This article is published in Journal of Research in Personality

---

Full article available here:

https://www.sciencedirect.com/science/article/abs/pii/S0092656619300224

Satchell, L., Davis, J. P., Julle-Danière, E., Tupper, N., & Marshman, P. (2019). Recognising faces but not traits: Accurate personality judgment from faces is unrelated to superior face memory. Journal of Research in Personality, https://doi.org/10.1016/j.jrp.2019.02.002

This is a pre-print version of the paper

Any questions please e-mail j.p.davis@gre.ac.uk

Recognising faces but not traits:

Accurate personality judgment from faces is unrelated to superior face memory

Running head: Unrelated face and trait recognition

Liam Paul Satchell*[a]

Josh P Davis[b]

Eglantine Julle-Danière[c]

Nina Tupper[d]

Paul Marshman[e]

*Corresponding author.

Can be contacted at: liam.satchell@winchester.ac.uk, University of Winchester, Sparkford Road, Winchester, Hampshire, UK, SO22 4NR

[a] Department of Psychology, University of Winchester

[b] Department of Psychology, Social Work and Counselling, University of Greenwich

[c] Department of Psychology, University of Portsmouth

[d] Department of Clinical Psychological Sciences, Maastricht University

[e] Department of Sports and Exercise Science, University of Portsmouth

Acknowledgements: We would like to thank Diandra Bretfelean for their assistance with material preparation and data collection.

This study received no specific financial support.

Author notes:

The data for this research can be found on the Open Science Framework, here; https://osf.io/crjyf/?view_only=c460b4a12895481d8a6ae31e7855e3e6

This study was not pre-registered.

Author contributions were as follows: LS led study conceptualisation. Study methodology was designed by LS, JD, EJD, NT and PM. Extant face memory data collection was led by JD. Data collection for this study was conducted by LS, EJD and NT. Software for personality and self-selected photograph collection was designed by NT and PM. Data analysis was conducted by LS. First draft was written by LS. Substantial editing was contributed from JD, NT and EJD.

Recognising faces but not traits:
Accurate personality judgment from faces is unrelated to superior face memory

Abstract

It is suggested that accurate personality judgments of faces are driven by a morphological 'kernel of truth' from face shape. We hypothesised that this relationship could lead to those with better face identification ability being better at personality judgments. We investigated the relationship between face memory, face matching, Big Five personality traits, and accuracy in recognising Big Five personality traits from 50 photographs of unknown faces. In our sample (n = 792) there was overall good (but varying) face memory and personality judgment accuracy. However, there was convincing evidence that these two skills do not correlate (all $r < .06$). We also replicate the known relationship between extraversion and face memory ability in the largest sample to date.

Keywords

Personality judgment; Kernel of Truth; Face memory; Face perception

Recognising faces but not traits:

Accurate personality judgment from faces is unrelated to superior face memory

Introduction

A growing body of research studies individuals known as 'super-recognisers' (e.g., Bobak, Bennetts, Parris, Jansari, & Bate, 2016; Bobak, Hancock, & Bate, 2016; Davis, Lander, Evans, & Jansari, 2016; Russell, Duchaine, & Nakayama, 2009). These individuals are classified by their exceptionally good accuracy at recognising the faces of people they have not met, compared to the poor face recognition and face matching of people in general (Megreya & Burton, 2008). Elsewhere, research has demonstrated that personality traits can be recognised from photographs of faces (Borkenau, Brecke, Möttig & Paelecke, 2009; Gordon & Platek, 2009; Gosling, Augustine, Vazire, Holtzman, & Gaddis, 2011; Naumann, Vazire, Rentfrow, & Gosling, 2009). The 'accuracy' (the relationship between self-reported and other-judged personality traits) of such judgments has been explained by the 'kernel of truth' hypothesis (Berry, 1990; Penton-Voak, Pound, Little, & Perrett, 2006), which suggest that face biometrics are indicative of personality. Recent work by Kocsor and Bereckei (2017) has shown that perceptions of an unknown individual's traits is related to previously defined associations between face shape and disposition. In other words, individuals learn to link socially relevant information (i.e., judgements of personality traits) to face shapes and then project that social information onto new people they encounter whose face shapes are similar to those they have seen before. This is particularly interesting because face shapes are also relevant information for accurate recognition (memory and matching) of faces. If those with superior face memory are adept at recognising face shapes, and the personality trait recognition is supposedly supported by face shape, then this would suggest an association between these two skills. This current study investigates this previously unstudied relationship in order to understand if these two-person perception skills are complementary, unique, or unrelated.

Individual differences in face recognition. Individual differences paradigms suggest that most abilities vary across the population, so that few individuals perform exceptionally poorly or well at a task and most perform somewhere in between. We can see this in face recognition where some individuals with no known brain damage have notably poor face recognition accuracy (so called face-blindness or *developmental prosopagnosia*) and some individuals can be identified as face super-recognisers (e.g. Russell et al., 2009). Face recognition is a term that encompasses both the ability to remember previously seen faces and the ability to identify two images of faces as belonging to the same person. The variability in face recognition ability is associated with many factors, including face processing style (e.g. holistic *vs*. feature-based processing; e.g. DeGutis, Wilmer, Mercado, & Cohan, 2013; Richler, Cheung, & Gauthier, 2011; Wang, Li, Fang, Tian, & Liu, 2012), age (Germine, Duchaine, & Nakayama, 2011) and sex (Sommer et al., 2013). It is heritable (Willmer et al., 2010; Zhu et al., 2010) and there is evidence of genetic facial recognition specificity (Shakeshaft & Plomin, 2015). When examining individual differences, face memory and matching abilities relate to the personality trait of extraversion (Lander & Poyarekar, 2015; Li, Tian, Fang, Xu, Li & Liu, 2010) and social anxiety (Davis et al., 2011; Megreya & Bindemann, 2013; however see also Bobak, Pampoulov, & Bate 2016). There is also some evidence that super-recognition may be a unique skill, in that super-recognition may not co-occur with other cognitive skills (Bobak, Bennetts, Parris, Jansari & Bate, 2016). For example, super-recognisers do not demonstrate superior recognition or memory for objects. Additionally, while some individuals working in forensic settings possess exceptional face matching skills with no known superior memory for faces (White et al., 2015; see White, Norell, Phillips, & O'Toole, 2017 for a review), others with outstanding face memory perform relatively poorly at face matching (e.g. Davis et al., 2016).

Accurate face recognition supported by face shape. Those who have notably good face recognition ability (super-recognisers) can excel in surveillance roles, especially in making suspect identifications from CCTV footage (Davis et al., 2016; Robertson, Noyes, Dowsett, Jenkins, & Burton, 2016). CCTV footage is often poor quality and it is notable that super-recognisers retain a strong identification rate when dealing with low resolution images of 15-year old degraded familiar faces (Davis et al., 2016). It is important to note that degraded images do not contain detailed face features, suggesting that super-recognisers use holistic mechanisms to process general face shape more than features of the face (see also Russell et al., 2009). This suggestion is supported by face processing research (DeGutis et al., 2013; Ellis, Shepherd & Davies, 1979; Megreya & Bindemann, 2009). Because these same facial features are those implicated in the 'kernel of truth' hypotheses of personality trait recognition (detailed below), it is possible that the style of face processing used by super-recognisers could benefit them in recognising personality traits. If super-recognisers are being used in surveillance roles, and such roles could include asking security personnel to detect impending criminal incidents (Troscianko et al., 2004), then it is of interest to know if super-recognisers are advantaged in detecting personality traits. As superior face memory and matching do not always co-occur, tests measuring each skill were included in the current study.

Individual differences in personality judgment accuracy. There is a long history of research demonstrating that people can make accurate personality judgments with limited previous interaction (Albright, et al., 1988; Back, Schmukle, & Egloff, 2010; Funder, 1980; 2012; Kenny & Albright, 1987; Levesque & Kenny, 1993; Molen et al, 2017; Vazire, 2010). In the modern age of online interaction and social media, people increasingly perform their first impressions based on photographs of faces (Vazire & Gosling, 2004; Naumann et al., 2009). With this, there has been an increase in the psychological study of these first impressions (Borkenau, et al., 2009; Carré & McCormick, 2008; Carré, et al., 2009; Gordon & Platek, 2009; Gosling, et al., 2011). Generally, participants in these studies show good accuracy at detecting the 'Big Five' personality of targets from observing photographs of unknown faces (e.g. Naumann et al., 2009).

The methodology of these studies is relatively consistent. Participants are presented with photos of unfamiliar target faces, and judge the personality traits (or everyday adjectives relevant to the personality traits) of the targets. The 'accuracy' of the personality judgment is the relationship between the judged personality and the self-reported personality of the targets. This judgment accuracy is known to naturally vary across the population. As is the case with an individual's ability for face memory, personality judgment accuracy is known to be affected by the judge's age (Boshyan, Zebrowitz, Franklin, McCormick, & Carré, 2013), personality (Christiansen, Wolcott-Burnam, Janovics, Burns, & Quick, 2005; Wall, Taylor, & Campbell, 2016), and their ability to create context for personality in interactions (Letzring, 2008). Currently, little is known as to which factors may enable some people to become a 'good judge' (Funder, 2012) of personality. This current study will investigate superior face recognition as a potential correlate.

Accurate personality judgments supported by face shape. Given that judgments of faces can be accurate, researchers formed the 'kernel of truth' hypothesis (Berry, 1990; Penton-Voak et al., 2006). This theory suggests that, if one only needs to see a face to form an accurate personality judgment, then faces should contain fundamental structural information indicative of personality traits. For example, Penton-Voak et al. (2006) observed that hormones such as testosterone and cortisol, growth hormone, and oestrogen have been considered to relate to both face shape and behaviour. Carré, McCormick, and Mondloch (2009) specifically suggest that accurate judgments of aggression from another person's face are supported by this mechanism. For their hypothesis, they combined previous literature finding a relationship between face shape and testosterone, and between testosterone and aggression, to suggest judgments are driven by a mutual hormone mechanism. There is currently not the evidence base to suggest any clear hormonal system that would affect both facial morphology and common personality traits (such as the Big Five). However, research using facial composites has shown that personality can be accurately inferred from faces created from the average face of high and low scoring Big Five trait people (Little & Perrett, 2007), and that internal facial features are particularly important for these judgments (Kramer & Ward, 2010).

There is also a broad previous literature that argues that face morphometry (e.g., face width-to-height measurements) is related to accurate trait judgments (Haselhuhn & Wong, 2011; Jia, Lent, & Zeng, 2014; Ma, Xu, & Luo, 2015; Rule, Krendl, Ivcevic, & Ambady, 2013; Stirrat & Perrett, 2010; Wong, Ormiston, & Haselhuhn, 2011; Zilioli et al., 2015) and is a signal of an individual's internal state (Geniole, Denson, Dixson, Carré, & McCormick, 2015; Hehman, Leitner, Deegan, & Gaertner, 2013; Lefevre, Lewis, Perrett, & Penke, 2013; Stirrat & Perrett, 2012; Whitehouse et al., 2015). However, this literature is not without critique, and there are different social perception findings using similar methodology (Deaner, Goetz, Shattuck, & Schnotala, 2012; Efferson & Vogt, 2013; Gómez-Valdés et al, 2013; Özener, 2012) or varying the standard experimental materials of static, face-on, photographs of faces (Hehman, Flake, & Freeman, 2015; Kosinksi, 2017; Sanchez-Pages, Rodriguez-Ruiz, & Turiegano, 2014; Todorov & Porter, 2014). Much more work needs to be conducted to understand the relationship between facial morphometry, personality traits and social perception. Understanding the role of the perceiver's ability to recognise faces would assist this developing area.

In summary, this extant literature suggests that personality judgments from photographs of faces can be accurate and that these judgments could be driven by biometric facial structures or features, the same features that those who are better at face memory may more readily identify (see above).

Current study. The current study brings together the research on individual differences in face recognition (memory and matching) and personality trait judgment accuracy. We ask three main questions. First, in an attempt to replicate previous research in a larger sample, do Big Five personality traits relate to individual differences in face recognition (both face memory and face matching)? Second, do Big Five personality traits relate to individual differences in personality judgment accuracy? And third, does a heightened ability to recognise faces correlate with a heightened ability to recognise these traits in others? From the 'kernel of truth' hypothesis we would predict that these two abilities would positively correlate.

Method

Ethical approval was obtained for the target and the participant data collection phases. Targets volunteered their self-selected photographs on a custom website. They gave informed consent before submitting their photograph and their personality data. Ethical permission was obtained from the University of [REDACTED]'s Science Faculty Ethics Committee, with the approval code XXXX XXXX-XXX. The face recognition data collected prior to the current project was approved by the University of [REDACTED]'s Research Ethics Committee. Both institutions approved the collection of person judgment data using these materials. The data used in this study can be found on the Open Science Framework; https://osf.io/crjyf/?view_only=c460b4a12895481d8a6ae31e7855e3e6

Participants. All participants provided informed consent before participating. We had no expectations of a size of effect, given the lack of previous research. We intended to recruit at least N= 175 so as to be adequately powered to detect the average effect size in social and personality psychology (r= .21, Richard, bond & Stokes-Zoota, 2003). We did not set an upper limit on our sample size, instead we accepted response to the study for 1 month (July 2017). We recruited our sample from a database of over 80,000 individuals, who had previously engaged with face processing tests (see below) and stated that they would be interested in participating in future research. This database was used because it allowed access to a large sample that would vary in face recognition ability, including individuals classified as super-recognisers. Emails were sent to a random subset of the database (4,140 individuals) asking them to take part in a 'selfie' judgment study. Of those contacted, 864 took part in the study, however 72 participants had incomplete person judgment data for analysis. This left a final sample of 792 participants (M_{Age}= 33.55, SD_{Age}= 10.15, Female = 476). This 'overpowers' our study looking for r= .21 and enables to demonstrate small effects, if present.

Materials.

Extant face memory data. Participants had previously completed the face processing measures used in this study as part of unpublished research. They provided permission to access scores. All had originally completed a fun, 14-trial, anonymous, *Could you be a super-recogniser? Test* linked to media reports about the ability. On completion, they were invited to contribute to online research, which included the Cambridge Face Memory Test: Extended (Duchaine & Nakayama, 2006, Russell et al., 2009); the Glasgow Face Matching Test (short version: Burton*,* White & McNeil, 2010), and two pilot tests (a bespoke Mooney Face and Guitar Recognition tests) not analysed here.

Cambridge Face Memory Test: Extended (CFMT+) (Russell et al., 2009): The CFMT+ is an extended 102-trial version of the CFMT (Duchaine & Nakayama, 2006). Participants are first familiarised with six white-Caucasian male targets, and the test consists of four increasingly difficult sections. The extended section contains heavily degraded images with larger variations in facial expressions and viewpoints and increases in distractor repetitions, all designed to promote holistic facial processing, and to better discriminate between good and exceptional participants.

Most previous research has employed the CFMT+ to assign participants to categorical 'super-recogniser' or control groups. Here we principally analyse the data using the participant's raw score (to avoid reducing variance with categorical data). However, in line with precedent in using the CFMT+, we create groups of 'super-recognisers' and controls based on further analysis. Bobak, et al. (2016) suggested that the super-recogniser scores should be categorised at 95/102 as two standard deviations above their mean (M = 70.72, SD = 12.32). In our sample, overall mean accuracy was far higher than in most studies (M = 83.85, SD = 10.47, Min = 46, Max = 102, see figure 1), perhaps unsurprising due to our participants self-selecting to take part in super-recognition research. For analysis of the CFMT+ we use *i*) the raw scores, *ii*) 'corrected' scores (to normalise the skewed distribution by raising the scores to the power of three) and *iii*) compare categories of super-recognisers (by Bobak et al's above 95/102 criteria, *n*= 106) with a control group (*n*= 289) who performed within Bobak et al.'s normal range ($58 \leq score \leq 83$).

[Insert Figure 1 here]

Glasgow Face Matching Test (GFMT) (short version: Burton *et al.*, 2010). This self-paced, standardised face matching test consists of 40 pairs of simultaneously presented white-Caucasian faces in greyscale. Half the trials are matched (i.e., simultaneously show two different photos of the same person), and half mismatched (i.e., simultaneously show two different photos of different persons who look alike). Participants respond 'same' (20 faces) if they believe the photos to be a match, or 'different' (20 faces) if they believe the photos to be a mismatch. Burton et al. published normalised GFMT scores (*n* = 194; *M* = 81.2% [raw score of 32.48], *SD* = 9.7). As with the CFMT+, our sample scored quite highly (using raw scores; *M* = 37.01, *SD* = 2.42, Min = 27, Max = 40, see figure 2). Again, following the extant analyses of the GFMT, we analyse our data using *i*) raw GFMT scores, *ii*) 'normalised' GFMT scores (raw scores raised to the power of three) and *iii*) we compare 'super-recognisers' (those who scored 100%, *n*= 103) with a control group who performed within Burton et al.'s normal range ($28 \leq score \leq 36$, *n*= 264).

[Insert Figure 2 here]

Targets' self-selected photographs. The photographs of the individuals who acted as targets for the personality judgments were gathered through an online 'Selfies for Science' campaign. The targets were asked to provide a self-selected photograph of themselves in any context as long as the photograph *i*) contained only the target and no other persons and *ii*) was a direct photograph of their face, from the front. These criteria allow many possibilities for targets to express their individuality in the photographs, perhaps helping participants judge their traits (increasing the 'availability' [Funder, 2012] of their traits). A total of 50 targets provided a photograph that met these criteria (Female=44, Male= 6 and M_{Age}= 26.6, SD_{Age}= 9.35). The targets also completed the brief 10 item Big Five Inventory (Rammstedt & John, 2007) when submitting their photograph. Rammstedt and John (2007) highlight good test-retest reliability and self-other agreement on the BFI-10. As others have noted, using short form measures have limitations, such as losing nuance in trait reporting (Smith, McCarthy, & Anderson, 2000). This brief measure of traits was used for efficiency of data collection to maximise responding.

The targets stated the extent to which they *strongly agreed* (5) to *strongly disagreed* (1) that statements described them. The aggregate response to the traits were retained for analysis (Conscientiousness: M = 3.70, SD = 0.92, Agreeableness: M = 3.33, SD = 1.20, Neuroticism: M = 3.24, SD = 1.00, Openness: M = 3.76, SD = 0.93, Extraversion: M = 3.49, SD = 0.98).

Procedure. After consenting to take part in the judgment study, participants firstly self-assessed their own personality using the 10 item Big Five inventory, analysed in the same manner as above (giving self-rated scores of Conscientiousness: M = 3.61, SD = 0.84, Agreeableness: M = 3.24, SD = 0.86, Neuroticism: M = 2.89, SD = 1.02, Openness: M = 3.80, SD = 0.89, Extraversion: M = 2.98, SD = 1.03).

Secondly the participants were presented with the 50 self-selected target photographs. For each photo, participants provided a single, socially-relevant judgment of each of the Big Five personality traits using five-point rating scales. They were asked: "In general, do you think this person is often…" Organised-Disorganised (i.e. Conscientious), Friendly-Unfriendly (i.e. Agreeableness), Anxious-Calm (i.e. Neuroticism), Creative-Not Creative (i.e. Openness) and Extraverted-Introverted (as this definition reflects lay understanding).

Analysis. To compute a measure of personality judgment accuracy, we calculated 'idiographic' correlations (Kolar, Funder, & Colvin, 1996) between participants' trait judgments of the target photos and the targets' self-assessments of their own personality traits (see Brand & Bradley, 2012; Hirschmüller, Egloff, Schmukle, Nestler, & Back, 2015; Kolar et al., 1996; Monin & Oppenheimer, 2005; Satchell, Morris, Akehurst, & Morrison, 2017). Thus, each individual participant will have an accuracy score (derived by Pearson r correlation) between r = 1 (linear accuracy; i.e., more extraverted individuals are rated as more extraverted) to r = -1 (linear inaccuracy; more extraverted individuals are rated as more introverted), with r = 0 indexing no relationship between traits and ratings. Calculating accuracy in this way allows us to describe study performance at the participant level (thus 'idiographic'). This avoid the issue of overestimating sizes of accuracy effects by using a large sample's average (see Monin & Oppenheimer, 2005) whereby a 'wisdom of crowds' effect would make sample-wide accuracy appear better than it is for many of our 792 participants. Further, idiographic accuracy values are necessary for testing our research question, as we can relate each participants' personality judgment accuracy to their own personality traits and their face recognition ability.

Trait judgment accuracy was tested against zero in a one sample *t*-test to demonstrate if accuracy is meaningfully above chance levels. These trait judgment accuracy ratings were then correlated with face recognition accuracy, and the participants' personality traits for analysis (with analysis based on non-zero overlap of 95% CI Pearson's *r*).

Results.

Participant traits and face recognition. The participant performance on the face matching (GFMT) and face memory (CFMT+) tasks was correlated (r= .45 95% CI [.39, .52]). The correlations between participant personality traits and their scores on the face processing tests are depicted in Table 1. Overall, there was only evidence that extraversion related to scores on the CFMT+. No other personality traits correlated with face processing. This pattern of results was the same when analysing the data with Pearson's *r* and Spearman's ranked correlations and was maintained even when transforming the skewed face recognition tests to normal distributions. There was no convincing scale evidence of a general relationship between personality and face recognition.

To fully investigate the differences between super-recognisers and a typical population on personality traits, we computed more statistically liberal comparisons between classifications of participants. A series of independent-measures *t*-tests, with effect sizes measured using Cohen's *d* (hereafter *d*), comparing the personality traits of super-recognisers with average-ability face memorisers all showed non-notable effects (all $t \leq 1.58$, $d \leq .18$). Super-matchers did not differ from average-ability face matchers in self-reported personality (all $t \leq 1.22$, $d \leq .14$). Finally, a created group (n = 35), who excelled at both face memory and matching, did not differ in personality from participants meeting average-ability criteria on both tests (n = 134, all $t \leq 1.52$, $d \leq .29$).

Taken together, these results make a convincing case that super-recognisers do not differ from the general population on personality traits. There was evidence of a weak, but present relationship between extraversion and face recognition across the population.

Table 1.

Correlations [95% CI] between the face processing tests and the Big Five personality traits of participants

	Face Processing Tests			
	Raw scores		Transformed-to-normal scores	
Personality Trait	CFMT+	GFMT	CFMT+^{3}	GFMT^3
Pearson's *r* correlations				
Conscientiousness	.02 [-.06, .09]	.04 [-.04, .11]	.02 [-.05, .09]	.04 [-.03, .11]
Agreeableness	-.02 [-.09, .05]	-.06 [-.12, .01]	-.02 [-.09, .05]	-.06 [-.13, .02]
Neuroticism	-.05 [-.13, .03]	.00 [-.06, .07]	-.05 [-.12, .03]	.01 [-.06, .08]
Openness	.05 [-.03, .11]	.01 [-.06. .08]	.05 [-.03, .12]	.02 [-.05, .09]
Extraversion	**.13 [.05, .20]**	.02 [-.06, .10]	**.12 [.05, .19]**	.03 [-.05, .11]
Spearman's ranked correlations				
Conscientiousness	.02 [-.05, .09]	.04 [-.03, .10]	.02 [-.05, .09]	.04 [-.04, .11]
Agreeableness	-.02 [-.09, .07]	-.05 [-.12, .02]	-.02 [-.09, .06]	-.05 [-.12, .03]
Neuroticism	-.05 [-.12, .03]	.00 [-.07, .07]	-.05 [-.12, .02]	.00 [-.07, .07]
Openness	.05 [-.03, .11]	.03 [-.04, .10]	.05 [-.02, .11]	.03 [-.04, .09]
Extraversion	**.12 [.05, .18]**	.04 [-.03, .11]	**.12 [.04, .19]**	.04 [-.04, .12]

Bold indicates non-zero overlap of 95% CI

The skewness for the raw scores of CFMT+ (-.92) and GFMT (-1.14) was reduced by raising the raw scores to the power of three. This changed the skewness for CFMT+^{3} (-.36) and GFMT^3 (-.75)

Personality trait judgment accuracy. Histograms depicting participants' accuracy at judging each Big Five personality trait are depicted in Figure 3. Participants were least accurate at detecting Agreeableness, with some evidence of inaccuracy (M = -.03, SD = .13, t(791) = -5.93, d = .21). Participants showed better accuracy at detecting Neuroticism (M = .05, SD = .13, t(791) = 11.44, d =

.41) and Extraversion (M = .05, SD = .11, t(791) = 13.04, d = .46). However, there was generally stronger evidence that participants could recognise Conscientiousness (M = .18, SD = .12, t(791) = 41.41, d = 1.47) and Openness (M = .14, SD = .12, t(791) = 33.05, d = 1.17). Overall, this shows that participants were generally accurate at recognising some of the Big Five traits, with sufficient variability to relate these individual differences to other individual differences measures in this study.

[Insert Figure 3 here]

There was no noteworthy evidence that participants' own personality traits affected their trait judgment accuracy. The largest absolute correlation between any participant trait and any trait judgment accuracy was the *negative* relationship between participant Conscientiousness and Conscientiousness judgment accuracy (r(791) = -.11, 95% CI [-.18, -.04]).

Participants were not consistent in their ability to recognise the different traits, highlighted in a notably small intraclass correlation coefficient (ICC = .06 CI [-.05, .16]). Only one participant met potential criteria for 'super-trait-judgment-recognition' (using Z scores $\geq$ 2.00 as our criteria) in more than one trait domain (judging Conscientiousness and Agreeableness). A total of 68 participants (8.6% of sample) met these criteria for super-recognising a single trait judgment domain but no others.

Trait judgment accuracy and face recognition accuracy. There was no evidence that any of the trait judgment accuracy values correlated with participant face processing test scores (see Table 2). There was no evidence of any correlations when the raw scores were correlated with both Pearson's and Spearman's ranked correlations, nor was there any evidence when 'normalised' distributions were used for analysis.

We conducted independent-measures t-tests comparing CFMT+ super-recognisers with average-ability face memorisers (all $t \leq 0.86$, $d \leq .10$), GFMT super-matchers with average-ability face matchers (all $t \leq 1.43$, $d \leq .17$), and those who met both tests' criteria with participants meeting average-ability criteria (all $t \leq 1.55$, $d \leq .30$) on trait judgement accuracy. All tests demonstrated no notable differences with negligible effect sizes.

Even when comparing the co-occurrence between super-recognisers and the participants categorised as high performing judges of traits we found no effects (all $\chi^2 \leq .28$, $\varphi \leq .03$). The same is found when investigating the categorical super-matchers (all $\chi^2 \leq .61$, $\varphi \leq .04$), and super-processors (all $\chi^2 \leq 3.48$, $\varphi \leq .16$).

Overall, using a series of tests of varying statistical liberalism, we can consider our findings robust evidence that recognition and matching of faces and recognition of personality traits in faces are distinct abilities.

Table 2.

Correlations [95% CI] between the face processing tests and the accuracy of participants at judging the Big Five personality traits

	Face Processing Tests			
	Raw scores		Transformed-to-normal scores	
Judged Trait	CFMT+	GFMT	$CFMT+^3$	$GFMT^3$
Pearson's *r* correlations				
Conscientiousness	.02 [-.05, .09]	-.01 [-.08, .07]	.02 [-.05, .09]	-.00 [-.08, .07]
Agreeableness	-.06 [-.13, .01]	-.06 [-.14, .02]	-.06 [-.12, .01]	-.06 [-.14, .01]
Neuroticism	.02 [-.06, .09]	.02 [-.05, .09]	.01 [-.06, .08]	.02 [-.05, .09]
Openness	-.01 [-.08, .07]	.01 [-.06, .09]	-.00 [-.08, .08]	.01 [-.06, .08]
Extraversion	.01 [-.06, .08]	-.02 [-.08, .06]	.01 [-.06, .07]	-.01 [-.08, .06]
Spearman's ranked correlations				
Conscientiousness	.01 [-.07, .08]	.01 [-.06, .08]	.01 [-.07, .08]	.01 [-.06, .08]
Agreeableness	-.03 [-.10, .04]	-.06, [-.13, .01]	-.03 [-.10, .04]	-.06 [-.13, .02]
Neuroticism	-.00 [-.07, .08]	.02 [-.05, .09]	-.00 [-.07, .08]	.02 [-.05, .09]
Openness	-.00 [-.08, .07]	-.00 [-.05, .09]	-.00 [-.07, .08]	-.00 [-.07, .07]
Extraversion	.00 [-.07, .07]	.01 [-.06, .07]	.00 [-.07, .07]	.01 [-.06, .07]

Bold indicates non-zero overlap of 95% CI

The skewness for the raw scores of CFMT+ (-.92) and GFMT (-1.14) was reduced by raising the raw scores to the power of three. This changed the skewness for $CFMT+^3$ (-.36) and $GFMT^3$ (-.75)

Additional analyses controlling for age and sex of targets

Further to our planned analyses, a peer reviewer requested that we explore the potential influence of age and sex of the targets on personality judgment accuracy. For example, as older adults are generally more Conscientious (e.g., Donnellan & Lucas, 2008), it could be the case that accuracy in judging Conscientiousness is not supported by facial morphometry, but by apparent age. In fact, in our current sample of targets with diverse ages, we found that age was positively associated with Conscientiousness ($r(50)= .44$, 95% *CI* [.26, .61]), but not clearly associated with Agreeableness ($r(50)= -.02$, 95% *CI* [-.27, .23]), Neuroticism ($r(50)= -.21$, 95% *CI* [-.44, .03]), Openness ($r(50)= -.12$, 95% *CI* [-.42, .37]) and Extraversion ($r(50)= -.25$, 95% *CI* [-.49, .02]). There are also known sex differences in Big Five scores (Schmitt, Realo, Voracek, & Allik, 2008) although we lack variability in sex to test this specifically in our population.

In order to assess the effect of target age and sex on participant personality judgment accuracy, we additionally conducted idiographic partial correlations. These are partial correlations computed on the participant level, where each particular participants' judgments of the target stimuli were correlated with the target's traits whilst controlling for the targets' age and sex. We can then describe the performance of the sample in those partial correlations and test for differences from the non-partial correlations.

We found that the partial correlations of accuracy were notably different from the standard correlations above, highlighting the importance of age and sex of target in personality judgment accuracy. The partial correlations were largely the opposite direction to the non-partial correlations. In fact, all correlations between participants partial and non-partial accuracy rates were strongly negatively correlated (all $r \leq -.86$). For example, the average idiographic accuracy partial correlation for Conscientiousness across the sample was negative ($M = -.11$, $SD = .13$) as opposed to the sample-wide tendency for positive accuracy demonstrated above. Similarly, the accuracy partial correlations for Openness was now generally negative ($M = -.13$, $SD = .12$). There was no notable sample-wide partial correlation accuracy for Extraversion ($M = -.01$, $SD = .11$), Neuroticism ($M = -.04$, $SD = .13$) and Agreeableness ($M = .03$, $SD = .13$) when controlling for age and sex of targets.

Overall this further analysis showed that age and sex of target were key communicative features of the targets. These features were related to the personality traits and the perceived traits of the targets and so were important for facilitating accurate judgments.

Discussion

In our large study of face recognition accuracy and trait judgment accuracy, we find *i*) substantial individual differences in personality judgment accuracy, with many displaying good ability, *ii*) wide individual differences in face recognition ability, with many showing better-than average ability and *iii*) convincingly, no evidence that these skills are correlated. This is in contradiction to our predictions based on the 'kernel of truth' hypothesis, but in line with the suggestion that face recognition ability is unrelated to other cognitive skills (Bobak et al., 2016). We do find evidence supporting the previous finding that the Big Five personality trait extraversion relates to face recognition ability (Lander & Pyarekar, 2015; Li et al., 2010), with a small but notable correlation in the largest sample to date. Additional analyses suggest that the age of the target person has an important role in facilitating accurate person perception.

Theoretically, the kernel of truth hypothesis of personality trait judgments would suggest that superior recognisers of faces are at an advantage for recognising traits from faces. Good facial memory is associated with a focus on face shape (Burton et al., 1999; Ellis et al., 1979), which has also been implicated in containing trait-relevant information (Little & Perrett, 2007; Kramer & Ward, 2010). The kernel of truth hypothesis would suggest that the recognition of these biometric qualities would lead to more accurate judgments of traits (Penton-Voak et al., 2006). With our findings that those with superior face recognition were not better at trait recognition, we may need to consider other theories of trait judgment accuracy from photographs of faces. We note that our participants were not inaccurate at detecting traits in general, demonstrating that the stimuli provided a suitable test of this skill. In fact, it appears that participants were making use of more general visual information from the targets, such as age and sex, for their judgments. However, trait judgments are not enhanced by superior face processing, (as measured using a face memory and a face matching test).

In our study, targets were asked to submit self-selected photographs to increase the *available* (Funder, 2012) information for personality judgments. Zebrowtiz, Collins, and Dutta (1998) suggest that individuals may change their appearance to be reflective of their personality (termed the 'Dorian Gray' effect), an alternative explanation for how stimuli based on faces may generate personality-relevant criteria. In their longitudinal research on appearance and personality, Zebrowitz et al. (1998) found that women in their 50s had a more attractive appearance if they reported a more 'attractive' personality (sociability) in their 30s. This increased facial attractiveness was attributed to increased cosmetic use. Much like other Dorian Gray effect research (Feingold, 1992; Zebrowtiz, Voinescu & Collins, 1996), there is evidence that how an individual self-presents could be indicative of their personality. This would offer an answer as to how personality judgments from faces could be possible, without reliance on facial morphometry.

Perhaps, in the case of our person judgment paradigm, there were non-biometric properties available in the photographs that were preferentially used by the participants to reach their judgments. For example, additional analyses highlighted the importance of target age in accurate judgments. Given that age is known to be related to personality traits, the apparent age of a face can act as a relevant cue to personality. This is evident in our data, as the personality traits that were more accurately at detected are those best known to vary with age, such as Conscientiousness (Donnellan & Lucas, 2008). As Funder proposes in the four stage Realistic Accuracy Model (2012), judgments of personality are accurate when *i*) relevant behaviours are *ii*) available to be *iii*) detected and *iv*) 'utilized' for a judgment. Behaviours are relevant if they are known to correlate with personality (such as age). These are then made available for detection through the presentation of stimulus people (a methodological consideration) which can then be utilized by participants to reach a judgment (giving rise to individual differences in judges). This model also highlights the possibility that participants are not detecting the available facial morphometry and instead focusing on other information available in the stimuli. Nevertheless, our results still suggest that there is no benefit of superior face recognition on personality judgment accuracy.

One critique of our target presentations could be that there is a variety of information beyond participant face shape that could be affecting perceptions of personality. Going forward, standardising presentation of photographs could eliminate any cosmetic or grooming effects that may be ameliorating personality judgment accuracy. This is the typical approach in experimental research to exploring facial cues to traits. However, this is notably artificial and does not reflect everyday person judgments. In fact, if research on kernel of truth requires abstract, artificial, presentations of individuals, then it suggests this model is not applicable for explaining everyday person judgment accuracy. Regardless of methodology, more research should focus on understanding the validity of the kernel of truth hypothesis, by focusing on different individual differences in the characteristics of judges.

Our results replicate extant work in personality and face recognition ability. Previous research has identified that the personality traits Extraversion (Lander & Pyarekar, 2015; Li et al., 2010) and social-anxiety (Davies et al., 2011; Megreya & Bindemann, 2013; see also Bobak et al., 2016) are associated with face recognition ability. We find that Extraversion does correlate with face recognition but Neuroticism did not. This could be due to the fact that the two Neuroticism questions in the 10 item Big Five inventory do not contain a social component (*I see myself as someone who… is relaxed, handles stress well* and *gets nervous easily;* Rammstedt & John, 2007). This has been suggested by others, who have cautioned against using short-form personality measures as they may not capture all domains of the trait (Smith, McCarthy & Anderson, 2000). In our case, the 10 item Big Five inventory allowed us to efficiently collect an overview of the key personality traits and replicate the extraversion correlation in a larger sample. However, future research should continue to explore socially-relevant anxiety and interpersonal behaviour as potential correlates for face recognition ability.

We are mindful of constraints on the generality of our current paper (Simons, Shoda, & Lindsay, 2017). Our study is reliant on volunteer participation for the face processing data, self-selected photograph donation and personality judgment data aspects of the study. We do observe a trend in that higher scoring face processors chose to engage with our study. Further, all three participant groups of the study were primarily recruited from the UK. We can consider our large dataset, with a broad range of ages and personalities, to be reasonably representative of individuals within the UK who are interested in psychological research. This population is the same as those typically used in personality judgment and face memory like our own. We suggest that caution should be used when applying our results to other countries and languages, but highlight the relevance of our work to similar face processing and personality judgment research.

Conclusions

This paper has explored the relationship between face recognition, personality trait judgment accuracy and personality. With our large sample size, we can convincingly suggest that there is no relationship between heightened trait and heightened face recognition. These results should be noted in applied (super-recogniser surveillance personnel may not be better at detecting threats in general) and theoretical research (the validity of the kernel of truth hypothesis) settings. Importantly, we find further evidence that extraversion is an important correlate of face memory ability.

References

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, *55*(3), 387–395. doi:10.1037/0022-3514.55.3.387

Back, M. D., Schmukle, S. C., & Egloff, B. (2010). Why are narcissists so charming at first sight? Decoding the narcissism–popularity link at zero acquaintance. *Journal of Personality and Social Psychology*, *98*(1), 132–145. doi:10.1037/a0016338

Berry, D. S. (1990). Taking people at face value: Evidence for the kernel of truth hypothesis. *Social Cognition*, *8*(4), 343-361. https://doi.org/10.1521/soco.1990.8.4.343

Bobak, A. K., Bennetts, R. J., Parris, B. A., Jansari, A., & Bate, S. (2016). An in-depth cognitive examination of individuals with superior face recognition skills. *Cortex*, *82,* 48-62. doi:10.1016/j.cortex.2016.05.003

Bobak, A. K., Hancock, P. J. B., & Bate, S. (2016). Super-Recognizers in action: Evidence from face matching and face memory tasks. *Applied Cognitive Psychology, 30*(1), 81-91. doi: 10.1002/acp.3170

Bobak, A. K., Pampoulov, P., & Bate, S. (2016). Detecting superior face recognition skills in a large sample of young British adults. *Frontiers in Psychology, 22*(7), 1-16. doi:10.3389/ fpsyg.2016.01378

Borkenau, P., Brecke, S., Möttig, C., & Paelecke, M. (2009). Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of Research in Personality*, *43*(4), 703-706. http://dx.doi.org/10.1016/j.jrp.2009.03.007

Boshyan, J., Zebrowitz, L. A., Franklin, R. G., Jr., McCormick, C. M., & Carré, J. M. (2013). Age similarities in recognizing threat from faces and diagnostic cues. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69(5), 710-718. doi:10.1093/geronb/gbt054

Brand, A., & Bradley, M. T. (2012). More voodoo correlations: when average-based measures inflate correlations. *The Journal of General Psychology*, *139*(4), 260-272. doi:10.1080/00221309.2012.703711

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods, 42*, 286–291. doi:10.3758/BRM.42.1.286

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science*, *10*(3), 243–248. doi: 10.1111/1467-9280.00144

Carre, J. M., & McCormick, C. M. (2008). In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1651), 2651–2656. doi:10.1098/rspb.2008.0873

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science*, 20(10), 1194–1198. doi:10.1111/j.1467-9280.2009.02423.x

Christiansen, N. D., Wolcott-Burnam, S., Janovics, J. E., Burns, G. N., & Quirk, S. W. (2005). The good judge revisited: individual differences in the accuracy of personality judgments. *Human Performance, 18*(2), 123–149. doi:10.1207/s15327043hup1802_2

Davis, J. M., McKone, E., Dennett, H., O'Connor, K. B., O'Kearney, R., & Palermo, R. (2011). Individual differences in the ability to recognise facial identity are associated with social anxiety. *PloS one*, *6*(12), doi:10.1371/journal.oone.0028800

Davis, J. P., Lander, K., Evans, R., & Jansari, A. (2016). Investigating predictors of superior face recognition ability in police super-recognisers. *Applied Cognitive Psychology, 30*(6)*,* 827–840. doi: 10.1002/acp.3260

Deaner, R. O., Goetz, S. M. M., Shattuck, K., & Schnotala, T. (2012). Body weight, not facial width-to-height ratio, predicts aggression in pro hockey players. *Journal of Research in Personality, 46*(2), 235–238. doi:10.1016/j.jrp.2012.01.005

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*, *126*(1), 87-100. DOI: 10.1016/j.cognition.2012.09.004

Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the big five across the life span: Evidence from two national samples. *Psychology and Aging, 23*(3), 558–566. doi:10.1037/a0012897

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*(4), 576-585. DOI:10.1016/j.neuropsychologia.2005.07.001

Efferson, C., & Vogt, S. (2013). Viewing men's faces does not lead to accurate predictions of trustworthiness. *Scientific Reports, 3*, 1047. doi: 10.1038/srep01047

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*, *8*(4), 431-439. DOI:10.1068/p080431

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin*, *111*(2), 304.

Funder, D. C. (1980). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *Journal of Personality*, *48*(4), 473–493. doi:10.1111/j.1467-6494.1980.tb02380.x

Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, *21*(3), 177-182. doi: 10.1177/0963721412445309Geniole, S. N., Denson, T. F., Dixson, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from Meta-Analyses of the Facial Width-to-Height Ratio as an Evolved Cue of Threat. *PLOS ONE, 10*(7), e0132726. doi:10.1371/journal.pone.0132726

Germine, L. T., Duchaine, B., & Nakayama, K. (2011). Where cognitive development and aging meet: Face learning ability peaks after age 30. *Cognition*, *118*(2), 201-210. DOI:10.1016/j.cognition.2010.11.002

Gómez-Valdés, J., Hünemeier, T., Quinto-Sánchez, M., Paschetta, C., de Azevedo, S., González, M. F., … González-José, R. (2013). Lack of Support for the Association between Facial Shape and Aggression: A Reappraisal Based on a Worldwide Population Genetics Perspective. *PLoS ONE, 8*(1), e52317. doi:10.1371/journal.pone.0052317

Gordon, D. S., & Platek, S. M. (2009). Trustworthy? The brain knows: Implicit neural responses to faces that vary in dark triad personality characteristics and trustworthiness. *Journal of Social, Evolutionary, and Cultural Psychology*, *3*(3), 182. http://dx.doi.org/10.1037/h0099323

Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N., & Gaddis, S. (2011). Manifestations of personality in online social networks: self-reported Facebook-related behaviors and observable profile information. *Cyberpsychology, Behavior, and Social Networking*, *14*(9), 483–488. doi:10.1089/cyber.2010.0087

Haselhuhn, M. P., & Wong, E. M. (2011). Bad to the bone: facial structure predicts unethical behaviour. *Proceedings of the Royal Society B: Biological Sciences, 279*(1728), 571–576. doi:10.1098/rspb.2011.1193

Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and Dynamic Facial Cues Differentially Affect the Consistency of Social Evaluations. *Personality and Social Psychology Bulletin, 41*(8), 1123–1134. doi:10.1177/0146167215591495

Hehman, E., Leitner, J. B., Deegan, M. P., & Gaertner, S. L. (2013). Facial Structure Is Indicative of Explicit Support for Prejudicial Beliefs. *Psychological Science, 24*(3), 289–296. doi:10.1177/0956797612451467

Hirschmüller, S., Egloff, B., Schmukle, S. C., Nestler, S., & Back, M. D. (2015). Accurate judgments of neuroticism at zero acquaintance: A question of relevance. *Journal of Personality*, *83*(2), 221-228. doi:10.1111/jopy.12097

Jia, Y., Lent, L. V., & Zeng, Y. (2014). Masculinity, Testosterone, and Financial Misreporting. *Journal of Accounting Research*, *52*(5), 1195–1246. doi:10.1111/1475-679x.12065

Kocsor, F., & Bereczkei, T. (2017). First impressions of strangers rely on generalization of behavioral traits associated with previously seen facial features. *Current Psychology, 36*(3), 385–391. doi:10.1007/s12144-016-9427-1

Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality*, *64*, 311-338. doi:10.1111/j.1467-6494.1996.tb00513.x

Kosinski, M. (2017). Facial Width-to-Height Ratio Does Not Predict Self-Reported Behavioral Tendencies. *Psychological Science, 28*(11), 1675–1682. doi:10.1177/0956797617716929

Kramer, R. S. S., & Ward, R. (2010). Internal facial features are signals of personality and health. *Quarterly Journal of Experimental Psychology*, *63*(11), 2273–2287. doi:10.1080/17470211003770912

Lander, K., & Poyarekar, S. (2015). Famous face recognition, face matching and extroversion. *The Quarterly Journal of Experimental Psychology, 68*(9), 1769-1776. https://doi.org/10.1080/17470218.2014.988737

Latinus, M., & Taylor, M. J. (2005). Holistic processing of faces: Learning effects with Mooney faces. *Journal of Cognitive Neuroscience, 17*, 1316–1327. doi:10.1162/0898929055002490

Latinus, M., & Taylor, M. J. (2006). Face processing stages: Impact of difficulty and the separation of effects. *Brain Research*, 1123, 179–187. doi:10.1016/j.brainres.2006.09.031

Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: facial width is associated with testosterone levels in men. *Evolution and Human Behavior, 34*(4), 273–279. doi:10.1016/j.evolhumbehav.2013.03.005

Letzring, T. D. (2008). The good judge of personality: Characteristics, behaviors, and observer accuracy. *Journal of Research in Personality, 42*(4), 914–932. doi:10.1016/j.jrp.2007.12.003

Levesque, M. J., & Kenny, D. A. (1993). Accuracy of behavioral predictions at zero acquaintance: A social relations analysis. *Journal of Personality and Social Psychology*, *65*, 1178.

Li, J., Tian, M., Fang, H., Xu, M., Li, H., & Liu, J. (2010). Extroversion predicts individual differences in face recognition. *Communication and Integrative Biology, 3,* 295-298. http://dx.doi.org/10.4161/cib.3.4.12093

Little, A. C., & Perrett, D. I. (2007). Using composite images to assess accuracy in personality attribution to faces. *British Journal of Psychology*, *98*(1), 111–126. doi:10.1348/000712606x109648

Ma, F., Xu, F. & Luo, X. (2015) Children's and adults' judgments of facial trustworthiness: The relationship to facial attractiveness. *Perceptual and Motor Skills* 211. 179–98. doi: 10.2466/27.22.PMS.121c10x1

Megreya, A. M., & Bindemann, M. (2013). Individual differences in personality and face identification. *Journal of Cognitive Psychology*, *25*(1), 30-37. http://dx.doi.org/10.1080/20445911.2012.739153

Megreya, A. M., & Burton, A. M. (2008). Matching faces to photographs: Poor performance in eyewitness memory (without the memory). *Journal of Experimental Psychology: Applied, 14*(4), 364-372. DOI: 10.1037/a0013464

Monin, B., & Oppenheimer, D. M. (2005). Correlated averages vs. averaged correlations: Demonstrating the warm glow heuristic beyond aggregation. *Social Cognition*, *23*(3), 257-278. doi:10.1521/soco.2005.23.3.257

Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality Judgments Based on Physical Appearance. *Personality and Social Psychology Bulletin, 35*(12), 1661–1671. doi:10.1177/0146167209346309

Özener, B. (2012). Facial width-to-height ratio in a Turkish population is not sexually dimorphic and is unrelated to aggressive behavior. *Evolution and Human Behavior, 33*(3), 169–173. doi:10.1016/j.evolhumbehav.2011.08.001

Penton-Voak, I. S., Pound, N., Little, A. C., & Perrett, D. I. (2006). Personality judgments from natural and composite facial images: More evidence for a "kernel of truth" in social perception. *Social Cognition*, *24*(5), 607-640. https://doi.org/10.1521/soco.2006.24.5.607

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. doi:10.1016/j.jrp.2006.02.001

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable Facial Configurations Affect Strategic Choices in Trust Games with or without Information about Past Behavior. *PLoS ONE, 7*(3), e34293. doi:10.1371/journal.pone.0034293

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, 7(4), 331–363. doi:10.1037/1089-2680.7.4.331

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science*, *22*(4), 464-471. DOI: 10.1177/0956797611401753

Robertson, D. J., Noyes, E., Dowsett, A. J., Jenkins, R., & Burton, A. M. (2016). Face recognition by Metropolitan Police super-recognisers. *PloS One, 11*(2), e0150036–8. doi: 10.1371/journal.pone.0150036

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology, 104*(3), 409–426. doi:10.1037/a0031050

Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review, 16*, 252-257. doi:10.3758/PBR.16.2.252

Sanchez-Pages, S., Rodriguez-Ruiz, C., & Turiegano, E. (2014). Facial Masculinity: How the Choice of Measurement Method Enables to Detect Its Influence on Behaviour. *PLoS ONE, 9*(11), e112157. doi:10.1371/journal.pone.0112157

Satchell, L., Morris, P, Akehurst, L., & Morrison, E. (2017). Can judgments of threat reflect an approaching person's trait aggression? *Current Psychology.* doi: 10.1007/s12144-016-9557-5

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of personality and social psychology*, *94*(1), 168. doi: 10.1037/a0014651Shakeshaft, N. G., & Plomin, R. (2015). Genetic specificity of face recognition. *Proceedings of the National Academy of Sciences*, *112*(41), 12887-12892. https://doi.org/10.1073/pnas.1421881112

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on Generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science, 12*(6), 1123–1128. doi:10.1177/1745691617708630

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, *12*(1), 102–111. doi:10.1037/1040-3590.12.1.102

Sommer, W., Hildebrandt, A., Kunina-Habenicht, O., Schacht, A., & Wilhelm, O. (2013). Sex differences in face cognition. *Acta Psychologica*, *142,* 62–73. doi: 10.1016/j.actpsy.2012.11.001

Stirrat, M., & Perrett, D. I. (2010). Valid Facial Cues to Cooperation and Trust. *Psychological Science*, *21*(3), 349–354. doi:10.1177/0956797610362647

Stirrat, M., & Perrett, D. I. (2012). Face Structure Predicts Cooperation. *Psychological Science, 23*(7), 718–722. doi:10.1177/0956797611435133

Todorov, A., & Porter, J. M. (2014). Misleading First Impressions. *Psychological Science, 25*(7), 1404–1417. doi:10.1177/0956797614532474

Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., & Wilson, A. (2004). What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception*, 33(1), 87–101. doi:10.1068/p3402

Vander Molen, R. J., Kaplan, S., Choi, E., & Montoya, D. (2018). Judgments of the Dark Triad based on Facebook profiles. *Journal of Research in Personality, 73,* 150–163. doi:10.1016/j.jrp.2017.11.010

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*, 281-300. DOI: 10.1037/a0017908

Vazire, S., & Gosling, S. D. (2004). e-Perceptions: personality impressions based on personal websites. *Journal of Personality and Social Psychology*, *87*(1), 123-132. http://dx.doi.org/10.1037/0022-3514.87.1.123

Wall, H. J., Taylor, P. J., & Campbell, C. (2016). Getting the balance right? A mismatch in interaction demands between target and judge impacts on judgement accuracy for some traits but not others. *Personality and Individual Differences*, *88*, 66-72. https://doi.org/10.1016/j.paid.2015.08.037

Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science*, *23*(2), 169–177. DOI: 10.1371/journal.pone.0058253

White, D., Norell, K., Phillips, P. J., & O'Toole, A. J. (2017). Human factors in forensic face identification. In *Handbook of Biometrics for Forensic Science* (pp. 195-218). Springer International Publishing.

White, D., Phillips, P. J., Hahn, C. A., Hill, M., O'Toole, A. J., & White, D. (2015). Perceptual expertise in forensic facial image comparison. *Proceedings of the Royal Society B: Biological Sciences*, *282*, 20151292. DOI: 10.1098/rspb.2015.1292

Whitehouse, A. J. O., Gilani, S. Z., Shafait, F., Mian, A., Tan, D. W., Maybery, M. T., … Eastwood, P. (2015). Prenatal testosterone exposure is related to sexually dimorphic facial morphology in adulthood. *Proceedings of the Royal Society B: Biological Sciences, 282*(1816), 20151351. doi:10.1098/rspb.2015.1351

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., ... & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences*, *107*(11), 5238-5241. https://doi.org/10.1073/pnas.0913053107

Wong, E. M., Ormiston, M. E., & Haselhuhn, M. P. (2011). A Face Only an Investor Could Love. *Psychological Science, 22*(12), 1478–1483. doi:10.1177/0956797611418838

Zebrowitz, L. A., Collins, M. A., & Dutta, R. (1998). The relationship between appearance and personality across the life span. *Personality and Social Psychology Bulletin*, *24*(7), 736-749. https://doi.org/10.1177/0146167298247006

Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-Eyed" and" Crooked-Faced": determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin*, *22*(12), 1258-1269. http://dx.doi.org/10.1177/01461672962212006

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., ... & Liu, J. (2010). Heritability of the specific cognitive ability of face perception. *Current Biology*, *20*(2), 137-142. doi: 10.1016/j.cub.2009.11.067

Zilioli, S., Sell, A. N., Stirrat, M., Jagore, J., Vickerman, W., & Watson, N. V. (2014). Face of a fighter: Bizygomatic width as a cue of formidability. *Aggressive Behavior, 41*(4), 322–330. doi:10.1002/ab.21544

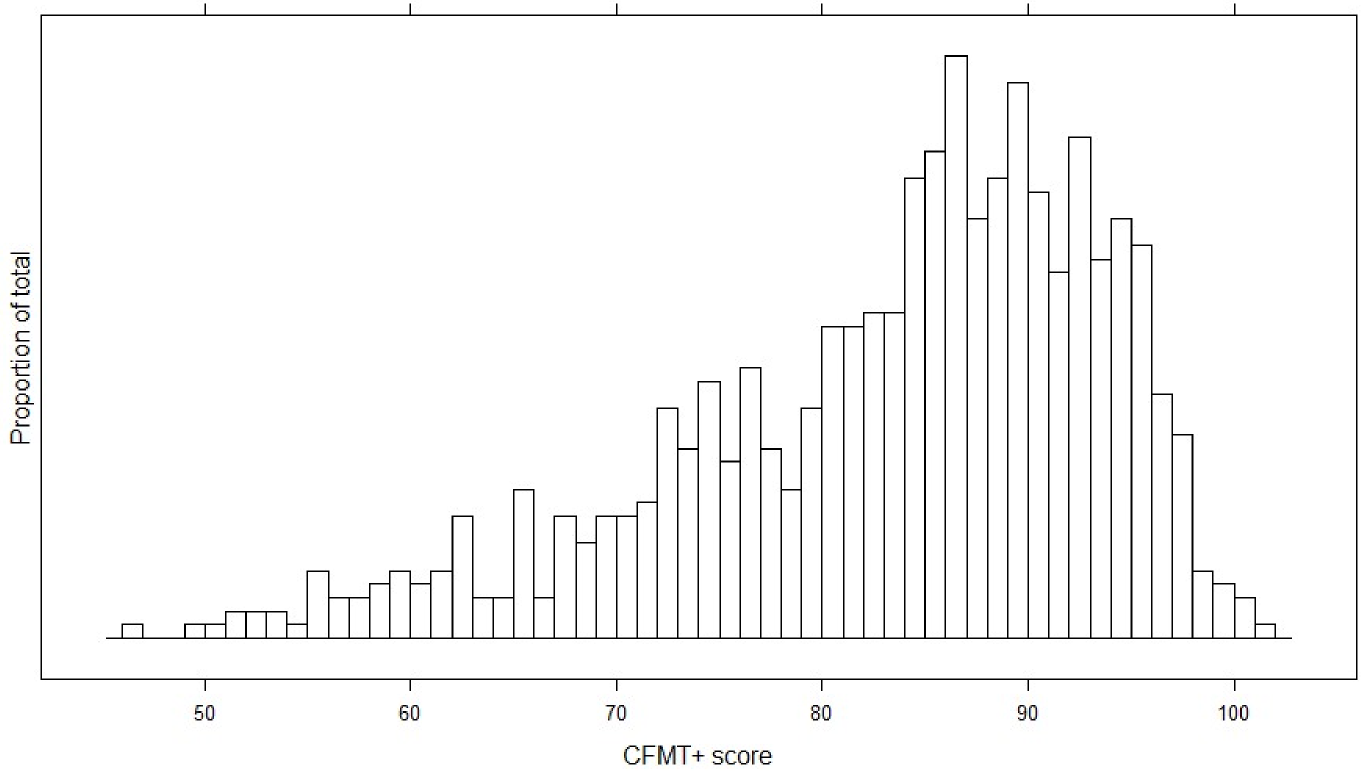

Figure 1. The skewed distribution of Cambridge Face Memory Test: Extended in our sample.

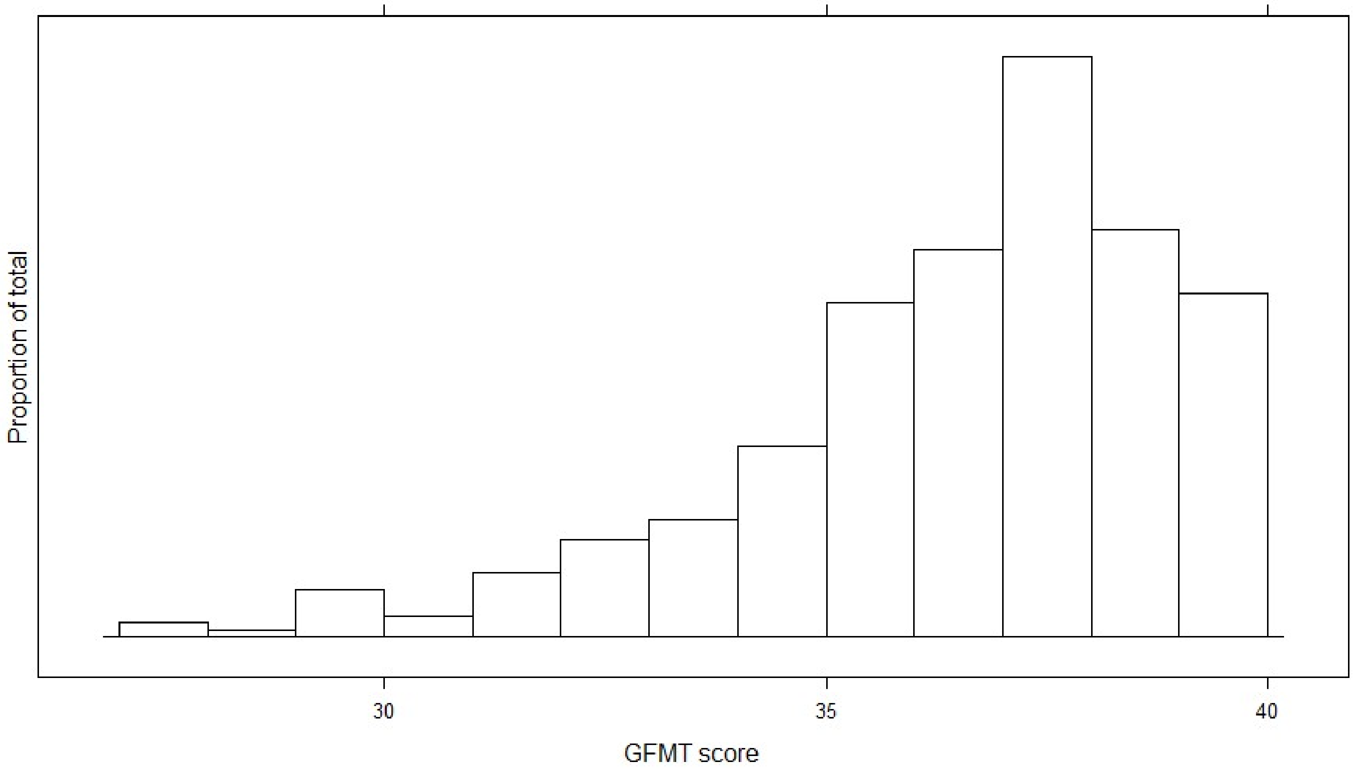

Figure 2. The skewed distribution of Glasgow Face Matching Test in our sample.

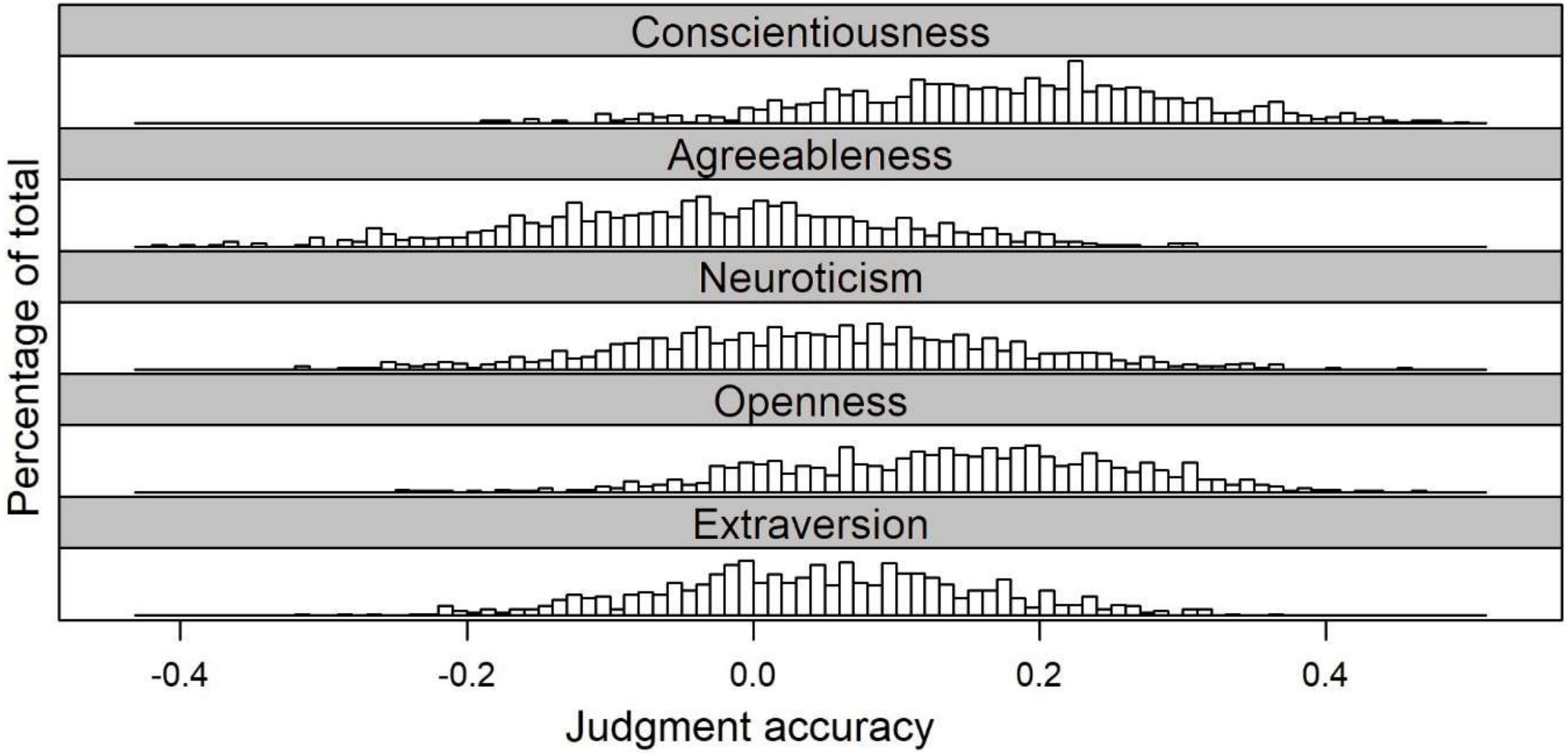

Figure 3. Histograms with equalised axes for the accuracy of our sample at judging Conscientiousness, Agreeableness, Neuroticism, Openness and Extraversion. Note that $r = 0$ represents no accuracy, $r > 0$ is more accurate and $r < 0$ is more inaccurate responding.