demonstrations of male bias (e.g. Wild et al., 2000) infer that participants overestimate the frequency of males relative to females, suggesting that the bias is determined by factors other than long term experience. This calls into question the stability of sex perception performance relative to changing perception in the short term.

The question of just how susceptible sex judgements are to short term manipulations of sex ratio – or *prior target probability* (PTP) manipulations – was first addressed by Gaetano and colleagues (2016). That study revealed that PTP has no systematic bearing on sex judgement *bias* – the tendency to judge a signal (e.g. a person's face or hand) as female or male. Independent from bias outcomes, *sensitivity* indicates the accuracy of sex judgements, both in terms of true positive decisions (e.g. viewing a male and deciding 'yes, it is male'), and true negative decisions (e.g. viewing a female and deciding 'no, it is not male'). To date, it is not known whether sex judgement sensitivity is dynamic, and thus can change relative to PTP. In summary, it is possible that sensitivity to sex cues might be tuned not only to long term or developmental experience, but also to recent experience, such that sensitivity may fluctuate as a function of the sex ratio to which participants are exposed.

The current study aims to test the extent to which sex judgements depend on (i) the participant's long term familiarity with stimuli, and (ii) the relative frequency of certain sex cues in the short term. In parallel to Gaetano and colleagues' (2016) study, these experiments involve a *cross-race sample* of adult female *and* male participants and non-face stimuli, allowing the influence of long term experience on sex discrimination accuracy to be investigated.

Firstly, assuming that ORA is a phenomenon general to sex processing, it should occur for visual non-face stimuli and *between* participant groups. Specifically, it is hypothesised that when colour and texture cues are available, sensitivity will be higher for own-race participants relative to other-race participants. When sex cues are more difficult to discern within- (silhouette conditions) or between-groups (shorter presentation durations), the advantage is expected to dissipate. Secondly, assuming that the female judgement advantage is also generalisable beyond face-based stimuli, it should manifest in relation to non-face stimuli in the current study. Specifically, it is hypothesised that female participants will exhibit higher sensitivity for hand stimuli – absence of colour and textures cues or short presentation durations will negate the effect. Third and finally, assuming sex perception sensitivity is dynamic in the short term, then performance should be affected by more or less ex-

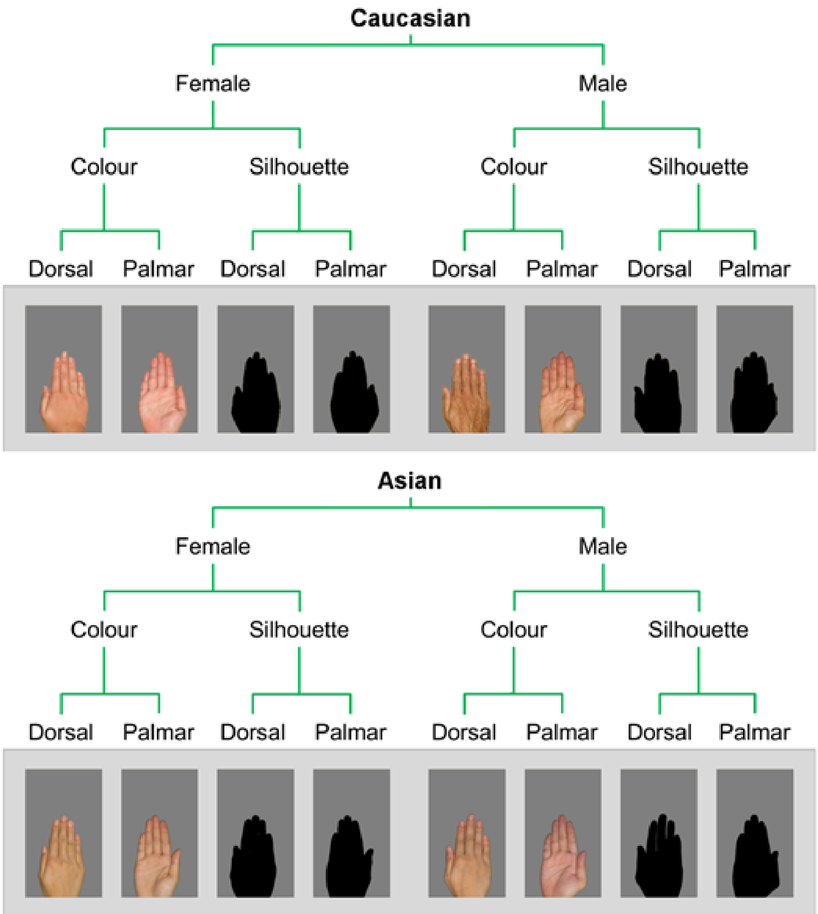

Figure 1. Caucasian (top panel) and Asian (bottom panel) stimuli used in the current study. Each image was reduced to the size of the smallest Caucasian exemplar while preserving natural aspect ratios. Within each stimulus condition 15 female and 15 male exemplars were represented.

posure to female (or male) own-race cues. Specifically, it is hypothesised that within a subset of participant groups (i.e. own-race participants), PTP will not affect sensitivity rates when participants are asked to target male or female hand stimuli. In this case, testing the null hypothesis is reasonable, in light of the negligible effect of PTP on bias outcomes observed in Gaetano and colleagues' (2016) study.

Method

Ethics Statement

All participants gave written, informed consent prior to participating in the study. All experiments were approved by the Human Research Ethics Committee, SCU (Approval numbers: ECN-11-236; ECN-12-280; ECN-13-032; ECN-14-028). In addition, all experiments conducted in Hong Kong were approved by the Human Research Ethics Committee for Non-Clinical Faculties, University of Hong Kong. This study complies with the ethical standards specified by the Declaration of Helsinki.

Participants and Materials

Throughout this study, race was operationalised from a social constructivist perspective, in line with contemporary cited studies and ethical research protocols (e.g. Brielmann, Bülthoff, & Armann, 2014; Cao, Contreras-Huerta, McFadyen, & Cunnington, 2015; Gaetano et al., 2016). Here, participants and hand stimuli models who self-identified culturally or ethnically as Australian or Hong Kongese formed the Caucasian or Asian study groups, respectively. All Caucasian participants reported being Australian citizens. Of those, a single participant (1%) reported spending one year in Hong Kong and/or China; all other Caucasians indicated living in Australia between 18 and 65 years (M = 31.29, SD = 10.28). The majority of Asian participants (71%) reported being permanent residents of Hong Kong or Chinese citizens. Caucasian participants reported living 18 to 30 years in Hong Kong and/or China (M = 21.46, SD = 2.70), all of whom reported spending no time in Australia.

Participants were 80 Caucasians (47 female) and 80 Asians (39 female), on average aged 32.49 (SD = 11.08) and 21.50 (SD = 2.72), respectively. The age difference was found to be significant ($F_{1,157} = 72.50$, $p < .001$) and although the role of age in sex judgements is not a current theoretical focus, it was explored in an unplanned manner (see Appendix).

Thirty Caucasian (15 female) and 30 Asian (15 female), size-standardised individual hands formed the basis of the stimulus set used in the present experiment. Exemplars were reduced to the size (as indexed by total pixel count) of the smallest (female) Caucasian hand (105,069px at 70.87px/cm resolution), as per the method developed by Gaetano et al. (2014), such that natural aspect ratios were preserved. They were presented centrally on a CRT monitor with 1024 × 768 px display resolution. The width and height of the grey background framing the stimulus subtended 15.74° and 25.70°, respectively, with an average distance of 57 cm between participant and monitor. Images were presented with all hue and texture information preserved ('colour' condition), and also with those cues removed ('silhouette' condition). Thus, for each experimental group, the omnibus stimulus set comprised 120 images (30 Caucasian or Asian hands [15 female, 15 male] × 2 surfaces [dorsal, palmar] × 2 conditions [colour, silhouette]). Stimulus exemplars are depicted in Figure 1.

Procedure and Analyses

An equal number of Caucasian and Asian participants were assigned randomly to one of two experiments that differed only by stimulus presentation duration (Experiment 1: 1000 ms; Experiment 2: 125 ms). Within each, participants were further equally and randomly divided into an *own-race* (i.e. Caucasian/Asian participants of Caucasian/Asian hands) or *other-race* (i.e. Caucasian/Asian participants of Asian/Caucasian hands) group. With each participant race (Caucasian or Asian) and sex (female or male) treated as an independent group, there were 16 quasi-experimental groups in total (i.e. 2 presentation duration experiments × 2 participant races × 2 stimulus races × 2 participant sexes).

Each experimental trial comprised in chronological order: a blank screen for 1000 ms, a stimulus presentation lasting 125 ms or 1000 ms, and a response screen (centred cross, +, on black background) that extinguished when either the participant made a response or 1000 ms had passed. At the response screen of each trial, the participant's task was to indicate via key press whether the image represented a target ('yes') or not ('no'). 'Targets' were defined as either female or male stimuli across separate blocks. Trials were blocked by target sex (female, male) and prior target probability (25%, 50%, 75%). Thus, each experiment consisted of 720 trials in total: 30 Caucasian or Asian individual hands (15 of each sex) × 2 hand surfaces (dorsum, palm) × 2 hue/texture conditions (colour, silhouette) × 2 target sex blocks (female, male) × 3 target probability blocks (25%, 50%, 75%). Stimuli were presented in random order within blocks, and block order and response key alternatives were counterbalanced across participants.

Participant sex discrimination ability was measured using the standardised (z-score) sensitivity measure d-prime (*d'*; Gaetano, 2017; Stanislaw & Todorov, 1999). Performances by each participant were calculated as an average on all (palmar and dorsal) trials on each condition of interest. For the sake of analytic parsimony, *between-group* prediction tests were applied only to selected conditions. Specifically, only data from blocks in which the target-to-lure ratio was equal were subjected to cross-race comparisons. Further, sensitivity was averaged across target sex conditions (female, male), because this factor was consistently found to not affect within-group sensitivity in a study that used identical stimuli (Gaetano et al., 2014).

In the subsequent *within-group* analyses pertaining to each experiment, female and male participant data were combined to form a Caucasian and an Asian group. Both independent groups included only participants of own-race hands. Within each group, performance was contrasted (i) across 25% and 75% PTP conditions and (ii) across those conditions combined and 50% PTP conditions, separately for each level of ambiguity (colour, silhouette) and target sex (female, male). Target sex conditions were also included for statistical comparison.

Predictions were tested via planned contrasts (Winer, 1962) using the PSY software package (Bird, 2004). The assumption of orthogonality was satisfied for all between- and within-group contrasts, hence no correction was made to the pairwise criterion of significance ($\alpha = .05$). For every contrast, r was calculated as the measure of effect size, expressing the magnitude of relationship between contrasted variables (Gonzalez, 2009).

Results

Experiment 1

Participants in this experiment were afforded a full second (1000 ms) to view each hand stimulus and subsequently identify it as a target (female or male) or not. In line with the general expectation that ORA is not specific to faces, sensitivity rates were first contrasted as a function of participant race (Caucasian, Asian), hand stimulus race (Caucasian, Asian) and the *interaction* between those factors. Then, female and male participant sensitivity rates were compared within Caucasian and Asian, own- and other-race groups. Those seven planned, between-group contrasts were applied independently to conditions of hue/texture (colour, silhouette), because sensitivity has consistently shown to separate between those respectively less and more ambiguous conditions (Gaetano et al., 2014). After those between-group tests, sensitivity rates were compared within each group of interest and across target sex and PTP conditions. It was expected that performance would not fluctuate as a function of those conditions.

Between-group outcomes. The *d'* statistics ($M \pm SE$) corresponding to each participant race and sex are presented in Figure 2 as a function of viewing condition. In the less ambiguous colour condition (left panel), Caucasian female participants (own-race hands: 1.47 ± 0.13; other-race hands: 1.06 ± 0.12) discriminated sex with greater average sensitivity than did Caucasian males (own-race: 0.99 ± 0.12; other-race: 0.78 ± 0.15). Asian females (own-race: 1.10 ± 0.07; other-race: 1.12 ± 0.10) also seemed more sensitive than Asian males (own-race: 0.88 ± 0.07; other-race: 0.32 ± 0.12).

Similarly in the silhouette condition (right panel), Caucasian females (own-race: 0.84 ± 0.08; other-race: 0.69 ± 0.07) performed with higher discriminability than did Caucasian males (own-race: 0.43 ± 0.12; other-race: 0.19 ± 0.12), and Asian females (own-race: 0.58 ± 0.08; other-race: 0.49 ± 0.15) outperformed Asian males (own-race: 0.53 ± 0.07; other-race: 0.12 ± 0.13).

Group performances in response to the colour hand cues are contrasted here first. Overall, though Caucasian participants were more sensitive than Asian participants ($F_{1,144} = 8.22$, $p = .005$, $r = .23$), no sensitivity difference was observed across stimulus races ($F_{1,144} = 0.06$, $p = .807$, $r = .02$). Importantly, the interaction between participant and hand race was significant ($F_{1,144} = 14.13$, p < .001, $r = .30$). Thus, performance was characterised by an ORA (see Figure 2, left panel). That is, participants of own-race hands discriminated sex with higher accuracy than did other-race participants, with one exception – Asian females did not show an ORA, as revealed via post hoc comparison (F1,17 = 0.02, $p = .879$, r < .01). Considering now just the own-race participants, Caucasian females were more sensitive on average than were Caucasian males ($F_{1,144} = 9.47$, $p = .002$, $r = .25$); no such female advantage was found for Asian participants ($F_{1,144} = 2.02$,

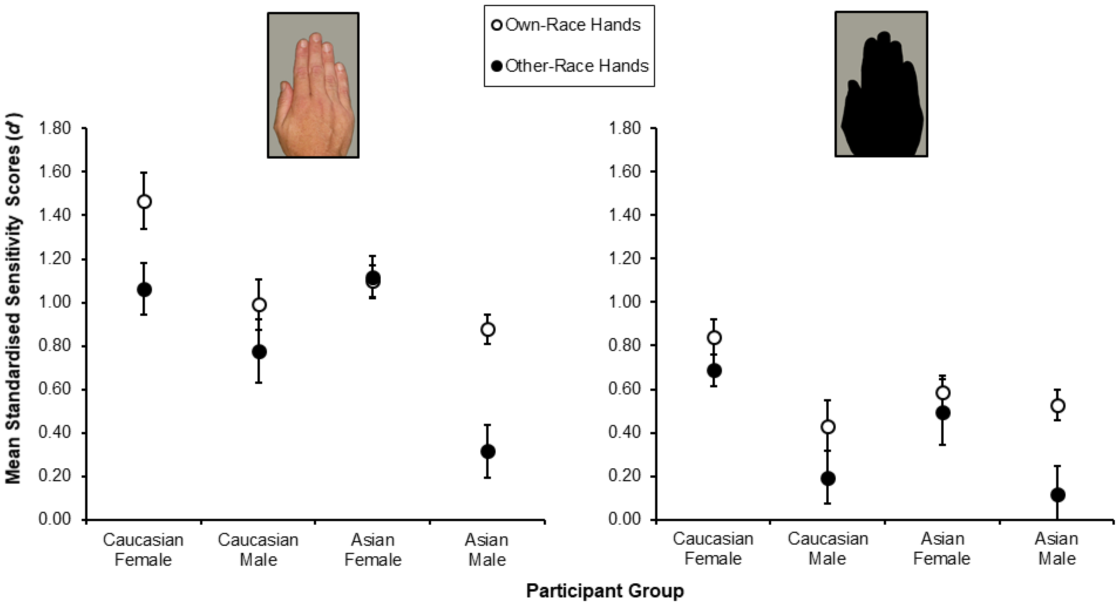

Figure 2. Group measures of sex judgement sensitivity, for 1000 ms presentations of hands shown in colour (left panel; less ambiguous condition) and in silhouette (right panel; more ambiguous condition). Sensitivity rates (*d'*) are grouped by participant race, participant sex, and stimulus familiarity (open circles: own-race hands; filled circles: other-race hands). Vertical bars represent ±1 *SE*.

p = .158, r = .12). Of the other-race participants, whilst Caucasian females seem to have had higher sex discriminability than Caucasian males, the effect did not reach significance ($F_{1,144}$ = 3.46, p = .065, r = .15). Finally, the female advantage was deemed significant among Asian participants of other-race hands ($F_{1,144}$ = 27.02, p < .001, r = .40).

Performance under conditions in which hue/texture cues were removed from hand stimuli were considered next. Overall, tests revealed that sensitivity rates varied neither by participant race ($F_{1,144}$ = 1.61, p = .207, r = .11) nor hand stimulus race ($F_{1,144}$ = 0.10, p = .748, r = .03). Nonetheless a significant interaction between those factors was found ($F_{1,144}$ = 6.92, p = .009, r = .21): that is, an ORA was surprisingly in evidence in the ambiguous, *silhouette* condition (see Figure 2, right panel).

Within participants of own-race stimuli, Caucasian females were more sensitive sex discriminators than were Caucasian males ($F_{1,144}$ = 5.79, p = .017, r = .20). By contrast, participant sex did not overall mediate Asian own-race participant performance ($F_{1,144}$ = 0.11, p = .741, r = .03). With respect to other-race participants, female sex discrimination advantage was found for both Caucasians ($F_{1,144}$ = 8.46, p = .004, r = .24) and Asians ($F_{1,144}$ = 4.94, p = .028, r = .18).

Within-group outcomes. Sensitivity statistics (M ± *SE*) for Caucasian participants judging both silhouette and colour hands appear in Figure 3 (A) and (B). Referring to the colour conditions (A), sensitivity decreased as a function of PTP (25%; 50%; 75%) when female hands were defined as the target (1.31 ± 0.11; 1.20 ± 0.11; 1.10 ± 0.12) but not when male hands were targets (1.13 ± 0.12; 1.21 ± 0.11; 1.15 ± 0.10). In the silhouette conditions (B), the trend between PTP and sensitivity was positive when female hands were targeted (0.53 ± 0.11; 0.56 ± 0.10; 0.83 ± 0.16), and negative when participants targeted male hands (0.82 ± 0.16; 0.67 ± 0.10; 0.51 ± 0.08).

Average *d'* values corresponding to Asian participants are depicted in Figure 3 (C) and (D). In the colour conditions (C), performance was lower when the target-to-lure ratio was equal (50% female targets: 1.02 ± 0.08; 50% male targets: 0.93 ± 0.08), than when it tipped in favour of lures (25% female targets: 1.30 ± 0.07; 25% male targets: 1.13 ± 0.09), or targets (75% female targets: 1.57 ± 0.18; 75% male targets: 1.14 ± 0.13). With hue/texture information not present (D), a similar trend arose when participants were asked to target females: They discriminated sex with higher sensitivity when PTP was 25% (0.99 ± 0.14) or 75% (0.51 ± 0.13) than when it was 50% (0.45 ± 0.09). Nevertheless, when asked to target silhouette males, group sensitivity seemed relatively stable across PTP conditions (25%: 0.65 ± 0.11; 50%: 0.66 ± 0.06; 75%: 0.57 ± 0.11).

Planned orthogonal contrasts revealed, first of all, that in the Caucasian group, the instruction to target either females or males had no systematic impact on performance across either the colour ($F_{1,19}$ = 0.29, p = .598, r = .12) or silhouette conditions ($F_{1,19}$ = 0.10, p = .753, r = .07). Once target sex was collapsed, and with hue/texture cues preserved, mean sensitivity was found to be uniform across PTP conditions (linear trend: $F_{1,19}$ = 1.57, p = .225, r = .28; quadratic: $F_{1,19}$ = 0.28, p = .602, r = .12). When hue/texture was removed from the hands, mean sensitivity did not differ as a linear ($F_{1,19}$ < 0.01, p = .975, r = .01) nor quadratic ($F_{1,19}$ = 0.65, p = .431, r = .18) function of PTP.

Turning now to the Asian participant group, contrasts revealed that when colour/hue cues were visible, performance unexpectedly diverged by target sex: Participants targeting female-present trials did so with greater sensitivity than when they were asked to target male hands ($F_{1,19}$ = 5.79, p = .027, r = .48). Unplanned F tests indicated that this difference was likely significant when PTP was 75% ($F_{1,19}$ = 5.66, p = .028, r = .23) and not 25% (p = .250) or 50% (p = .450), though not at the alpha level corrected for multiple comparisons (α = .017). Across colour trials, sex discrimination was just as proficient when target trials were sparse (25%) or frequent (75%), meaning that no significant linear trend was found ($F_{1,19}$ = 1.59, p = .223 r = .28). However, group performance did change as a *quadratic* function of PTP ($F_{1,19}$ = 15.09, p = .001 r = .67); performance was *worse* when targets and lures were equally probable relative to deviant (25% or 75%).

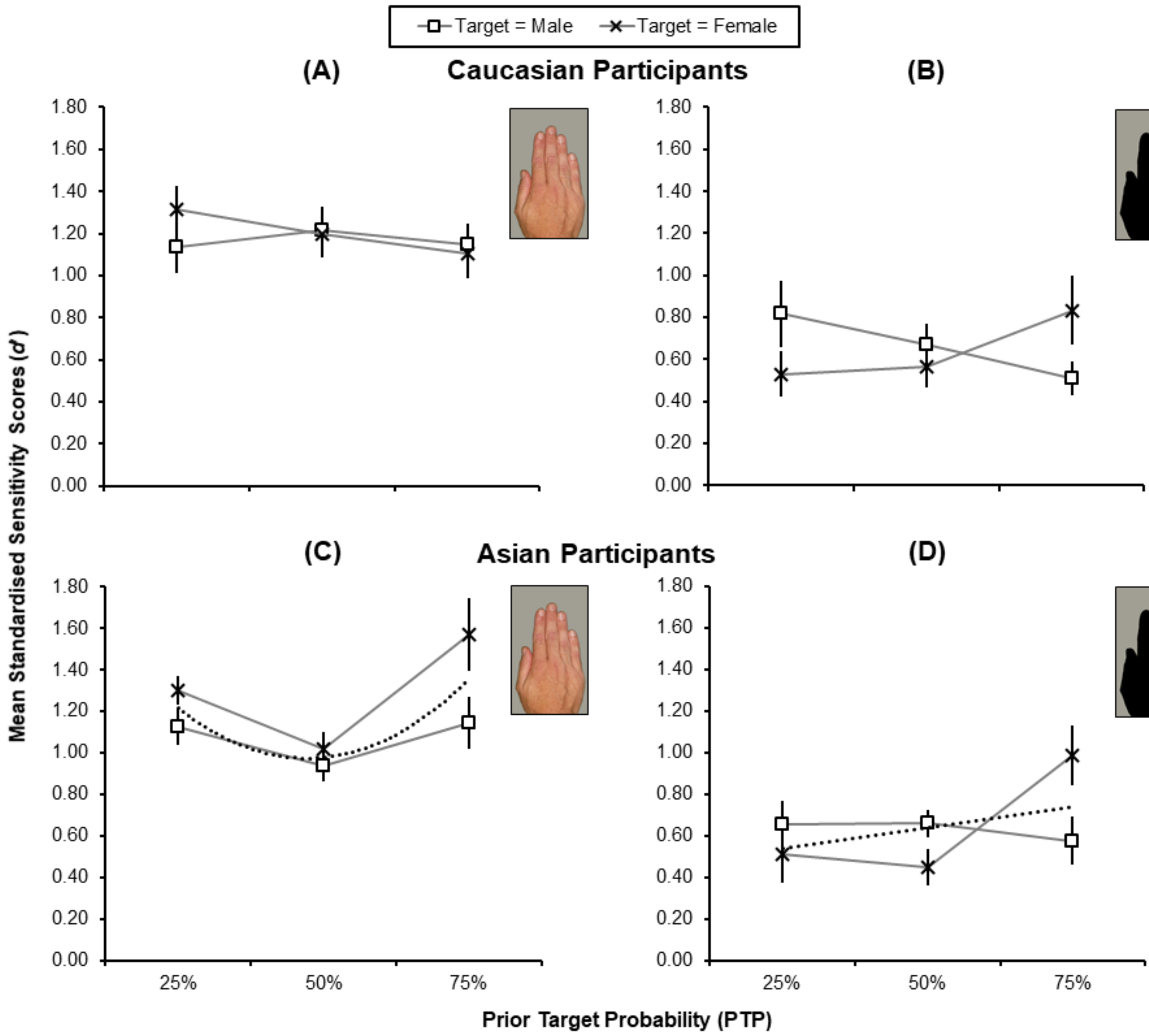

Figure 3. Within-group measures of sex judgement sensitivity for 1000 ms presentations of own-race hands. Sensitivity scores (d') corresponding to Caucasian (top panels) and Asian (bottom panels) participants are averaged over participant sex, and plotted as a function of target sex (crosses: female; squares: male), prior target probability (PTP: 25%, 50%, 75%), and whether hands were presented with (left panels) or without (right panels) hue and texture information. Broken lines represent significant polynomial trends fitted to the PTP marginal means (quadratic: panel C; linear: panel D). Vertical bars represent ± 1 *SE*.

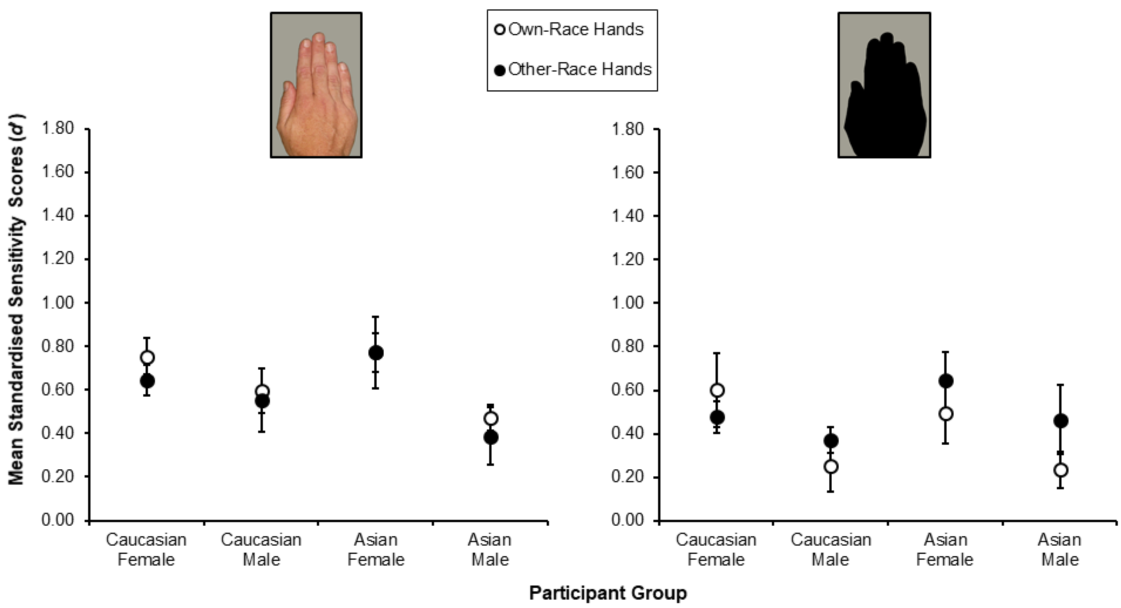

Figure 4. Group measures of sex judgement sensitivity for 125 ms colour (left panel) and silhouette (right panel) hand presentations. Sensitivity scores (d') are grouped by participant race, participant sex, and stimulus familiarity (open circles: own-race hands; filled circles: other-race hands). Vertical bars represent ± 1 *SE*.

Finally, when hue/texture cues were not available for observation, sensitivity collapsed by target sex ($F_{1,19} = 0.06$, $p = .809$ $r = .06$). Across those silhouette trials, participants discriminated sex more sensitively as PTP increased linearly ($F_{1,19} = 4.48$, $p = .048$ $r = .44$). Finally, sensitivity in the silhouette condition did not differ as a quadratic function of PTP ($F_{1,19} = 2.30$, $p = .146$ $r = .33$).

Interim discussion. The significant effects of ORA and PTP are summarised here, saving discussion of mixed or unplanned effects and trends for the *Discussion*. To summarise, sex judgements from hands each presented for 1000 ms is subject to ORA; the more experienced *own-race* participants were more sensitive to the differences between target and distractor sex of hands. This effect was not specific to one or the other race of participant – Asians and Caucasians exhibited ORA and did so independent of stimulus ambiguity.

Considering just the own-race data, one surprising effect was that Asian but not Caucasian sensitivity tracked target-to-lure stimulus ratio via quadratic and linear trends under certain conditions, which are depicted in Figure 3 (C & D; dotted lines). When the probability of fe/male stimuli deviated from the norm, Asian participants used the signal to their advantage (e.g. quadratic trend in Figure 3 [C]). In particular it can be seen that Asian participants had heightened sensitivity when asked to discriminate *common* (PTP: 75%) female targets (from male lures) relative to male targets (from female lures). What these PTP effects seem to indicate is a dynamic learning difference across cultures – a notion entertained further on in the Discussion.

To summarise Experiment 1 outcomes, ORA appears to be a true perceptual phenomenon, given that the race of hands was manipulated across groups who were not made aware of the variable (cf. O'Toole et al., 1996), and considering that the predicted outcome was produced even when sex signals were weak (as in the 'silhouette' condition). Furthermore, this is the first time that the sex judgement ORA has shown to be pan-stimulus in nature – it arose here without the assistance of familiar facial features or their associated stereotypes, and so appears to genuinely be a result of dynamic sex processing mechanisms.

Experiment 2

In Experiment 2, the parameters of the sensitivity effects noted above are probed further. Specifically, the stimulus inspection time is here limited to an eighth (i.e. 125 ms) of that used in Experiment 1, to investigate the extent to which the sex categorisation ORA is dependent on processing time. If ORA is weaker at 125 ms, it would suggest that the advantage incurs a time-expense associated with comparing current sensory data with a stored norm of sex signals (Valentine & Endo, 1992).

Between-group outcomes. The sensitivity statistics obtained from 125 ms hand participants are shown in Figure 4. When hands were presented with hue/texture intact (left panel), Caucasian females (own-race: 0.75 ± 0.08; other-race: 0.64 ± 0.07) discriminated sex with heightened sensitivity group scores compared with Caucasian males (own-race: 0.60 ± 0.10; other-race: 0.55 ± 0.15). Asian females (own-race: 0.77 ± 0.09; other-race: 0.77 ± 0.17) similarly had higher sensitivity rates than did Asian males (own-race: 0.47 ± 0.06; other-race: 0.39 ± 0.13).

When hue/texture cues were eliminated from the hands (right panel), the same trend emerged: Caucasian female participants (own-race: 0.60 ± 0.17; other-race: 0.48 ± 0.07) had higher group sensitivity rates than did Caucasian males (own-race: 0.25 ± 0.12; other-race: 0.37 ± 0.06); likewise Asian females (own-race: 0.50 ± 0.14; other-race: 0.65 ± 0.13) on average performed better than Asian males (own-race: 0.23 ± 0.08; other-race: 0.46 ± 0.16). Overall, the standard range of sensitivity means is narrow ($d'_{\max}$ - $d'_{\min} = 0.54$) compared to the range across 1000 ms groups (1.35; see *Experiment 1: Between-group outcomes*).

Tests applied to decisions made in the *colour* conditions revealed, for the most part, null effects. Race of participant ($F_{1,144} = 0.19$, $p = .660$, $r = .04$) and hand familiarity ($F_{1,144} = 0.04$, $p = .840$, $r = .02$) did not systematically impact overall group performance, and the non-significant interaction between those factors ($F_{1,144} = 0.54$, $p = .462$, $r = .06$) provides evidence against the existence of an ORA at short durations (i.e. 125 ms). Of the own-race groups, for both Caucasian ($F_{1,144} = 0.97$, $p = .327$, $r = .08$) and Asian ($F_{1,144} = 3.79$, $p = .054$, $r = .16$) participants, mean sensitivity did *not* diverge by participant sex. Referring to other-race stimulus groups, Caucasian participant sensitivity was not on average different between females and males ($F_{1,144} = 0.27$, $p = .606$, $r = .04$). However, a difference was found among other-race, Asian participants ($F_{1,144} = 6.24$, $p = .014$, $r = .20$): Within that group, females judged sex with higher sensitivity than did males.

Contrasts of performance under *silhouette* conditions also resulted in a lack of systematic differences. Sensitivity varied neither by participant race ($F_{1,144} = 0.15$, $p = .704$, $r = .03$) nor by hand stimulus race ($F_{1,144} = 1.16$, $p = .284$, $r = .09$), and no interaction between those factors was found ($F_{1,144} = 1.14$, $p = .288$, $r = .09$).

Considering performance relating to familiar (own-race) hands, Caucasian female participants were slightly advantaged compared to Caucasian males, though not significantly so ($F_{1,144} = 3.85$, $p = .052$, $r = .16$). Similarly, Asian own-race performance did not diverge by participant sex ($F_{1,144} = 2.39$, $p = .124$, $r = .13$). Finally, female advantage was not detected in either Caucasian ($F_{1,144} = 0.29$, $p = .594$, $r = .04$) or Asian ($F_{1,144} = 1.17$, $p = .281$, $r = .09$) participants of unfamiliar (other-race) hands. As mentioned, the difference between the largest and smallest 125 ms sensitivity mean spans about half a standard deviation (0.54), thus the range in which a true effect can be detected is small.

Within-group outcomes. The d' statistics for *Caucasian* participants of stimuli each presented for 125 ms are represented in Figure 5 (A) and (B). When hands were judged in colour (A), different trends emerged depending on target sex. When participants were asked to target female hands, sensitivity peaked in the condition of equal target versus lure trials (50%; 0.80 ± 0.08), and dropped when fewer (25%) or more (75%) female targets were present (respectively: 0.69 ± 0.09; 0.58 ± 0.10). The opposite trend was in evidence when male hands were targets: Participants performed worse given a balanced PTP (50%; 0.60 ± 0.09) than when given a diminished (25%) or augmented (75%) one (respectively: 0.73 ± 0.12; 0.77 ± 0.10). When hue/texture cues were omitted from the hands (B), there was a slight, positive trend between PTP (25%; 50%; 75%) and sensitivity for deciding hands were female (0.48 ± 0.11; 0.54 ± 0.15; 0.55 ± 0.08). A similarly weak yet opposite trend emerged when male hands were being targeted (0.55 ± 0.13; 0.42 ± 0.12; 0.41 ± 0.13).

The mean d' values for *Asian* participants of briefly presented (125 ms) stimuli are represented in Figure 5 (C) and (D). When hue/texture was preserved (C), judgement sensitivity was lowest when targets and lures were presented in equal number (female targets: 0.62 ± 0.06; male targets: 0.63 ± 0.09), and improved as PTP either decreased (female targets: 0.81 ± 0.15; male targets: 0.72 ± 0.13), or increased (female targets: 0.84 ± 0.15; male targets: 0.87 ± 0.15). In the absence of hue/texture (D), sensitivity diminished as PTP grew (25%; 50%; 75%) when target sex was male (0.55 ± 0.10; 0.43 ± 0.14; 0.30 ± 0.14) but not female (0.30 ± 0.14; 0.30 ± 0.08; 0.70 ± 0.19).

A set of orthogonal contrasts tested the within-group predictions described above, first for Caucasian then Asian participants. First, as expected, Caucasian sensitivity rates did not differ as a function of target sex. This was the case both when hue and texture was present ($F_{1,19} = 0.02$, $p = .892$, $r = .03$) and when absent ($F_{1,19} = 0.90$, $p = .356$, $r = .21$). In the colour condition, Caucasian participants' sensitivity did not overall differ as a linear function of PTP ($F_{1,19} = 0.15$, $p = .707$, $r = .09$); and group performance did not change in a quadratic direction either ($F_{1,19} = 0.01$, $p = .944$, $r = .02$). Finally, in the silhouette condition, group sensitivity did not differentiate across PTP blocks (linear: $F_{1,19} = 0.15$, $p = .705$, $r = .09$; quadratic: $F_{1,19} = 0.03$, $p = .874$, $r = .04$).

Orthogonal contrasts within the Asian participant data revealed that sex discrimination performance did not diverge by target sex, regardless of whether hue and texture cues were shown ($F_{1,19} = 0.02$, $p = .879$, $r = .04$) or not ($F_{1,19} = 0.01$, $p = .914$, $r = .03$). When hue/texture cues were visible to Asian participants, their average judgement sensitivity did not change linearly by PTP ($F_{1,19} = 0.43$ $p = .521$, $r = .15$). However, sensitivity did change in a *quadratic* fashion such that judgement performance was worse in the condition with 50% targets and lures ($F_{1,19} = 6.21$, $p = .022$, $r = .50$). Finally, when hue/texture cues were absent, no linear ($F_{1,19} = 0.36$, $p = .554$, $r = .14$) or ($F_{1,19} = 0.88$, $p = .359$, $r = .21$) quadratic trend was detected in the sensitivity data across PTP blocks.

Interim discussion. Experiment 2 outcomes showed that limiting hand presentations to just 125 ms rendered the ORA non-significant, especially so when hue/texture properties were absent among stimuli. Therefore, in conjunction with the positive result found when hands were presented for 1000 ms (Experiment 1), it seems that the advantage afforded by expertise with own-race cues involves a processing time cost. This could be explained in terms of a dynamic sex cue space model. Sensory evidence – in this case, a hand shape – is matched against stored fe/male norms that are tuned by ever-accumulating experience; if the evidence is too fleeting (e.g. 125 ms), it is not able to be processed as a familiar exemplar, and hence, does not lead to any behavioural advantage.

Similarly, the female participant advantage was in most cases nullified by the reduced presentation duration, though in every group, the female participant mean superseded that of male participants. Accounts of female advantage might seem unsuited to the dynamic sex cue processor model, as it is reasonable to assume that adult participants have approximately as much experience with either female or male adult cues. However, Loven and colleagues (2012) have suggested that the encoding of stimuli via experience-tuned perceptual norms is further mediated by motivation: female participants essentially enhance their social categorisation acuity by paying more attention to social (female and male) cues.

Finally, Experiment 2 found additional evidence supporting the theory, that human sex judgement abilities are dynamic in relation to short term PTP changes. Again, the qualifying factor is, mysteriously, participant race: Asians but not Caucasians showed higher sensitivity when target hands were uncommon (25% PTP) or common (75% PTP), relative to equiprobable (50% PTP). This U-shape shift in decision-making was weaker in Experiment 2, because of the quick stimulus exposure time (125 ms); group sensitivity traced a U-shape only when Asian participants were assisted by the presence of texture and colour signals.

Discussion

The broad aim of this study was to explore the extent to which the ability to judge sex is shaped by relative, changing experience with certain signals. The specific objective was to determine whether sex judgements from hands – like faces – are influenced by racial familiarity (long term experience) as well as PTP (short term experience). Participants were asked to report whether or not each presentation of a hand depicted a target sex, with the primary prediction that an ORA would be observed. That prediction was mostly supported: Given sufficient viewing time (1000 ms), Caucasian and Asian participants were more sensitive targeting sex from own-race hands than they were performing the same task with respect to hands of the other race. Furthermore, it was predicted that the variable probability of target sex – which here represents change to real-time experience and not a priori knowledge – would not alter sex discriminability rates within groups. Here, some unexpected trends were detected. Intriguingly, those were race-specific: Sensitivity changed for Asians but not Caucasians as a function of PTP manipulations. Before those two key outcomes are discussed, two periphery findings should at least be mentioned.

Firstly, an overall female judge advantage was found – the trend was apparent in almost every condition and for Caucasians and Asians alike, and in many cases the trend was significant. Gaetano et al. (2014) had speculated that sex judgements would differ between female and male participants, but lacked statistical power to definitively test that possibility. Previous studies have reported systematic differences between female and male cortical structure (Wang, Shen, Tang, Zang, & Hu, 2012) and functions (Canli, Desmond, Zhao, & Gabrieli, 2002). In terms of perceptual dimorphism, female and male participants have been found to inspect different areas of the face when categorising sex (Armann & Bülthoff, 2009), and females appear to have superior memory for faces, especially if those are female (Herlitz & Lovén, 2013; Lewin & Herlitz, 2002; Rehnman & Herlitz, 2007). The superior perceptual performance of females in the present study is the first using hands as stimuli (cf. Schouten, Troje, Brooks, van der Zwan, & Verfaillie, 2010), so it would be of theoretical interest to study which region of the hands females and males are focussing on.

Secondly, unlike the participant-mediated effects of ORA and female advantage, participant age did not seem to affect sensitivity measures (see Appendix). However, this could be an artefact of each median split reducing the power of analyses. Based on the lack of support from the sex judgement literature and the non-definitive findings here, a systematic role for age in these effects seems unlikely. That said, a future study could enlist separate

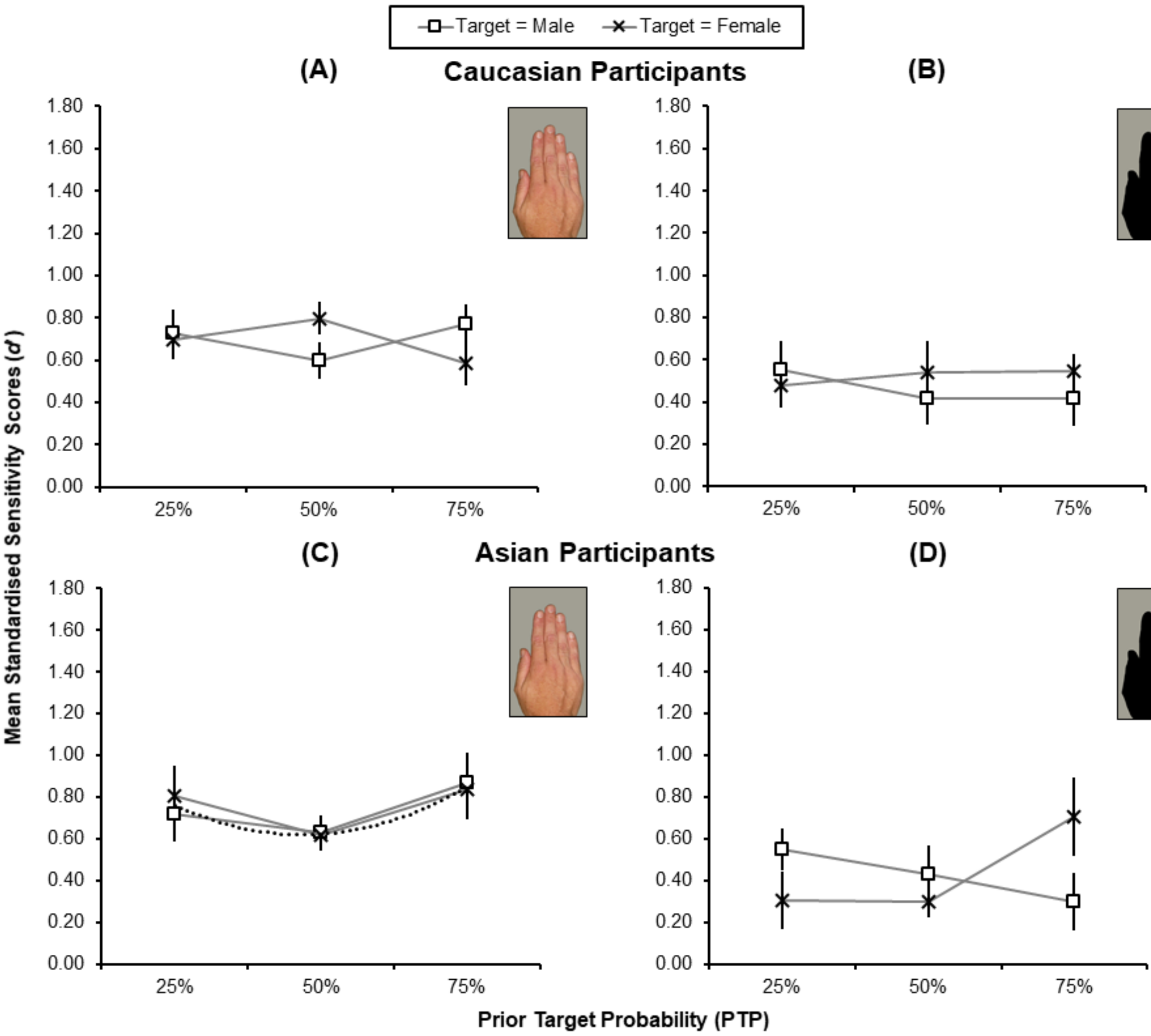

Figure 5. Sex judgement sensitivity measures corresponding to 125 ms presentations of own-race hands. Caucasian (top panels) and Asian (bottom panels) participants' sensitivity scores (*d'*) are averaged over participant sex, and plotted as a function of prior target probability (PTP: 25%, 50%, 75%), target sex (crosses: female; squares: male), and whether hands were presented with (left panels) or without (right panels) hue and texture information. The broken line (panel C) represents a significant quadratic trend fitted to the PTP marginal means. Vertical bars represent ± 1 *SE*.

participant age groups to systematically explore the relationships.

The own-race advantage in sex classification

Whilst the ORA has been demonstrated under a range of different conditions, the present data represent the first demonstration of an ORA with respect to judging the sex of human hands. Indeed, the effect was detected in response to cues presented for 1000 ms, but no advantage was apparent given a much shorter processing time (125 ms). By contrast, O'Toole's (1996) *face-based* study evoked the effect with an exposure time of just 75ms. There though, participants were primed with the information that the aim of the task was to measure accuracy in response to own- versus other-race faces. So, whilst participants in that study were aware racial congruency was being manipulated across blocks of trials, participants in the present study viewed *either* own-race or other-race stimuli, and were not informed that stimulus race was a variable. The difference in participant expectation between these studies may explain why ORA occurred at a brief presentation duration in the previous (O'Toole et al., 1996) but not the present set of observations.

A further explanation is that participants have more expertise viewing faces relative to hands per se, and so race-selectivity in sex judgement is nullified given a brief exposure time of the latter. Certainly, this idea is supported by perceptual data: Caucasian and Asian *face participants* in O'Toole's (1996) study achieved sex classification sensitivity rates of $d' > 2.00$, whereas *hand participants* afforded almost double the exposure time (125 ms as opposed to 75ms) averaged only $d' = 0.62$.

The sex classification ORA may have an upper bound as well. In one study, Chinese students were afforded *unrestricted* time to categorise each Chinese and Caucasian face by sex (Zhao & Hayward, 2010). On average, overall sex discriminability was markedly high for *intact* faces ($d' \geq 3.00$), yet participants did not exhibit an advantage for the own-race subset (Zhao & Hayward, 2010). Nevertheless, under certain degraded signal conditions, the match between participant and face race did benefit sex categorisation (Zhao & Hayward, 2010; cf. Hayward, Rhodes, & Schwaninger, 2008). In sum, despite the methodological differences between the current study, O'Toole et al.'s (1996) study, and Zhao and Hayward's (2010), together they support the notion that deciding someone is female or male is a matter of accumulating experience.

Surprisingly, the ORA does not explain the judgements of one subgroup in the current study: Asian females. To the author's knowledge, studies of ORA do not typically compare measures across participant sexes or races. One study has investigated Caucasian females' proneness to ORA when judging faces, but does not comment on whether findings would generalise to Asian females (Wallis, Lipp, & Vanman, 2012). Thus, explanations of the current finding are speculative without further evidence, that Asian males but not Asian females possess an ORA for sex cues. If this finding cannot be replicated, it could reflect an enculturated strategy specific to Hong Kong (i.e. where the current Asian participants were recruited). For example, there may be more selection pressure for Hong Kongese males to identify in-group versus out-group membership, as males are the minority in the Hong Kongese population (CIA, 2017).

On the basis of these findings, it is plausible that there exist mechanisms which process sensory input from face- (e.g. FFA; Kanwisher, McDermott, & Chun, 1997) or hand-selective (e.g. left lateral occipitotemporal cortex; Bracci, Ietswaart, Peelen, & Cavina-Pratesi, 2010) regions via a dynamic, sex signal space. Such a space has already been modelled for face perceptions (Campanella, Chrysochoos, & Bruyer, 2001; Johnston, Kanazawa, Kato, & Oda, 1997; Valentine & Endo, 1992). According to the norm-based model of face recognition (e.g. Valentine & Endo, 1992), faces are encoded in a hypothetical space as points located around a population norm – those points are more densely clustered surrounding other-race prototypes than own-race prototypes, facilitating judgements about own-race exemplars on various dimensions (e.g. sex, age) in the space. The sex judgement ORA from face and non-face, male and female signals suggests there could be a pan-stimulus sex processor. If so, such a framework could be used to test predictions about how sex processing functions as part of a wider person judgement matrix (e.g. Freeman & Ambady, 2011).

Indeed, the ORA as described here and previously in O'Toole et al.'s (1996) work is a theoretical element of the wider, experience-dependent nature of how humans judge those around them. For instance, emerging research has shown that differential experience with racial groups can affect neural correlates of perceiving pain in other persons (Contreras-Huerta, Baker, Reynolds, Batalha, & Cunnington, 2013; Contreras-Huerta, Hielscher, Sherwell, Rens, & Cunnington, 2014), such that the neural bias associated with other-race faces is reduced as the level of everyday contact is increased (Cao et al., 2015). In contrast to this support for the contact hypothesis, other-race effects can dissipate if participants are *told* that they share intrinsic characteristics with other-race individuals (Zhou, Pu, Young, & Tse, 2015). More broadly, participants seem better able to process biological stimuli that are more familiar to them not just by race (Meissner & Brigham, 2001) but also age (Rhodes & Anastasi, 2011) and species (Dahl, Chen, & Rasch, 2014; Sigala, Logothetis, & Rainer, 2011). In summary, these effects demonstrate that dynamic, pan-stimulus models of person perception – and sex perception in particular – are high in explanatory power for incorporating past judgements as a factor.

PTP effects within groups

Present evidence suggests that PTP did mediate sensitivity in some unexpected ways. Firstly, when PTP differed from the 50% level expected in binary decision tasks, sensitivity also changed for Asians – that is, it increased if target trials were fewer (25%) or many (75%) – but only when the hues and textures of hands were visible. Caucasians on the other hand showed no such sensitivity shift in response to colour hands.

Higher PTP equates to more trials in which the participant can make a 'hit', and less trials in which a 'false alarm' can be made. Yet paradoxically, when cues had hue/texture preserved, Asians did better in *both* the 25% and 75% conditions relative to the 50% condition, revealing a U-shape sensitivity pattern. That quadratic trend was significant irrespective of viewing time (125 ms or 1000 ms), but was stronger when participants were allowed a complete 1000 ms per stimulus view. In sum, this result provides tentative support for the novel suggestion that Asians adopt a different strategy when performing the task: Compared to Caucasians, they discriminated sex 'online' or adaptively, by matching the dynamic proportion of targets to lures – and despite no explicit instruction that the proportion was shifting across experiment blocks.

It is uncertain why Asian participants might have behaved in this manner. Speculating on the causes, cultural variation in problem solving strategies, or even the significantly unbalanced sex ratio in the population of Hong Kong may play a role. On the latter, the Hong Kongese population consists of only 87 males for every 100 females (CIA, 2017). This female population bias has increased over time and is projected to continue increasing. In Australia, the ratio of 101 males per 100 females is statistically balanced, and matches closely the global statistic (i.e. 102:100; CIA, 2017). In the current study, Hong Kongese participants may have learned to judge sex with greater care when the PTP was unbalanced (25% or 75%), because a balanced PTP (50%) does not agree with everyday experience of the true sex ratio in Hong Kong.

The observed effects of PTP on Asian hand judges may also be explained by differing enculturated attentional strategies between participant races. For instance, convergent evidence from eye-tracking studies indicate that Asians tend to scan facial images in a more holistic manner than do Caucasians (e.g. Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Brielmann et al., 2014). It is untested though indeed possible that race-based tracking differences exist for hand stimuli as well, and if so, could explain the Asian PTP effect in the present study. It is also possible that in general, Asians are more likely than Caucasians to distinguish *people* by sex holistically. Asians may have an advantage exploiting signals from hands and other areas as well as the face, and if so, then they may show heightened sensitivity to a dynamic PTP. Although such hypotheses are beyond the scope of the present study, they are at least consistent with the flattening of sensitivity patterns observed when colour and texture cues were removed.

A parallel study has been conducted to investigate whether this quadratic trend has any association with sex classification *bias* (Gaetano et al., 2016). This is a question of legitimate theoretical concern, because the chosen index of sensitivity (d') works on the premise that target and lure distributions are normal-shaped and have equal variances; violations of either condition will permit d' to vary with response bias (c; Stanislaw & Todorov, 1999). Paradoxically, such violations are more likely to occur when sex signals are difficult to discern, which in turn is also when male bias is more likely to arise (e.g. Gaetano et al., 2014). Nevertheless, present data reveals a completely unexpected effect of PTP. In contrast to bias outcomes, which were found to be *static* relative to PTP changes (Gaetano et al., 2016), sensitivity outcomes in the present study changed non-linearly as a function of participant race.

Breaking the Asian-specific PTP phenomena down further, performance was compared across the 'uncommon' and 'common' conditions, ignoring those conditions in which the target-to-lure ratio was equal. When presentations contained hue/texture, sensitivity was generally no different between 25% and 75% PTP blocks. When presented with silhouette hands, the 1000 ms Asian participant group discriminated sex with greater acuity when PTP was high compared to low. That said, it is difficult to determine whether or not this is a true effect. For instance, the associated significance value of .048 is close to the threshold of .050; the effect can explain only 19% of that particular group's sensitivity variance, which is small in comparison to the 44% explained by the same group's U-shaped effect mentioned above. Certainly, this positive linear effect was not demonstrated within any of the other groups or ambiguity conditions. So in total, Asian participants are better at discriminating sex from *unambiguous* hands when the probability of fe/male targets is deviant (i.e. 25% or 75%).

Finally, overall group performances did not vary by target sex, with just one exception: Asian participants of stimuli presented for 1000 ms were on average more sensitive targeting females than targeting males when hue/texture cues were preserved. The relatively female-saturated population of Hong Kong that these participants were exposed to could explain this result. Nevertheless, the effect seemed to manifest only when there were fewer targets (25%) per block, and only if a liberal significance value was chosen (see Figure 3 (C)). In sum then, as expected, sensitivity rates are uniform irrespective of whether the participant is looking for females or males. On the contrary, it has been demonstrated consistently that target sex *does* affect response criteria (Gaetano et al., 2014; Gaetano et al., 2016). The bulk of the evidence in the present study suggest that any such changes in decision bias occur independent of decision sensitivity.

Conclusion

In summary, consistent with a general theory of sex judgement (Freeman & Ambady, 2011), the present data provide empirical support for the notions that sex categorisations: (i) partially dependent on the participant's long term perceptual expertise with certain groups of dimorphic cue, and (ii) may fluctuate as a function of short term probabilistic changes across cues, at least for certain groups of participants. With respect to (i), it has been shown that ORA is a pan-stimulus phenomenon that affects not just face judgements, but more generally sex judgements. Of particular note, this phenomenon does not require participants to be aware of stimulus race manipulations. Regarding (ii), the present study has revealed some interesting patterns of PTP-mediated sex judgement, which apparently arise for Asians but not Caucasians. Specifically, when sex cues are relatively intact, Asians adaptively change their decision-making acuity in a curvilinear fashion as PTP increases; when cues are degraded, they decrease their sensitivity linearly as PTP increases. Caucasians, despite being afforded the same variable likelihoods of making a correct decision, overall did not change their sensitivity. These findings extend on the notion of a sex processing model analogous to the face-space model: Human sex judgement depends not only on how different the female and male signals are in this space, but also on the participant's dynamic experience with signals in the long and short term.

Acknowledgements: The author dedicates this work to the memory of Dr Kevin Minotti. The author thanks former supervisors Anna Brooks and Rick van der Zwan for their foundational supervisory support and guidance. Thanks also to William G. Hayward and Matthew Oxner for their assistance with data collection. Finally, thank you to Ross Cunnington, for sharing his insights on ORA at the 12th International Conference on Cognitive Neuroscience, Brisbane, Australia.

Declaration of conflicting interests: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be constructed as a potential conflict of interest.

Handling editor: Andreas Fischer

Citation: Gaetano, J. (2019). Evidence for the dynamic human ability to judge another's sex from ambiguous or unfamiliar signals. *Journal of Dynamic Decision Making, 5*, 3. doi:10.11588/jddm.2019.1.61118

Received: 04 Apr 2019
Accepted: 28 Jun 2019
Published: 09 Jul 2019

References

Armann, R., & Bülthoff, I. (2009). Gaze behavior in face comparison: The roles of sex, task, and symmetry. *Attention, Perception and Psychophysics*, *71*(5), 1107–1126. doi:10.3758/app.71.5.1107

Bailey, A. H., LaFrance, M., & Dovidio, J. F. (2018). Is man the measure of all things? A social cognitive account of androcentrism. *Personality & Social Psychology Review*. doi:10.1177/1088868318782848

Bayet, L., Pascalis, O., Quinn, P. C., Lee, K., Gentaz, E., & Tanaka, J. W. (2015). Angry facial expressions bias gender categorization in children and adults: behavioral and computational evidence. *Frontiers in Psychology*, *6*. doi:10.3389/fpsyg.2015.00346

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior*, *25*(5), 475–491. doi:10.1023/a:1012840831846

Bird, K. D. (2004). *Analysis of variance via confidence intervals*. London: Sage Publishing. doi:10.4135/9781849208598

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE*, *3*(8), e3022. doi:10.1371/journal.pone.0003022

Bracci, S., Ietswaart, M., Peelen, M. V., & Cavina-Pratesi, C. (2010). Dissociable neural responses to hands and non-hand body parts in human left extrastriate visual cortex. *Journal of Neurophysiology*, *103*(6), 3389–3397. doi:10.1152/jn.00215.2010

Brielmann, A. A., Bülthoff, I., & Armann, R. (2014). Looking at faces from different angles: Europeans fixate different features in Asian and Caucasian faces. *Vision Research*, *100*, 105–112. doi:10.1016/j.visres.2014.04.011

Brielmann, A. A., Gaetano, J. M., & Stolarova, M. (2015). Man, you might look like a woman - if a child is next to you. *Advances in Cognitive Psychology*, *11*(3), 84–96. doi:10.5709/acp-0174-y

Campanella, S., Chrysochoos, A., & Bruyer, R. (2001). Categorical perception of facial gender information: Behavioural evidence and the face-space metaphor. *Visual Cognition*, *8*(2), 237–262. doi:10.1080/13506280042000072

Canli, T., Desmond, J. E., Zhao, Z., & Gabrieli, J. D. E. (2002). Sex differences in the neural basis of emotional memories. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(16), 10789–10794. doi:10.1073/pnas.162356599

Cao, Y., Contreras-Huerta, L. S., McFadyen, J., & Cunnington, R. (2015). Racial bias in neural response to others' pain is reduced with other-race contact. *Cortex*, *70*, 68–78. doi:10.1016/j.cortex.2015.02.010

Central Intelligence Agency [CIA]. (2017). *The world fact book: Sex ratio*. Retrieved from: https://www.cia.gov/library/publications/the-world-factbook/geos/xx.html

Contreras-Huerta, L. S., Baker, K. S., Reynolds, K. J., Batalha, L., & Cunnington, R. (2013). Racial bias in neural empathic responses to pain. *PLoS ONE*, *8*(12). doi:10.1371/journal.pone.0084001

Contreras-Huerta, L. S., Hielscher, E., Sherwell, C. S., Rens, N., & Cunnington, R. (2014). Intergroup relationships do not reduce racial bias in empathic neural responses to pain. *Neuropsychologia*, *64*, 263–270. doi:10.1016/j.neuropsychologia.2014.09.045

Cook, S. W., & Tanenhaus, M. K. (2009). Embodied communication: Speakers' gestures affect listeners' actions. *Cognition*, *113*, 98–104. doi:10.1016/j.cognition.2009.06.006

Dahl, C. D., Chen, C.-C., & Rasch, M. J. (2014). Own-race and own-species advantages in face perception: A computational view. *Scientific Reports*, *4*. doi:10.1038/srep06654

Davidenko, N. (2007). Silhouetted face profiles: A new methodology for face perception research. *Journal of Vision*, *7*(4), 1–17. doi:10.1167/7.4.6

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, *118*(2), 247–279. doi:10.1037/a0022327

Gaetano, J. M. (2017). Signal detection theory calculator (1.2) [Microsoft Excel workbook]. Retrieved from https://www.researchgate.net/publication/316642315_Signal_detection_theory_calculator_12 doi:10.13140/RG.2.2.26215.85926

Gaetano, J. M., van der Zwan, R., & Brooks, A. R. (2012). Perceiving other people on the basis of categorical multisensory data: Towards a unified theory of person perception. In R. van der Zwan (Ed.), *Current trends in experimental and applied psychology* (Vol. 1, pp. 105–115). Brisbane, QLD: Primrose Hall.

Gaetano, J. M., van der Zwan, R., Oxner, M., Hayward, W. G., Doring, N., Blair, D., & Brooks, A. R. (2016). Converging evidence of ubiquitous male bias in human sex perception. *PLoS ONE*, *11*(2), e0148623. doi:10.1371/journal.pone.0148623

Gaetano, J. M., van der Zwan, R., Blair, D., & Brooks, A. R. (2014). Hands as sex cues: Sensitivity measures, male bias measures, and implications for sex perception mechanisms. *PLoS ONE*, *9*(3), e91032. doi:10.1371/journal.pone.0091032

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*(6), 568–573. doi:10.1038/9224

Gittings, N. S., & Fozard, J. L. (1986). Age related changes in visual acuity. *Experimental Gerontology*, *21*(4–5), 423–433. doi:10.1016/0531-5565(86)90047-1

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction*, *9*(3), 201–219. doi:10.1207/s1532690xci0903_2

Gonzalez, R. (2009). Orthogonal, planned and unplanned comparisons. In *Data analysis for experimental design* (pp. 211–241). New York, NY: Guilford Press.

Hacker, G., Brooks, A., & van der Zwan, R. (2013). Sex discriminations made on the basis of ambiguous visual cues can be affected by the presence of an olfactory cue. *BMC Psychology*, *1*. doi:10.1186/2050-7283-1-10

Haegerstrom-Portnoy, G., Scheck, M. E., & Brabyn, J. A. (1999). Seeing into old age: Vision function beyond acuity. *Optometry & Vision Science*, *76*(3), 141–158. doi:10.1097/00006324-199903000-00014

Hayward, W. G., Rhodes, G., & Schwaninger, A. (2008). An own-race advantage for components as well as configurations in face recognition. *Cognition*, *106*(2), 1017–1027. doi:10.1016/j.cognition.2007.04.002

Herlitz, A., & Lovén, J. (2013). Sex differences and the own-gender bias in face recognition: A meta-analytic review. *Visual Cognition*, *21*(9–10), 1306–1336. doi:10.1080/13506285.2013.823140

Hess, U., Adams, R. B., Jr., Grammer, K., & Kleck, R. E. (2009). Face gender and emotion expression: Are angry women more like men? *Journal of Vision*, *9*(12), 1–8. doi:10.1167/9.12.19

Jaquet, E., Rhodes, G., & Hayward, W. G. (2007). Opposite aftereffects for Chinese and Caucasian faces are selective for social category information and not just physical face differences. *Quarterly Journal of Experimental Psychology*, *60*(11), 1457–1467. doi:10.1080/17470210701467870

Jaquet, E., Rhodes, G., & Hayward, W. G. (2008). Race-contingent aftereffects suggest distinct perceptual norms for different race faces. *Visual Cognition*, *16*(6), 734–753. doi:10.1080/13506280701350647

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology*, *102*, 116–131. doi:10.1037/a0025335

Johnson, K. L., McKay, L. S., & Pollick, F. E. (2011). He throws like a girl (but only when he's sad): emotion affects sex-decoding of biological motion displays. *Cognition*, *119*(2), 265–280. doi:10.1016/j.cognition.2011.01.016

Johnston, R. A., Kanazawa, M., Kato, T., & Oda, M. (1997). Exploring the structure of multidimensional face-space: The effects of age and gender. *Visual Cognition*, *4*(1), 39–57. doi:10.1080/713756750

Junger, J., Pauly, K., Bröhr, S., Birkholz, P., Neuschaefer-Rube, C., Kohler, C., . . . Habel, U. (2013). Sex matters: Neural correlates of voice gender perception. *Neuroimage*, *79*, 275–287. doi:10.1016/j.neuroimage.2013.04.105

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

Kovács, G., Gulyás, B., Savic, I., Perrett, D. I., Cornwell, R. E., Little, A. C., . . . Vidnyánszky, Z. (2004). Smelling human sex hormone-like compounds affects face gender judgment of men. *NeuroReport*, *15*(8), 1275–1277. doi:10.1097/01.wnr.0000130234.51411.0e

Kozlowski, L. T., & Cutting, J. E. (1977). Recognising the sex of a walker from a dynamic point-light display. *Perception and Psychophysics*, *21*(6), 575–580. doi:10.3758/bf03198740

Lewin, C., & Herlitz, A. (2002). Sex differences in face recognition—Women's faces make the difference. *Brain and Cognition*, *50*(1), 121–128. doi:10.1016/S0278-2626(02)00016-7

Li, X., Logan, R. J., & Pastore, R. E. (1991). Perception of acoustic source characteristics: Walking sounds. *Journal of the Acoustical Society of America*, *90*(6), 3036–3049. doi:10.1121/1.401778

Loven, J., Rehnman, J., Wiens, S., Lindholm, T., Peira, N., & Herlitz, A. (2012). Who are you looking at? The influence of face gender on visual attention and memory for own- and other-race faces. *Memory*, *20*(4), 321–331. doi:10.1080/09658211.2012.658064

McGugin, R. W., Newton, A. T., Gore, J. C., & Gauthier, I. (2014). Robust expertise effects in right FFA. *Neuropsychologia*, *63*, 135–144. doi:10.1016/j.neuropsychologia.2014.08.029

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, & Law*, *7*, 3–35. doi:10.1037/1076-8971.7.1.3

O'Toole, A. J., Peterson, J., & Deffenbacher, K. A. (1996). An 'other-race effect' for categorizing faces by sex. *Perception*, *25*(6), 669–676. doi:10.1068/p250669

Pezdek, K., Blandon-Gitlin, I., & Moore, C. (2003). Children's face recognition memory: More evidence for the cross-race effect. *Journal of Applied Psychology*, *88*(4), 760–763. doi:10.1037/0021-9010.88.4.760

Rehnman, J., & Herlitz, A. (2007). Women remember more faces than men do. *Acta Psychologica*, *124*(3), 344–355. doi:10.1016/j.actpsy.2006.04.004

Rhodes, M. G., & Anastasi, J. S. (2011). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin*, *138*, 146–174. doi:10.1037/a0025750

Schouten, B., Troje, N. F., Brooks, A. R., van der Zwan, R., & Verfaillie, K. (2010). The facing bias in biological motion perception: Effects of stimulus gender and participant sex. *Attention, Perception, and Psychophysics*, *72*(5), 1256–1260. doi:10.3758/app.72.5.1256

Sigala, R., Logothetis, N. K., & Rainer, G. (2011). Own-species bias in the representations of monkey and human face categories in the primate temporal lobe. *Journal of Neurophysiology*, *105*(6), 2740–2752. doi:10.1152/jn.00882.2010

Smith, S. M., Lindsay, R. C. L., Pryke, S., & Dysart, J. E. (2001). Postdictors of eyewitness errors – Can false identifications be diagnosed in the cross-race situation? *Psychology Public Policy and Law*, *7*, 153–169. doi:10.1037//1076-8971.7.1.153

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. doi:10.3758/bf03207704

Taylor, A. J. G. (2017). The role of fixations and face gender in facial expression categorization. *Cognition, Brain, Behavior: An Interdisciplinary Journal*, *21*(2), 101–115. doi:10.24193/cbb.2017.21.07

Troje, N. F., Sadr, J., Geyer, H., & Nakayama, K. (2006). Adaptation aftereffects in the perception of gender from biological motion. *Journal of Vision*, *6*(8), 850–857. doi:10.1167/6.8.7

Tsang, T., Ogren, M., Peng, Y., Nguyen, B., Johnson, K. L., Johnson, S. P. (2018). Infant perception of sex differences in biological motion displays. *Journal of Experimental Child Psychology*, *173*, 338–350. doi:10.1016/j.jecp.2018.04.006

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology*, *43*(2), 161–204. doi:10.1080/14640749108400966

Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *44*(4), 671–703. doi:10.1080/14640749208401305

Wallis, J., Lipp, O. V., & Vanman, E. J. (2012). Face age and sex modulate the other-race effect in face recognition. *Attention, Perception and Psychophysics*, *74*(8), 1712–1721. doi:10.3758/s13414-012-0359-z

Wang, L., Shen, H., Tang, F., Zang, Y., & Hu, D. (2012). Combined structural and resting-state functional MRI analysis of sexual dimorphism in the young adult human brain: An MVPA approach. *Neuroimage*, *61*(4), 931–940. doi:10.1016/j.neuroimage.2012.03.080

Wells, G. L., & Olson, E. A. (2001). The other-race effect in eyewitness identification - What do we do about it? *Psychology Public Policy and Law*, *7*, 230–246. doi:10.1037//1076-8971.7.1.230

Wenzlaff, F., Briken, P., & Dekker, A. (2018). If there's a penis, it's most likely a man: Investigating the social construction of gender using eye tracking. *PLoS ONE*, *13*(3), e0193616. doi:10.1371/journal.pone.0193616

White, H., Hock, A., Jubran, R., Heck, A., & Bhatt, R. S. (2018). Visual scanning of male and female bodies in infancy. *Journal of Experimental Child Psychology*, *166*, 79–95. doi:10.1016/j.jecp.2017.08.004

Wild, H. A., Barrett, S. E., Spence, M. J., O'Toole, A. J., Cheng, Y. D., & Brooke, J. (2000). Recognition and sex categorization of adults' and children's faces: Examining performance in the absence of sex-stereotyped cues. *Journal of Experimental Child Psychology*, *77*(4), 269–291. doi:10.1006/jecp.1999.2554

Winer, B. J. (1962). *Statistical principles in experimental design*. New York, NY: McGraw-Hill. doi:10.1037/11774-000

Yamaguchi, M. K., Hirukawa, T., & Kanazawa, S. (1995). Judgment of gender through facial parts. *Perception*, *24*, 563–575. doi:10.1068/p240563

Zhao, M. T., & Hayward, W. (2010). Holistic processing underlies gender judgments of faces. *Attention, Perception and Psychophysics*, *72*(3), 591–596.

Zhao, M. T., Hayward, W. G., & Bülthoff, I. (2014). Holistic processing, contact, and the other-race effect in face recognition. *Vision Research*, *105*, 61–69. doi:10.1016/j.visres.2014.09.006

Zhou, G., Pu, X., Young, S. G., & Tse, C.-S. (2015). Effects of divided attention and social categorization on the own-race bias in face recognition. *Visual Cognition*, *22*(9–10), 1296–1310. doi:10.1080/13506285.2014.998324

Appendix

Ancillary analyses of participant age

The aim of this supplementary study was to test whether participant age might confound the sensitivity effects of interest presented in the main text. To that end, parallel analyses in which participant age was included as covariate were run. Because negative links between age and general visual acuity measures have been documented (e.g. Gittings & Fozard, 1986; Haegerstrom-Portnoy, Scheck, & Brabyn, 1999), participant age in the present study was analysed in a gross sense. More to the point, there is no specific reason to suspect age should systematically affect hand-based sex classifications, and so the current tests were conducted in a post hoc manner.

Specifically, sensitivity data corresponding to colour and silhouette conditions were each subjected to a 2 (own-other race participant group) × 2 (presentation duration group) ANCOVA. To simplify analyses, data were not partitioned further by participant race or sex; each ANCOVA significance criterion was .050 (rather than .025), in order to maximise the overall power of detecting age confounds. The post hoc prediction was hereby tested that participant age was not driving the ORA described within the main text (for a face-based analogue, cf. O'toole et al., 1996). Thus, it is expected that these a posteriori analyses will yield (a) non-significant outcomes for *participant age*, and (b) significant main and/or interaction effects of *stimulus familiarity* (i.e. own- vs. other-race hands), once participant age has been factored out.

Methods

Methods are described in full in the main text. Participants were 80 Caucasians (47 female) and 80 Asians (39 female), on average aged 32.49 ($SD = 11.08$) and 21.50 ($SD = 2.72$), respectively. The unanticipated age difference was found to be significant ($F_{1,157} = 72.50$, $p < .001$, $r = .56$), thus effects of participant age were explored via unplanned analyses.

Results

The outcomes of both sets of analyses supported the notion that ORA as described previously is unrelated to age. With respect first to the colour conditions, it was found that participant age did not influence sensitivity rates ($F_{1,153} = 2.53$, $p = .113$, $r = .13$), nor did its removal nullify the influence of own-other race ($F_{1,153} = 6.86$, $p = .010$, $r = .21$) or presentation duration ($F_{1,153} = 25.08$, $p < .001$, $r = .38$) group differences; the interaction between those factors was not, however, significant ($F_{1,153} = 2.75$, $p = .099$, $r = .13$). Turning now to the silhouette conditions, participant age did not affect group sensitivity for male or female cues ($F_{1,153} = 0.84$, $p = .361$, $r = .07$). When the covariate was factored out, a significant interaction between own-other race and presentation duration was detected ($F_{1,153} = 5.58$, $p = .019$, $r = .19$), though neither main effect was significant (own-other race: $F_{1,153} = 0.82$, $p = .366$, $r = .07$; presentation duration: $F_{1,153} = 0.07$, $p = .789$, $r = .02$). Summary

In sum then, these ancillary analyses at least rule out the possibility that the ORA is merely an artefact of systematic age differences across groups. Given that participant age had no systematic impact on predicted between-group sensitivity outcomes, the reader can be confident that this variable was not skewing outcomes in the corresponding main text. Moreover, with the age differences statistically accounted for, sensitivity rates differed as a function of familiarity but was in the silhouette condition qualified by stimulus exposure time; this agrees with outcomes in the main study in which participant age was ignored.