Do people have insight into their face recognition abilities?

Journal:	*Quarterly Journal of Experimental Psychology*
Manuscript ID	QJE-SIP 15-305.R1
Manuscript Type:	Special Issue Paper
Date Submitted by the Author:	15-Jan-2016
Complete List of Authors:	Palermo, Romina; The University of Western Australia, The School of Psychology, ARC Centre of Excellence in Cognition and its Disorders Rossion, Bruno; Institute of Research in Psychology (IPSY) and Institute of Neuroscience (IoNS), Rhodes, Gillian; University of Western Australia, Laguesse, Renaud; Université Catholique de Louvain, Tez, Tolga; Australian National University, Hall, Bronwyn; Australian National University, Albonico, Andrea; Università degli Studi di Milano-Bicocca, malaspina, manuela; Università degli Studi di Milano-Bicocca, Daini, Roberta; Università degli Studi di Milano-Bicocca, Irons, Jesscia; Australian National University, Al-Janabi, Shahd; Macquarie University, Taylor, Libby; The University of Western Australia, The School of Psychology, ARC Centre of Excellence in Cognition and its Disorders Rivolta, Davide; Macquarie University, McKone, Elinor; Australian National University,
Keywords:	face perception, self-evaluation, prosopagnosia, individual differences, metacognition, accuracy

Do people have insight into their face recognition abilities?

Romina Palermo[1,2,3], Bruno Rossion[4], Gillian Rhodes[1,2], Renaud Laguesse[4], Tolga Tez[3], Bronwyn Hall[3], Andrea Albonico[5], Manuela Malaspina[5], Roberta Daini[5], Jessica Irons[3], Shahd Al-Janabi[6], Libby C. Taylor[1,2], Davide Rivolta[6], Elinor McKone[1,3]

[1] ARC Centre of Excellence in Cognition and its Disorders
[2] School of Psychology, University of Western Australia, Perth, WA, Australia
[3] Research School of Psychology, The Australian National University, Canberra, ACT, Australia
[4] Institute of Psychology (IPSY) and Institute of Neuroscience (IoNS), Université Catholique de Louvain, Louvain-la-Neuve, Belgium
[5] Psychology Department, Università degli Studi di Milano-Bicocca, and NeuroMI - Milan Center for Neuroscience, Milano, Italy
[6] Department of Cognitive Psychology, Macquarie University, Sydney, NSW, Australia

Corresponding Author:
Associate Professor Romina Palermo
ARC Centre of Excellence in Cognition and its Disorders, and School of Psychology, University of Western Australia, 35 Stirling Highway, Crawley, WA, 6009, Australia
tel: +61 (0) 8 6488 3256
e-mail: romina.palermo@uwa.edu.au

Notes:
Jessica Irons is now at the Department of Psychology, The Ohio State University; Shahd Al-Janabi is at Department of Psychology, University of Wisconsin-Milwaukee and Davide Rivolta is at the School of Psychology, University of East London.

Running head:
Insight into face recognition ability

Acknowledgments:
This research was supported by the Australian Research Council (ARC) Centre of Excellence for Cognition and its Disorders (CE110001021) http://www.ccd.edu.au, an ARC Discovery Project grant to RP and EM (DP110100850), a ARC Queen Elizabeth II Fellowship to EM (DP0984558), an ARC Professorial Fellowship to GR (DP0877379), an ARC Discovery Outstanding Researcher Award to GR (DP130102300), and an ERC Grant (facessvep 284025) to BR.

Thank you to Ainsley Read for testing some of the participants in Study 3 and C. Ellie Wilson and Eleni Avard for testing some of the participants in Study 5.

Word count (main text): 8,175

Abstract

Diagnosis of developmental or congenital prosopagnosia (CP) involves self-report of everyday face recognition difficulties, which are corroborated with poor performance on behavioural tests. This approach requires accurate self-evaluation. We examine the extent to which typical adults have insight into their face recognition abilities across four studies involving nearly 300 participants. The studies used five tests of face recognition ability: two that tap into the ability to learn and recognise previously unfamiliar faces (the Cambridge Face Memory Test, CFMT, Duchaine & Nakayama, 2006 and a newly devised test based on the CFMT but where the study phases involve watching short movies rather than viewing static faces – the *CFMT-Films*) and three that tap face matching (Benton Facial Recognition Test, BFRT, Benton, Sivan, Hamsher, Varney, & Spreen, 1983; and two recently devised sequential face matching tests). Self-reported ability was measured with the 15-item Kennerknecht et al. (2008) questionnaire; two single-item questions assessing face recognition ability; and a new 77-item meta-cognition questionnaire). Overall, we find that adults with typical face recognition abilities have only modest insight into their ability to recognise faces on behavioural tests. In a fifth study, we assess self-reported face recognition ability in people with CP and find that some people who expect to perform poorly on behavioural tests of face recognition do indeed perform poorly. However, it is not yet clear whether individuals within this group of poor performers have greater levels of insight (i.e., into their *degree* of impairment) than those with more typical levels of performance.

Keywords: face perception; self-evaluation; prosopagnosia; individual differences; metacognition; accuracy

Recognising other people from their faces is extremely challenging because individual faces form a highly visually homogenous category. Moreover, a person's face changes all the time (e.g., viewpoint, lighting conditions, expression, ageing), making the task of individual face recognition very difficult. Despite this difficulty, for a long time it was considered that human adults were all experts at face recognition (Carey, 1992). However, recent studies have shown large individual differences in face identity recognition ability (e.g., Bowles et al., 2009; Herzmann, Danthiir, Schacht, Sommer, & Wilhelm, 2008; McKone & Palermo, 2010; Wilmer et al., 2012; Wilmer et al., 2010). Although some people, so-called super-recognisers, appear to be very good at face recognition (Russell, Duchaine, & Nakayama, 2009), there has been much more interest devoted to people who are apparently poor at recognising others by their face. By analogy to the neurological condition of prosopagnosia (face recognition impairment following brain damage, after Bodamer, 1947, see Ellis & Florence, 1990), these people have been defined as cases of developmental or congenital prosopagnosia (CP). People with CP report face recognition difficulties, often for as long as they can remember, without known and detectable brain injury and with typical visual acuity and intelligence (Bate, 2014; Behrmann & Avidan, 2005; Dalrymple & Palermo, 2016; Palermo & Duchaine, 2012; Rivolta, Palermo, & Schmalzl, 2013; Susilo & Duchaine, 2013).

Cases of CP are typically identified following their self-reports of poor face recognition abilities. Then, in many studies, their face recognition abilities are tested behaviourally, using face recognition tests, to support (or dismiss) their initial self-reports (e.g., Bate et al., 2014; Dalrymple & Palermo, 2016; Duchaine & Nakayama, 2006; Palermo, Rivolta, Wilson, & Jeffery, 2011). Common tests include tests of famous face recognition (e.g., the Macquarie Famous Face Test, MFFT, Palermo, Rivolta, et al., 2011), the Cambridge Face Memory Test (CFMT, Duchaine & Nakayama, 2006), a standardised episodic memory test that involves studying six male faces, followed by three-alternative forced choice test trials that increase in difficulty, and the

Cambridge Face Perception Test (CFPT, Duchaine, Germine, & Nakayama, 2007), in which faces are ordered for similarity to a target. However, in some studies, people have been defined as cases of CP based on self-report or semi-structured interviews only (e.g., Kennerknecht et al., 2006; Kennerknecht, Ho, & Wong, 2008). This latter way of proceeding might be problematic because there is limited evidence that self-reports of face recognition difficulty truly reflect difficulties in face recognition.

De Haan (1999) advises caution on relying on self-reports of face recognition difficulty, after finding that a member of a family in which three members show very poor recognition of familiar faces on a test showed typical face recognition performance despite noting everyday problems. De Haan states, "*It is probable that he incorrectly equated his own incidental recognition failures (which happen to all of us from time to time) with the severe recognition problems of his sisters… stresses the need for objective testing. It is hazardous to rely on subjective reports concerning face recognition difficulties of family members of developmental prosopagnosics*." (p.314). We also have experience of people reporting poor face recognition abilities yet performing at typical levels on a battery of behavioural tests (Palermo, personal correspondence, 2015). The opposite is also the case, with some people who would be classified as prosopagnosic on the basis of test scores unaware that their face recognition was poorer than that of others (e.g., Bowles et al., 2009). Similarly, Grueter and colleagues (2007, p. 746) state that; "*One of the most striking aspects of hereditary prosopagnosia is that, despite poor recognition abilities, most people are able to navigate daily life with relatively little impairment, and may even be unaware of any impairment until quite late in life*".

Kennerknecht and colleagues developed a short self-report screening measure for the presence of CP (published in Kennerknecht et al., 2008). The questionnaire contains 15 questions, which can be answered on a 5-point rating scale, resulting in scores between 15 to 75 points, with higher scores indicating more difficulty recognizing faces. Questions include; '*I can easily follow* *actors in a movie*' and '*I recognize famous people immediately*'. Eight people deemed prosopagnosic on the basis of another more detailed questionnaire and semi-structured interview reported a mean of 41.6 (SD = 4.6), whereas 186 non-prosopagnosic people (age range 18 – 25 years) reported a mean of 30.8 (SD = 6.9) (Kennerknecht et al., 2008). In a follow-up study, a significant correlation was found between scores on the questionnaire for 15 people with CP and face recognition performance (z-score of performance combined over multiple tests), $r(13) = -0.55$, $p = 0.03$ (Stollhoff, Jost, Elze, & Kennerknecht, 2011). However, it was not reported if there was a similar relationship for people who did not report long-life difficulties in face recognition.

The Kennerknecht et al. (2008) questionnaire includes questions that are not related to face identity recognition (e.g., '*I can easily form a mental picture of a red rose*'), in an attempt to exclude individuals who have other difficulties in addition to CP. In contrast, Shah, Gaule, Sowden, Bird, and Cook (2015) recently developed a questionnaire (Prosopagnosia Index, PI20) in which all 20-items are focussed on face identity recognition ability. A group of people who suspected that they had CP (and some for whom this was confirmed by poor performance on behavioural tests) scored significantly higher on the PI20 than controls, indicating that this self-report measure may tap into the everyday face recognition difficulties that are part of CP.

People with CP are not the only group poor at recognising facial identity. Adults who were deprived of early visual input for a time as children due to bilateral congenital cataracts are also slower and less accurate on tests of face recognition memory (i.e., famous face memory tests; CFMT) and their deficit has been described as "prosopagnosic-like", although they typically do not report face recognition difficulties (de Heering & Maurer, 2012). de Heering and Maurer (2012) recently developed a 10-item 7-point Likert scale "prosopagnosic questionnaire" and assessed the subjective impressions of a group of 12 people who had cataracts removed in childhood. On 8 of the 10 items, their responses were indistinguishable from those of visually-normal controls. Responses on the other two items differed from controls but inconsistently. The group reported poorer face memory on one question (‘*In general, do you have the impression of being less accurate than other people in recognizing familiar faces (family, friends, celebrities…*)’ yet better face memory on another (*‘Do you think you are very good at recognizing faces?)*.

These studies suggest that groups of people with very poor face recognition ability (e.g., CPs) may have insight into their deficit, while those with less severe difficulties (e.g., bilateral congenital cataracts) may not. Insight into face recognition abilities may be even more difficult for those within the typical ‘average’ range of performance. There are a number of reasons why it might be quite difficult for an individual to evaluate his/her own ability to learn and recognise faces. For instance, while other abilities, such as language competence, are often the subject of clear and consistent feedback in educational settings (which might be why they show one of the largest associations between self-perceptions and performance, e.g., .63, Zell & Krizan, 2014), face recognition abilities are not typically measured. Moreover, feedback in real-life may often confound face recognition ability with the ability to remember a person’s name *after* they have been recognised and/or may confound face recognition with *person* recognition (which can involve the voice, gait etc.).

Most studies have typically only used a single question to assess self-reported face recognition ability. Bowles et al. (2009) asked typical participants in their study to rate their ability to recognize faces in everyday life as “*compared to the average person*” on a 10-point scale where 0 was much worse than average and 10 was much better than average. It was clarified that the question related to the recognition of *faces* as familiar, not how well the participant remembered names. For young adults the mean rating was 6.8 (SD = 1. 6) and there was a significant, but small correlation between self-report and overall score on the CFMT, $r(113) = .22$, $p < .05$. Although in the expected direction (negative, as CFPT scores are errors so higher scores indicate poorer performance), there was no significant association between rating and performance on the CFPT, $r(27) = –.12$, n.s.).

Bindemann, Attard, and Johnston (2014) asked participants four questions (note that their responses were not combined but examined separately as single questions) – to rate their ability to recognise famous faces, family faces, unfamiliar faces seen once and unfamiliar faces seen several times on a seven-point scale from "very bad" to "very good". Self-report ratings of ability to recognise unfamiliar faces and family faces did not correlate with the later ability to recognise famous faces (r's (28) < .22, p's > .25) or correctly identify unfamiliar faces from line-ups (r's (38) < .15, p's > .10). However, self-reported ability to recognise famous faces was moderately related to later ability to recognise famous faces on two tests (r's (28) = .39 and .47, p's .03 and .0009) but not correctly identify unfamiliar faces from line-ups (r's (38) = .04, p = .82).

Rotshtein, Geng, Driver, and Dolan (2007) asked participants to rate their ability to recognize faces on a scale from 1 to 10 at the beginning of the test session, with 1 being ''*I cannot remember faces at all*'' and 10 being ''*I never forget a person's face once I met him or her*.'' Once again, it was clarified that this was not assessing their naming skills or their ability to retrieve semantic information. Using 1-tailed Spearman's rho (ρ), they found that self-reported ability did not correlate with behavioural measures of face memory performance (famous face recognition and incidental learning, ρ's (17) < .03, n.s). In a similar vein, McGugin, Richler, Herzmann, Speegle & Gauthier, 2012 (also see Gauthier et al., 2014) measured participant's self-reported *experience* (defined as "interest in, years exposure to, knowledge of, and familiarity'') with faces and other object categories, such as cars and owls) on a nine-point scales where 1 was the lowest expertise. Better performance on the CFMT was weakly associated with greater experience with faces, r(221) = .17 (p-value not reported).

Converging evidence from these studies using a single question is that typical adults have only minimal insight into their face recognition ability (with the possible exception of their ability to recognise famous faces, c.f. Bindemann et al., 2014). In contrast, the study using the recently developed PI20 questionnaire (Shah et al., 2015) suggests that people do have a great deal of insight into their face recognition abilities: scores on the PI20 correlated highly (and as expected, negatively) with scores on a famous face task ($r(171) = -.81$, $p < .001$) and the CFMT ($r(108) = -.68$, $p < .001$). This level of insight is much higher than that reported for most other abilities. For instance, a recent metasynthesis of 22 meta-analyses suggests the correspondence between self-evaluations and ability (in this case academic ability, intelligence, language competence, medical skills, sports ability, and vocational skills) average at .29 (SD = .11) (Zell & Krizan, 2014). The PI20 is a reliable measure (Cronbach's $\alpha = .96$ across a sample of typical and CP participants) and it is possible that people really do have insight into their abilities when asked a large number of very specific questions, such as *"Without hearing peoples' voices, I struggle to recognise them"*, "*I am better than most people at putting a 'name to a face'".*

However, it is important to note that the reported correlations include *both* typical individuals and those who suspect that they have CP, and the inclusion of the latter group appears to contribute greatly to the strength of the association (see Shah et al., 2015, Figure 2). Including different groups in correlational analyses sometimes occurs in the literature (e.g., Russell et al., 2009 show a correlation between the CFMT and the 'Before They Were Famous Test' but include both typical people and people with exceptionally good face recognition skills – 'super recognizers'). However, the correlation could be purely driven by mean differences between groups (e.g., Thomas et al., 2009 showed a correlation between errors in face recognition and a reduction in mean fractional anisotropy in the ventral visual stream which essentially reflected a difference between CPs and typical subjects) and thus do not add any information to a statistically different variable between groups.

Other than the recent study by Shah and colleagues (2015), most studies suggest that there is minimal, if any, relationship between one's self-evaluation of face recognition ability and behavioural measures of face recognition performance. Here, in a collective effort to address this important issue, we report the outcome of four studies, performed in different institutions, in which face recognition performance was assessed in large cohorts of *typical* participants, in parallel with self-reports of their face recognition abilities. A novel feature of this study is that we administered five tests of face recognition ability. Two of the tests are well-established: the CFMT (Duchaine & Nakayama, 2006) is a valid and reliable test of face learning and memory that is sensitive to variation in typical adults (e.g., Bowles et al., 2009; Wilmer et al., 2012) and the BFRT (Benton, Sivan, Hamsher, Varney, & Spreen, 1983) involves simultaneous face matching . Three other tests were newly developed: the CFMT-Films, which is based on the CFMT format but differs in that the study stage involves watching short film clips of people interacting (see Supplementary Materials) and two sequential face-matching tests, one in which front view faces are matched and another in which the faces vary in viewpoint from target to test. All of the tests used faces that were previously unfamiliar to participants. This is important since the tests involve matching (with or without delay) and old/new recognition, which could be performed using nonvisual codes (e.g., semantic information or labels) if using familiar faces (see reviews by Burton & Jenkins, 2011; Johnston & Edmonds, 2009 for differences in processing of unfamiliar and familiar faces). We also included the Cambridge Car Memory Test (CCMT) (Dennett et al., 2012) as a test of non-face recognition memory, to examine whether relationships were specific to faces or to visual memory in general.

Another novel feature of this study was that we used a variety of self-report measures of face recognition ability: the 15-item Kennerknecht et al. (2008) questionnaire; two single-item questions assessing overall face recognition ability with a 9-point likert scale (*How well do you think you will perform on studying and recognising faces from your own “race”?"* and "*Overall, how would you describe your general ability to recognise faces*?); and a new 77-item questionnaire of face recognition metacognition (knowledge about mental processes). We also included measures of social anxiety (Social Interaction Anxiety Scale, SIAS, Mattick & Clarke, 1998) and autistic traits (Autism-Spectrum Quotient, AQ, Baron-Cohen, Wheelwright, Skinner, Martin, &

Clubley, 2001) to investigate whether test scores and/or self-reported face recognition ability are associated with other traits.

Finally, we also report a separate study with people diagnosed with CP on the basis of behavioural tests and anecdotal self-report of everyday face recognition difficulties and examine whether very poor performance is linked with increased insight (as compared to those with typical range face recognition ability). This combination of measures and participant groups will provide a comprehensive analysis of people's insight into their face recognition abilities.

Methods

Tests

In the upright version of the *Cambridge Face Memory Test (CFMT)* (Duchaine & Nakayama, 2006), participants study six greyscale male target faces that have been cropped to remove non-face cues and then select the studied face from two distractors. The test section consists of three stages that increase in difficulty, for a total score out of 72. In the first 'learn' stage the faces are the same as those studied (score out of 18), in the second 'novel' stage the faces are seen under novel lighting and viewpoints (score out of 30), and in the third 'noise' stage, visual noise is overlaid on all the faces (score out of 24). The CFMT has been validated on the basis of its ability to diagnose people with acquired prosopagnosia (Liu-Shuang, Torfs & Rossion, in press; Susilo, Yovel, Barton, & Duchaine, 2013) and the much lower scores observed for inverted as compared to upright faces (Duchaine & Nakayama, 2006). Also, performance on the CFMT correlates with performance on a face matching test without any learning component (i.e., CFPT) (Bowles et al., 2009; Duchaine et al., 2007) but the test displays only modest correlations with measures of non-face visual memory (Dennett et al., 2012; Wilmer et al., 2010) and even weaker correlations with measures of verbal memory (Wilmer et al., 2012). The CFMT is also reliable, as measured with

Cronbach's alpha, a measure of internal consistency (α = .89, Bowles et al., 2009; α = .83, Herzmann, et al., 2008; α = .90, Wilmer et al., 2010).

The *CFMT-Films* is a newly devised test (see Supplementary Materials for details of the test development and psychometric data). This test is based on the CFMT format but the study stage involves watching short film clips of people interacting. The test section contains two stages (38 trials each) in which participants select previously seen faces from a line-up of three faces: a 'novel' stage and a 'noise' stage in which Gaussian noise was overlaid on the faces (76 trials in total). Analyses conducted with a sample of 89 participants indicate that the CFMT-Films displays excellent reliability (α = .89) and validity (correlates highly with the CFMT but only weakly with non-face visual memory), that average performance is at the psychometric 'sweet spot' mid-way between chance and ceiling (M = 68.57%) and that individual performance displays sufficient range to measure individual differences (SD = 14.64).

The Benton Facial Recognition Test (BFRT, Benton et al., 1983) is a classical test to assess face recognition impairments in brain-damaged patients. It involves the matching of a target face to either one face under the same viewpoint and lighting (6 items) or three of six faces that vary in viewpoint and lighting (16 items). All faces are presented simultaneously. The maximum score is 54, with a score between 41 and 39 considered as mildly impaired, and below 39 as severely impaired. Note that despite the difficulty of this test, cases of patients with acquired prosopagnosia (e.g., Delvenne, Seron, Coyette, & Rossion, 2004; Busigny & Rossion, 2010) or CP (Duchaine & Nakayama, 2004) can sometimes reach almost normal performance at simultaneous matching by using unusual strategies and taking abnormally long response times (see Delvenne et al., 2004; Busigny & Rossion, 2010 for discussion of this issue and the consideration of RT measures in the BFRT).

Study 4 also included *two matching tests*. In these tests, participants first see a target face presented for 200 ms followed by a brief mask (250ms) and then four faces arranged in a square around the centre of the screen, which remain visible until the participant's response. In the first task (4AFC-FF), all the faces are seen from front-on (exact same image at encoding and test). In the second task (4AFC-DR), the target face is rotated 30° to the left or right, and the test faces all face the opposite direction. All faces are presented in colour, and external features are cropped. Each test contains 56 face identities (28 males and 28 females). The faces are about 7 cm high and 5 cm wide on the screen at a distance of 70 cm from the screen. Each face is repeated four times, appearing once at each possible location on the screen. Participants are asked to answer as quickly and as accurately as possible, using both hands to press the keys numbered 1, 3, 4 or 6. The split-half reliability of these two delayed face matching tests is very high (.93 and .94 respectively, Table 1), and three well-described patients with acquired prosopagnosia (GG, Busigny, Joubert, Felician, Ceccaldi, & Rossion, 2010; LR, Busigny et al., 2014; PS, Rossion et al., 2003) perform very poorly on both tests (Laguesse & Rossion, unpublished data). PS's (scores of 46; 54, respectively) and GG's (47.3, 42.4) scores are well below 2 SDs of the mean (limits: 65.1; 61), while LR's low score for the first test was only was slightly above 2 SDs below normal range (69.6; 55.8). However, LR (as PS)'s correct RTs were more than 5 SDs below normal range. The *Cambridge Car Memory Test (CCMT)* (Dennett et al., 2012) was also included as a test of non-face recognition memory. This test is modelled on the CFMT, except that images of cars instead of faces are used. The reliability of the test is .84 (Dennett et al., 2012).

Questionnaires

The questionnaire developed by Kennerknecht et al. (2008) is used as a screening measure for the presence of CP, with higher scores indicating more difficulty recognising faces. Thus, a relationship between self-report ratings of face recognition ability and behavioural face recognition performance in this study would be seen via significant negative correlations.

We also included single questions assessing how well participants thought that they would perform at recognizing faces of their own race: "*How well do you think you will perform on studying and recognising faces from your own "race"?"* and their overall face recognition ability: "*Overall, how would you describe your general ability to recognise faces*?" Participants responded using a 9-point Likert scale ranging from 1 ("very poor") to 9 ("very good"). Thus, a relationship between ratings of face recognition ability and behavioural face recognition performance in this study would be seen via significant positive correlations.

A new questionnaire of *metacognition about face recognition* was also included. The questionnaire contains a total of 77 questions organised into various subcategories, with many questions worded in both positive and negative forms to be able to check for consistency (see Supplementary Materials for the French to English translation of the questionnaire).

The *Social Interaction Anxiety Scale* (SIAS) (Mattick & Clarke, 1998), which measures self-reported levels of social anxiety, and the *Autism-Spectrum Quotient (AQ)* (Baron-Cohen et al., 2001) which measures self-reported levels of autistic traits, were also included. These questionnaires allow us to investigate whether test scores and/or self-reported face recognition ability are also associated with social anxiety and/or autistic traits.

Participants and procedures for each study

Study 1

The Kennerknecht et al. (2008) questionnaire (translated into Italian) was administered to 490 Psychology students at the University of Milan-Bicocca. Ninety-six of these (16 males), aged between 19 and 28 years (M = 21.69, SD = 1.98), volunteered to return at a later time and complete computerised tests for course credit. All had normal or corrected to normal vision and no evidence of neurological deficit. These 96 participants completed the CFMT, both upright and inverted, and the BFRT (Benton et al., 1983). The order of test administration was counterbalanced.

Study 2

Eighty-nine adults of European descent (28 males), aged between 17 and 46 years (M = 21.44, SD = 5.64), with normal or corrected-to-normal vision, were tested in an individual session. Most participants were students at the Australian National University, recruited via flyers and received course credit or $30 for participating in a 2-hour session.

Participants completed four computerized tests on a 24 in iMac with a resolution of 1920 × 1200 pixels with brightness and contrast set to maximum. The tests were completed in the following order: the newly developed *CFMT-Films* (Supplementary Materials), a composite face test (not discussed here), the CFMT (Duchaine & Nakayama, 2006) and the CCMT (Dennett et al., 2012). The CFMT and CCMT were administered following the standard instructions. Ten participants who had already completed the CFMT and/or CCMT as part of another study did not re-take these tests and we used their score from this previous study. At the end of the session participants completed a pencil and paper demographic questionnaire, the face recognition questionnaire of Kennerknecht et al. (2008), and the SIAS (Mattick & Clarke, 1998).

Study 3

Fifty-seven adults of European descent (17 males), aged between 17 and 32 years (M = 18.77, SD = 2.57) were tested in an individual session. Most participants were students at the University of Western Australia recruited via an online sign-up system and received course credit or $20 for participating in a 2-hour session. For reasons not related to this study, some of the students were specifically invited to participate on the basis of their Autism-Spectrum Quotient (AQ) scores (Baron-Cohen et al., 2001), a questionnaire which was administered to the 1st year Psychology student pool in the weeks prior to testing. Nine students were selected for low scores

(11 or less) and ten for high scores (23 and over). No participant scored 32 or above, which is the cut-off indicative of a possible autism spectrum disorder (Baron-Cohen et al., 2001).

Participants completed two test sessions. In the first, they completed four computerized tasks on a 20 in iMac with a resolution of 1680 x 1050 pixels. The tasks were completed in the following order: an identity aftereffect task (not discussed here), the CFMT, the CCMT and a Car Makes and Models check. In this study the CFMT was re-programmed into SuperLab, and the stimuli in the Novel and Noise phases were randomly presented, rather than in the fixed order used in Studies 1 and 2 (note that due to a coding error, one of the trials in the Noise condition was repeated and the trial that should have been included was omitted). In the second session, exactly one week after the first session, participants completed two computerised tasks not discussed here – a face aftereffects task and a cross-race face recognition task. Prior to beginning the cross-race face recognition task, participants were asked to rate how they thought they would perform at recognizing faces of their own race ("*How well do you think you will perform on studying and recognising faces from your own race?"),* a different race ("*How well do you think you will perform on studying and recognising faces from a different race?*"; not discussed here as our interest is not with other-race face recognition), as well as rate their general ability to recognize faces *("Overall, how would you describe your general ability to recognise faces?"*). Participants responded using the keyboard with a 9-point Likert scale ranging from 1 (“very poor”) to 9 (“very good”). After completing the computer tasks the participants completed paper and pencil versions of the AQ, a demographic questionnaire, the SIAS, and then the Kennerknecht et al. (2008) questionnaire.

Study 4

Fifty-eight adults of European descent (9 males), aged between 18 and 25 years (M = 19, SD = 1.47) were tested in an individual session. Most participants were undergraduate students at the University of Louvain and received credit courses in exchange to their participation. Participants were asked to complete the Metacognition questionnaire, and then the CFMT and the two 4AFC tasks (FF and DR). The tasks were run on a PC with a 19 inches screen of 1680 x 1050 pixels of resolution.

Study 5

In this study we tested 13 people (4 males), aged between 23 and 60 years (M = 43.46, SD = 12.08) who registered with the online Australian Prosopagnosia Register (https://www.cogsci.mq.edu.au/research/projects/prosopagnosia/register/) because they experience significant everyday face recognition difficulties. Each individual completed several tests of face recognition: a test of famous face recognition (the MFFT 2008; Palermo, Rivolta et al., 2011), the CFMT (Duchaine & Nakayama, 2006), a version of the CFMT using faces well-matched to those typically seen in Australia (the CFMT-Australian, McKone et al., 2011) and the CFPT (Duchaine et al., 2007). Individuals were impaired across multiple tests, displaying performance at least 1.7 SDs below typical face recognisers on at least two tests (with the exception of F23 who was below this limit only one test. However, she was close to the limit on an additional two tests, and as such was included as a CP due to general overall pattern of poor performance). Most of the individuals displayed deficits that were relatively face-specific (i.e., not impaired on the CCMT, Dennett et al., 2012), with the exception of M57 who was also impaired at recognising cars. Given that his performance on the Raven's Progressive Matrices (Raven, 2008) did not indicate substantial general cognitive deficits we opted to keep him in our sample (see Bate, Haslam, Tree, & Hodgson; 2008; Dalrymple & Palermo, 2016 for discussion of diagnosis of CP).

Participants were provided with feedback on their poor face recognition performance at different points of the study. The first five people in Table 2 were given feedback prior to completing the Kennerknecht et al. (2008) questionnaire, whereas the other eight participants completed the face recognition tests and the questionnaires in the same test session, prior to receiving feedback on their performance.

Results

Table 1 displays descriptive statistics for the four studies with typical individuals. These include tests of face identity recognition (CFMT, CFMT-Films, BFRT, 4AFC-FF and –DR Matching), and in two studies, a test of object recognition (CCMT). Importantly, average performance was neither at floor or ceiling and all show sufficient range. Descriptive statistics are also shown for the face recognition ability questionnaires (Kennerknecht et al., 2008; own-race and general single item questionnaire, metacognition questionnaire) and those that measured levels of social anxiety (SIAS) and autistic traits (AQ). Cronbach's alpha, a measure of internal consistency, was average to high for all of the measures. Cronbach's alpha values were also used to calculate the theoretical upper bound (UB) of a correlation that could be obtained between two tests; calculated as the geometric mean of the two reliabilities (Kaplan & Saccuzzo, 2001).

When more than one face test was included in a study we first examine the relationships between the tests to determine whether they were measuring similar aspects of face recognition. If so, we would expect that self-report measures should correlate with both tests similarly. We then turn to our primary question: the relationship between test score and questionnaire self-report. We report both Pearson's r and Spearman's Rho (ρ, and in the case of partial correlations Kendall's Tau, pτ), as in some cases the distributions were not normal, or suffered from significant skew or kurtosis (Table 1).

1. Are scores on the Kennerknecht et al. (2008) questionnaire related to performance on the tests of face recognition ability?

Higher scores on the Kennerknecht et al. questionnaire indicate more difficulties with face recognition. Thus, a negative correlation indicates a relationship between self-report scores and ability. A large sample of participants completed both the CFMT and the Kennerknecht et al. questionnaire (N = 240, Studies 1-3), and there was a significant, yet small, relationship between the measures, r(238) = -.14, p = .03, UB = .86, ρ = -.12, p = .08. Given the large sample we also examined whether there were any sex differences (180 females; 60 males). The correlation was numerically smaller for females (r(178) = -.08, p = .27; ρ = -.11, p = .16) than for males (r(58) = -.28, p = .03; ρ = -.17, p = .21) but the difference was not significant (z = 1.36, p = .17).

These studies all tested unselected university students. Given estimates that 2.5% of the general population perform at prosopagnosic levels (see Bowles et al., 2009; Kennerknecht et al., 2006), we also examined correlations between the Kennerknecht et al. questionnaire and the CFMT excluding those who could "potentially" be prosopagnosic. Excluding the poorest 2 percent of participants (n = 5) did not change the strength of the correlation, r(233) = -.14, p = .03; ρ = -11, p = .10.

Some of these participants also completed the CCMT (Studies 2 and 3), as a measure of non-face visual memory well-matched in format to the CFMT (Dennett et al., 2012). The Kennerknecht questionnaire did not correlate with non-face visual memory (r(143) = -.08, p = .31; ρ = -.05, p = .55, UB = .83) and controlling for CCMT strengthened the correlation between the CFMT and the questionnaire, pr(141) = -.25, p = .003, pτ(141) = -.15, p = .009. Thus, the relationship with the Kennerknecht questionnaire appears to be specific to face recognition ability rather than general visual memory.

Davis et al. (2011) reported a small significant negative relationship between social anxiety (measured with the SIAS) and CFMT performance. Similarly, in this study we found a small but significant relationship between SIAS and CFMT, r(136) = -.22, p = .008, ρ = -.18, p = .03 (Studies 2 and 3). Additionally, here we show that scores on the Kennerknecht questionnaire also correlate with social anxiety, in that people with higher self-reported levels of social anxiety tended to report that they had more difficulty recognizing faces, r(137) = .30, p = .001, UB = .86, ρ = .28, p =

.001. Note that the association between the CFMT and the Kennerknecht questionnaire is still apparent when the SIAS was controlled for, $pr(135) = -.18$, $p = .04$; $p\tau(135) = -.11$, $p = .05$.

The Kennerknecht questionnaire also correlated with self-reported autistic traits (Study 3), in that people who scored more highly on the AQ tended to report that they had more difficulty recognizing faces, $r(55) = .41$, $p = .002$, UB = .80; $\rho = .39$, $p = .003$ (and this was reflected in performance, with a negative correlation between the CFMT and AQ, $r(55) = -.46$, $p = .001$; $\rho = -.38$, $p = .003$). These correlations are of a similar magnitude to the correlations between the CFMT and Kennerknecht questionnaire (i.e., AQ is as good as a predictor as the Kennerknecht questionnaire). The correlation between the Kennerknecht questionnaire and the CFMT is weaker after controlling for AQ, $pr(54) = -.25$, $p = .07$; $p\tau(54) = -.23$, $p = .01$.

Participants in Study 2 completed both the CFMT and the CFMT-Films. Performance on these tests was highly correlated, $r(86) = .64$, $p < .001$, UB = .89; $\rho = .65$, $p < .001$ (see Supplementary Materials for more details). However, scores on the CFMT-Films were not associated with the Kennerknecht et al. questionnaire, $r(86) = -.08$, $p = .46$, UB = .86; $\rho = -.01$, $p = .93$.

The CFMT was also moderately correlated with the BFRT (Study 1), $r(94) = .49$, $p < .001$; UB = .76; $\rho = .50$, $p < .001$. However, once again, scores on the BFRT were not associated with the Kennerknecht et al. questionnaire, $r(94) = -.00$, $p = .99$, UB = .77; $\rho = .00$, $p = .97$.

2. Are scores on the single-item questions related to performance on the tests of face recognition ability?

In Study 3, participants also completed two single-item questions assessing the self-reported ability to recognise own-race faces and faces in general. A positive relationship would indicate some insight into performance. However, the ability to recognise faces on the CFMT was not significantly correlated with either single-item self-reported own-race ($r(55) = .13$, $p = .35$; $\rho = .07$, $p = .61$) or general face recognition ability ($r(55) = .26$, $p = .05$; $\rho = .22$, $p = .10$).

Interestingly, responses on the Kennerknecht questionnaire were not significantly associated with the single-item self-report measure of own-race face recognition ($r = -.07$, $p = .63$; $\rho = -.08$, $p = .57$) and were only marginally associated with the single-item self-report measure of general face recognition ($r = -.23$, $p = .09$; $\rho = -.26$, $p = .053$).

3. Are scores on the meta-cognition questionnaire related to performance on the tests of face recognition ability?

The CFMT and the 4AFC-FF and –DR matching tasks were highly positively correlated, FF: $r = .68$, UB = .91, $\rho = .68$; DR: $r = .77$, UB = .91, $\rho = .78$, all p's = .001. Moreover, the two 4AFC tests were highly correlated despite the change in viewpoint in the latter task, both for accuracy ($r = .85$, $p = .001$, $\rho = .79$, $p = .001$) and correct RTs ($r = .73$, $p = .001$, $\rho = .77$, $p = .001$).

A positive relationship between the face recognition tests and the metacognition questionnaire is expected if the self-report measure is related to performance. A significant positive (small) relationship was observed for both the CFMT ($r(56) = .32$, $p < .05$, UB = .87; $\rho = .37$, $p < .01$) and the matching tests (4AFC-FF: $r(56) = .29$, $p < .05$, UB = .89; $\rho = .35$; 4AFC-DR: $r(56) = .32$, $p < .05$, UB = .90; $\rho = .40$, $p < .01$).

4. Do people with developmental/congenital prosopagnosia (CP) self-report face recognition difficulties?

Here, we examine how people who report everyday face recognition difficulties and perform poorly on tests of face recognition ability respond on self-report questionnaires. Note that comparing CPs and typical face recognisers is valuable, regardless of whether CP reflects the low end of continuous variation in face recognition ability or a distinct group (see Bowles et al, 2009 for discussion).

The average age of the 13 CPs (M = 43.5, SD = 12.1) was greater than that of the people in Studies 1-3. Bowles et al. (2009) recommend using age-adjusted z-scores for the CFMT and age- and sex-adjusted z-scores for the CFPT for those aged middle-aged or older. Thus, Table 2 includes both raw scores adjusted z-scores. As is evident, the CPs performed poorly on the face recognition tests – the MFFT 2008, the CFMT and the CFPT.

We have no reason to expect that responses on the Kennerknecht questionnaire would vary within that age-range and so compared the ratings for the CPs (M = 46.7 SD = 8.2) with those of the undergraduates (M = 29.1, SD = 7.3), and found that those of the CPs were significantly higher $t(252) = 8.38$, $p < .0001$.

AQ scores did not differ between the CPs (M = 19.1, SD = 8.6) and undergraduates (M = 15.7, SD = 6.7), $t(68) = 1.57$, $p = .12$. This replicates previous research with age-matched samples, which has not found any evidence of elevated AQ scores in groups of people with CP (Duchaine, Murray, Turner, White, & Garrido, 2009; Palermo, Willis, et al., 2011).

Social anxiety is more likely to be diagnosed in adolescence or young adults than in older adults (Kessler et al., 2005). However, given other evidence that age is not related to SIAS scores (Brown et al., 1997) we also compared the scores made by the group of CPs to the undergraduates. We expected to find a significant difference in light of the study by Yardley and colleagues (2008), where interviews suggested that many people with CP reported psychosocial difficulties, including social anxiety. However, while the SIAS scores were numerically higher for CPs (M = 33.15, SD = 16.05) than undergraduates (M = 28.31, SD = 14.80), they did not differ significantly, $t(150) = 1.12$, $p = .26$. Scores on the SIAS above 36 are indicative of social phobia (Peters, 2000) and 4 of the 13 CPs had scores above 36. This rate of 31% is higher than that of the Australian population reporting any anxiety disorder (14.4%) (The Mental health of Australians 2:

report on the 2007 national survey of mental health and wellbeing). However, we note here that the mean of the undergraduate sample is also relatively high (c.f. M = 19.0, SD = 10.1 for undergraduates in Mattick & Clarke, 1998; M = 22.61, SD = 12.69 in Davis et al., 2011) and 29% of the group had scores above 36, which suggests that our CPs could show significantly higher levels of social anxiety when compared to the typical population.

Finally, to test our assumption that the magnitude of the correlation in the study by Shah et al. (2015) was inflated by mixing CPs and those with typical face recognition skills, we added the 13 CPs to the data collected in Studies 1-3. The strength of the correlation between the CFMT and Kennerknecht et al. (2008) questionnaire significantly increased, $r(253) = -.31$, $p = .001$, $\rho = -.22$, $p = .001$ ($z = 1.98$, $p = .03$).

Discussion

To investigate how much insight typical individuals have in their ability to recognise faces collected data from nearly 300 participants, who completed five face recognition tests and three questionnaires Overall, our results suggest that individuals in the general population have only minimal to moderate insight into their face recognition abilities. We also found that longer questionnaires seem better able to tap into insight than shorter questionnaires, particularly single-item questions.

Across a general undergraduate sample unselected for face recognition ability we observed only a small, albeit significant, relationship between the 15-item Kennerknecht et al. (2008) questionnaire and the CFMT ($r = -.14$) yet no significant relationships between this questionnaire and either the CFMT-Films or the BFRT. The observed small relationship with the CFMT appears selective to faces rather than general visual memory, as there was no relationship between this questionnaire and a non-face version of the CFMT, the CCMT. In sum, results from the

Kennerknecht et al. (2008) questionnaire reveal only minimal insight into face recognition ability in the typical population.

The group of people with CP scored more highly on the Kennerknecht et al. (2008) questionnaire than our undergraduate comparison group. This difference supports the recent finding that people with very poor face recognition have insight into their (in)abilities (Shah et al., 2015), and/or that the questionnaire may tap into their specific deficits (which would not be surprising given that is what it was designed for). However, it is important to note that our data only indicate that this group know they are poorer relative to others, not that those *within* this group of people are aware of the relative severity of their impairment within the group of CPs. In our sample of 13 CPs there was no evidence of any correlation between their score on tests and their self-report score on the Kennerknecht et al. questionnaire. However, a correlation may be evident with a larger sample size (c.f., Shah et al. Figure 2; Stollhoff et al., 2011). It would be of interest for future studies to examine whether insight differs between those who are very poor (and very good) as compared to those who perform with normal limits. An additional question of interest would be to examine self-reported face recognition ability with that for other categories of objects in people with CP. On the one hand, it has been suggested that the relationship between self-report and recognition may be higher for faces than other non-face objects (Gauthier et al., 2014; McGugin et al., 2012; Van Gulick, McGugin, & Gauthier, in press). On the other hand, self-reported experts at non-face object recognition, car experts for instance, score much higher than self-reported novices at independent tasks only for the category of self-reported expertise (Rossion et al., 2007; Rossion & Curran, 2010), thus rather showing a good insight at their non-face object recognition ability. Assessing verbal semantic knowledge may also be valuable, given that that verbal semantic knowledge of car makes was a better predictor of visual recognition of cars than self-reported car expertise in typical individuals and that most people with acquired prosopagnosia did not show typical levels of car recognition once verbal semantic scores were adjusted for (Barton, Hanif, &Ashraf, 2009; see also Barton & Corrow, 2016).

The validity of the Kennerknecht et al. (2008) questionnaire to tap insight into face recognition ability has been previously criticized (e.g., Shah et al., 2015), because it includes some items that are unrelated to the core face recognition impairment in CP. When we used a newly devised 77-item novel questionnaire that includes questions from everyday life situations where impairments in face recognition should appear, we found much larger correlations with the CFMT and face-matching test (~.30). In contrast, ratings on the single-items asking about own-race and general face recognition ability were poor predictors of face recognition ability and were not related to CFMT performance. These results suggest that the new, longer questionnaire is better at tapping into the kinds of face recognition difficulties people are aware of.

In general, we found that the size of relationship between self-report ratings and tests of face recognition ability is commensurate with that seen for other types of memory. For instance, a meta-analysis has shown a small ($r = .15$), yet significant, relationship between memory self-efficacy and performance across 107 studies (Beaudoin & Desrichard, 2011). Our conclusions contrast with Shah et al. (2015), who have argued that people have a great deal of insight into their own face recognition abilities, based on correlations using their 20-item scale with the CFMT ($r = -.68$, $N = 173$) and a famous face test ($r = -.81$, $N = 110$). As noted earlier, their sample includes both people who suspect that they have CP ($n = 23$) and those who do not ($n = 87$). We suggest that combining the groups in this way is likely to have increased the strength of their correlation. In support of this argument, when we added the data from 13 people with CP to the data from the 240 typical individuals, the strength of the correlation between the Kennerknecht et al. (2008) questionnaire and the CFMT doubled in magnitude.

Some participants also completed the AQ, a self-report measure of autistic traits. Performance on the CFMT was negatively correlated with the AQ, which has been previously reported (Halliday, MacDonald, Sherf, & Tanaka, 2014; but less clearly in Rhodes, Jeffery, Taylor, & Ewing, 2013). Interestingly, the strength of the relationship between the CFMT and AQ (r = -.46, upperbound = .84) is of a similar magnitude to that between the CFMT and Kennerknecht et al. questionnaire (r = -.39, upper bound = .83). There was also a moderate correlation between the Kennerknecht et al. questionnaire and the AQ, with more self-rated autistic traits correlated with poorer self-reported face recognition ability.

Some participants also completed the SIAS, a self-report measure of social anxiety. Once again, performance on the CFMT was negatively correlated with SIAS scores (as in Davis et al., 2011) and the strength of this relationship (r = -.22, upperbound = .89) is similar to that between the CFMT and Kennerknecht et al. questionnaire (r = -.24, upperbound = .86). There was also a moderate correlation between the Kennerknecht et al. questionnaire and the SIAS, with higher levels of social anxiety associated with poorer self-reported face recognition ability. These patterns of results suggest that self-reported autistic traits/social anxiety is as predictive of performance on face recognition tests as self-reported face recognition ability, at least across those without levels of AQ indicative of autism (none of the participants scored above the cut-off indicative of autism as specified by Baron-Cohen et al., 2001). Note that we do not wish to argue that the AQ is a good proxy for face recognition self-evaluation, given that people with prosopagnosia do not typically score highly on the AQ – i.e., AQ scores in Study 5 did not differ between CPs and controls (also see Duchaine et al., 2009; Palermo, Willis et al., 2011). Similarly, while people with prosopagnosia often report social anxiety (Yardley et al., 2008), the group of CPs in Study 5 did not report significantly higher levels of social anxiety on the SIAS. For the CPs, the Kennerknecht et al. questionnaire was more useful than either the AQ or the SIAS.

We note that the first four studies were conducted with young adults. Face recognition ability as measured behaviourally peaks at around age 30 and then slowly declines (Susilo, Germine, & Duchaine, 2013). Bowles et al. (2009) reported a small (r= .22) but significant relationship between self-reported face recognition and CFMT score for young adults aged 18-35 years but there was no association for those aged 55-88 years (r = -.05). An interesting avenue for future research will be to examine whether age affects insight into face recognition ability.

In summary, the ability of self-report questionnaires to measure insight into face recognition ability, even when the items concentrate on face recognition (e.g., Study 4), appears limited. However, this conclusion is only valid if the behavioural tests used truly reflect daily-life face recognition abilities. Regarding this point, many of the tests used have been shown to measure both individual differences in ability (e.g., the CFMT, Wilmer et al., 2012) and are used to diagnose CP (e.g., the CFMT, Duchaine & Nakayama, 2006). Importantly, these tests also appear to identify cases of acquired prosopagnosia (e.g., CFMT, Liu-Shuang et al., in press; Susilo et al., 2013; 4AFC tests used in Study 4, Busigny et al., 2010, 2014; Rossion et al., 2003). People with very poor face recognition skills may be more aware of their difficulties than the typical population, but such insight is clearly not universal. The studies reported here have important implications for the diagnosis of CP, which relies upon poor performance on behavioural tests *and* accounts of everyday face recognition difficulties. It also goes some way to explaining why not all individuals with poor face recognition test scores would expect that they would perform poorly.

References

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of autism and developmental disorders, 31*(1), 5-17.

Barton, J. J. S., & Corrow, S. L. Selectivity in acquired prosopagnosia: The segregation of divergent and convergent operations. *Neuropsychologia*. doi:http://dx.doi.org/10.1016/j.neuropsychologia.2015.09.015

Barton, J.J.S., Hanif, H., Ashraf, S. (2009). Relating visual to verbal semantic knowledge: the evaluation of object recognition in prosopagnosia. *Brain*, *132*, 3456-3466; DOI: 10.1093/brain/awp252

Bate, S. (2014). Face facts: Understanding prosopagnosia. *Psychology Review, 19*, 18-21.

Bate, S., Cook, S. J., Duchaine, B., Tree, J. J., Burns, E. J., & Hodgson, T. L. (2014). Intranasal inhalation of oxytocin improves face processing in developmental prosopagnosia. *Cortex, 50*, 55-63. doi:10.1016/j.cortex.2013.08.006

Bate, S., Haslam, C., Tree, J. J., & Hodgson, T. L. (2008). Evidence of an eye movement-based memory effect in congenital prosopagnosia. *Cortex, 44*(7), 806-819.

Beaudoin, M., & Desrichard, O. (2011). Are memory self-efficacy and memory performance related? A meta-analysis. *Psychol Bull, 137*(2), 211-241. doi:10.1037/a0022106

Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: face-blind from birth. *Trends Cogn Sci, 9*(4), 180-187. doi:10.1016/j.tics.2005.02.011

Benton, A., Sivan, A., Hamsher, K., Varney, N., & Spreen, O. (1983). *Contribution to neuropsychological assessment*. New York: Oxford University Press.

Bindemann, M., Attard, J., & Johnston, R. A. (2014). Perceived ability and actual recognition accuracy for unfamiliar and famous faces. *Cogent Psychology, 1*(1), 986903. doi:10.1080/23311908.2014.986903

Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., . . . Yovel, G. (2009). Diagnosing prosopagnosia: effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. *Cogn Neuropsychol, 26*(5), 423-455. doi:10.1080/02643290903343149

Brown, E.J., Turovsky, J., Heimberg, R.G., Juster, H.R., Brown, T.A., & Barlow, D.H. (1997) Validation of the Social Interaction Anxiety Scale and the Social Phobia Scale Across the Anxiety Disorders. *Psychological Assessment, 9*: 21–27. doi:10.1037/ 1040-3590.9.1.21.

Burton, A.M. & Jenkins, R. (2011). Unfamiliar face perception. In A.J. Calder, G. Rhodes, M.H. Johnson & J. Haxby (Eds.). *The Oxford Handbook of Face Perception.* pp 287-306, Oxford University Press.

Busigny, T., Joubert, S., Felician, O., Ceccaldi, M., & Rossion, B. (2010). Holistic perception of the individual face is specific and necessary: evidence from an extensive case study of acquired prosopagnosia. *Neuropsychologia, 48*(14), 4057-4092. doi:10.1016/j.neuropsychologia.2010.09.017

Busigny, T. & Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*, 46, 965-981. doi:10.1016/j.cortex.2009.07.004

Busigny, T., Van Belle, G., Jemel, B., Hosein, A., Joubert, S., & Rossion, B. (2014). Face-specific impairment in holistic perception following focal lesion of the right anterior temporal lobe. *Neuropsychologia, 56*, 312-333. doi:http://dx.doi.org/10.1016/j.neuropsychologia.2014.01.018

Carey, S. (1992). Becoming a face expert. *Philos Trans R Soc Lond B Biol Sci, 335*(1273), 95-102; discussion 102-103. doi:10.1098/rstb.1992.0012

Dalrymple, K. A., & Palermo, R. (2016). Guidelines for studying developmental prosopagnosia in adults and children. *Wiley Interdisciplinary Reviews: Cognitive Science, 7,* 73–87. doi:10.1002/wcs.1374

Davis, J. M., McKone, E., Dennett, H., O'Connor, K. L., O'Kearney, R., & Palermo, R. (2011). Individual differences in face recognition ability are associated with social anxiety. *PLoS ONE, 12*, e28800. doi:10.1371/journal.pone.0028800.

De Haan, E. H. (1999). A familial factor in the development of face recognition deficits. *J Clin Exp Neuropsychol, 21*(3), 312-315. doi:10.1076/jcen.21.3.312.917

de Heering, A., & Maurer, D. (2012). Face memory deficits in patients deprived of early visual input by bilateral congenital cataracts. *Developmental Psychobiology*. doi:10.1002/dev.21094

Delvenne, J.-F., Seron, X., Coyette, F., Rossion, B. (2004). Evidence for perceptual deficits in associative visual (prosop)agnosia: a single-case study. *Neuropsychologia, 42*, 597-612. doi:http://dx.doi.org/10.1016/j.neuropsychologia.2003.10.008

Dennett, H. W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., & Duchaine, B. (2012). The Cambridge Car Memory Test: a task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. *Behav Res Methods, 44*(2), 587-605. doi:10.3758/s13428-011-0160-2

Duchaine, B., Germine, L., & Nakayama, K. (2007). Family resemblance: ten family members with prosopagnosia and within-class object agnosia. *Cogn Neuropsychol, 24*(4), 419-430. doi:10.1080/02643290701380491

Duchaine, B., Murray, H., Turner, M., White, S., & Garrido, L. (2009). Normal social cognition in developmental prosopagnosia. *Cogn Neuropsychol, 26*(7), 620-634. doi:10.1080/02643291003616145

Duchaine, B. C., & Nakayama, K. (2004). Developmental prosopagnosia and the Benton Facial Recognition Test. *Neurology, 62*(7), 1219-1220. doi: http://dx.doi.org/10.1212/01.WNL.0000118297.03161.B3

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

Ellis, H. D., & Florence, M. (1990). Bodamer's (1947) paper on prosopagnosia. *Cogn Neuropsychol, 7*(2), 81-105. doi:10.1080/02643299008253437

Gauthier, I., McGugin, R.W., Richler, J.J., Herzmann, G., Speegle, M. & VanGulick, A.E. (2014). Experience moderates overlap between object and face recognition, suggesting a common ability, *Journal of Vision, 14:7,* doi: 10.1167/14.8.7.

Grueter, M., Grueter, T., Bell, V., Horst, J., Laskowski, W., Sperling, K., . . . Kennerknecht, I. (2007). Hereditary prosopagnosia: the first case series. *Cortex, 43*(6), 734-749.

Halliday, D. W. R., MacDonald, S. W. S., Sherf, S. K., & Tanaka, J. W. (2014). A Reciprocal Model of Face Recognition and Autistic Traits: Evidence from an Individual Differences Perspective. *PLoS ONE, 9*(5), e94013. doi:10.1371/journal.pone.0094013

Herzmann, G., Danthiir, V., Schacht, A., Sommer, W., & Wilhelm, O. (2008). Toward a comprehensive test battery for face cognition: Assessment of the tasks. *Behavior Research Methods, 40*(3), 840-857. doi:10.3758/BRM.40.3.840

Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: a review. *Memory, 17*(5), 577-596. doi:10.1080/09658210902976969

Kaplan, R. M., & Saccuzzo, D. P. (2001). *Psychological Testing: Principles, Applications and Issues* (3rd ed.). California: Brooks/Cole Publishing Company.

Kennerknecht, I., Grueter, T., Welling, B., Wentzek, S., Horst, J., Edwards, S., & Grueter, M. (2006). First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). *Am J Med Genet A, 140*(15), 1617-1622. doi:10.1002/ajmg.a.31343

Kennerknecht, I., Ho, N. Y., & Wong, V. C. (2008). Prevalence of hereditary prosopagnosia (HPA) in Hong Kong Chinese population. *Am J Med Genet A, 146A*(22), 2863-2870. doi:10.1002/ajmg.a.32552

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). LIfetime prevalence and age-of-onset distributions of dsm-iv disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62*(6), 593-602. doi:10.1001/archpsyc.62.6.593

Liu-Shuang, J., Torfs, K., Rossion, B. (in press). An objective electrophysiological marker of face individualisation impairment in acquired prosopagnosia with fast periodic visual stimulation. *Neuropsychologia.*

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behav Res Ther, 36*(4), 455-470.

McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Gauthier, I. (2012). The Vanderbilt Expertise Test reveals domain-general and domain-specific sex effects in object recognition. *Vision Research, 69*, 10-22. doi:http://dx.doi.org/10.1016/j.visres.2012.07.014

McKone, E., & Palermo, R. (2010). A strong role for nature in face recognition. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1000567107

McKone, E., Hall, A., Pidcock, M., Palermo, R., Wilkinson, R. B., Rivolta, D., . . . O'Connor, K. B. (2011). Face ethnicity and measurement reliability affect face recognition performance in developmental prosopagnosia: evidence from the Cambridge Face Memory Test-Australian. *Cogn Neuropsychol, 28*(2), 109-146. doi:10.1080/02643294.2011.616880

Palermo, R., & Duchaine, B. (2012). Introduction to this special issue on developmental prosopagnosia. *Cogn Neuropsychol, 29*(5-6), 349-353. doi:10.1080/02643294.2012.744740

Palermo, R., Rivolta, D., Wilson, C. E., & Jeffery, L. (2011). Adaptive face space coding in congenital prosopagnosia: typical figural aftereffects but abnormal identity aftereffects. *Neuropsychologia, 49*(14), 3801-3812. doi:10.1016/j.neuropsychologia.2011.09.039

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia, 49*(5), 1226-1235. doi:10.1016/j.neuropsychologia.2011.02.021

Peters, L. (2000). Discriminant validity of the Social Phobia and Anxiety Inventory (SPAI), the Social Phobia Scale (SPS) and the Social Interaction Anxiety Scale (SIAS). *Behaviour Research and Therapy, 38*(9), 943-950. doi:http://dx.doi.org/10.1016/S0005-7967(99)00131-X

Raven, J.C. (2008) *Standard Progressive Matrices: Plus Version and Mill Hill Vocabulary Scale*. San Antonio, TX: Pearson.

Rhodes, G., Jeffery, L., Taylor, L., & Ewing, L. (2013). Autistic traits are linked to reduced adaptive coding of face identity and selectively poorer face recognition in men but not women. *Neuropsychologia, 51*(13), 2702-2708. doi:10.1016/j.neuropsychologia.2013.08.016

Rivolta, D., Palermo, R., & Schmalzl, L. (2013). What is overt and what is covert in congenital prosopagnosia? *Neuropsychol Rev, 23*(2), 111-116. doi:10.1007/s11065-012-9223-0

Rossion, B., & Curran, T. (2010). Visual expertise with pictures of cars correlates with RT magnitude of the car inversion effect. *Perception, 39*(2), 173-183. doi:10.1068/p6270

Rossion, B., Caldara, R., Seghier, M., Schuller, A.-M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain Journal of Neurology, 126*, 2381-2395.

Rossion, B., Collins, D., Goffaux, V., & Curran, T. (2007). Long-term expertise with artificial objects increases visual competition with early face categorization processes. *J Cogn Neurosci, 19*(3), 543-555. doi:10.1162/jocn.2007.19.3.543

Rotshtein, P., Geng, J. J., Driver, J., & Dolan, R. J. (2007). Role of features and second-order spatial relations in face discrimination, face recognition, and individual face skills: behavioral and functional magnetic resonance imaging data. *J Cogn Neurosci, 19*(9), 1435-1452. doi:10.1162/jocn.2007.19.9.1435

Russell, R., Duchaine, B. C., & Nakayama, K. (2009). Super-recognizers: people with extraordinary face recognition ability. *Psychonomic Bulletin & Review, 16*(2), 252-257. doi:10.3758/PBR.16.2.252

Shah, P., Gaule, A., Sowden, S., Bird, G., & Cook, R. (2015). The 20-item prosopagnosia index (PI20): a self-report instrument for identifying developmental prosopagnosia. *Royal Society Open Science, 2*(140343), 1-11.

Stollhoff, R., Jost, J., Elze, T., & Kennerknecht, I. (2011). Deficits in Long-Term Recognition Memory Reveal Dissociated Subtypes in Congenital Prosopagnosia. *PLoS ONE, 6*(1), e15702. doi:10.1371/journal.pone.0015702.g005

Susilo, T., & Duchaine, B. (2013). Advances in developmental prosopagnosia research. *Current opinion in neurobiology*, 1-7. doi:10.1016/j.conb.2012.12.011

Susilo, T., Germine, L., & Duchaine, B. (2013). Face recognition ability matures late: evidence from individual differences in young adults. *J Exp Psychol Hum Percept Perform, 39*(5), 1212-1217. doi:10.1037/a0033469

Susilo, T., Yovel, G., Barton, J. J. S., & Duchaine, B. (2013). Face perception is category-specific: Evidence from normal body perception in acquired prosopagnosia. *Cogniiton, 129*(1), 88-94. doi:http://dx.doi.org/10.1016/j.cognition.2013.06.004

Thomas, C., Avidan, G., Humphreys, K., Jung, K.-j., Gao, F., & Behrmann, M. (2009). Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nature Neuroscience, 12*(1), 29-31. doi:10.1038/nn.2224

Van Gulick, A.E., McGugin, R.W. & Gauthier, I. (in press). Measuring non-visual knowledge about object categories: The Semantic Vanderbilt Expertise Test, *Behavioral Research Methods.*

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Gerbasi, M., & Nakayama, K. (2012). Capturing specific abilities as a window into human individuality: the example of face recognition. *Cogn Neuropsychol, 29*(5-6), 360-392. doi:10.1080/02643294.2012.753433

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. *Proc Natl Acad Sci U S A, 107*(11), 5238-5241. doi:10.1073/pnas.0913053107

Yardley, L., McDermott, L., Pisarski, S., Duchaine, B., & Nakayama, K. (2008). Psychosocial consequences of developmental prosopagnosia: a problem of recognition. *J Psychosom Res, 65*(5), 445-451. doi:10.1016/j.jpsychores.2008.03.013

Zell, E., & Krizan, Z. (2014). Do People Have Insight Into Their Abilities? A Metasynthesis. *Perspectives on Psychological Science, 9*, 1745-6924.

Table 1. Descriptive statistics for variables across studies 1-4 with typical individuals: N, possible and observed range, mean and standard deviation (SD), Cronbach's alpha, skew and kurtosis.

Measures in each study	N	Chance – Ceiling	Observed Range	Mean (SD)	Cronbach's α	Skew	Kurtosis
Studies 1-3							
Face and object recognition tests							
CFMT (Studies 1-3)	241	24 – 72	30 – 72	56.56 (9.25)	.89	-.63	-.26
CFMT-Films (Study 2)	89	25 – 76	32 – 74	52.11 (11.13)	.89	.08	-1.21
BFRT (Study 1)	96	0 – 54	38 – 53	46.42 (3.31)	.69	-.33	-.71
CCMT (Studies 2-3)	146	24 – 72	32 – 72	50.64 (8.30)	.82	.41	-.72
Questionnaires							
Kennerknecht questionnaire (Studies 1-3)	241	15 – 75	16 – 62	29.12 (7.32)	.84	.82	1.33
Self-reported own-race face recognition (Study 3)	57	1 – 9	3 – 8	5.61 (1.30)	-	-.57	-.18
Self-reported general face recognition (Study 3)	57	1 – 9	3 – 8	5.56 (1.20)	-	-.51	-.27
SIAS (Studies 2-3)	139	0 – 80	4 – 68	28.31 (14.80)	.89	.10	-.52
AQ (Study 3)	57	0-50	2-31	15.68 (6.68)	.81	.29	-.47
Study 4							
Face recognition tests							
CFMT	58	24 – 72	34-72	56.66 (8.91)	.89	-.57	-.43
Face matching (4AFC-FF)	58	56 – 224	58.04 – 98.21	82.73 (7.72)	.93	-.96	1.11
Face matching (4AFC-DR)	58	56 – 224	54.91 – 95.98	82.71 (8.52)	.94	-1.16	2.12
Questionnaire							
Face Metacognition questionnaire	58	11 – 100	55.84 – 85.28	75.56 (6.81)	.86	-.54	.13

Notes: CFMT - Cambridge Face Memory Test (Duchaine & Nakayama, 2006); Benton Facial Recognition Test (BFRT, Benton et al., 1983); CCMT – Cambridge Car Memory Test (Dennett et al., 2012), Kennerknecht et al. questionnaire (Kennerknecht et al., 2008); SIAS – Social Interaction Anxiety Scale (Mattick & Clarke, 1998); AQ – Autism Quotient (Baron-Cohen et al., 2001);

Table 2. Raw scores and Z-scores for each CP, labelled by sex and approximate age when completing the tests.

	MFFT 2008		CFMT		CFMT-Aus		CFPT		CCMT		Questionnaires		
Sex/Age	Percent Correct	Z-score	Accuracy (/72)	Z-score	Accuracy (/72)	Z-score	Errors (/chance = 93.3)	Z-score	Accuracy (/72)	Z-score	Kennerknecht	SIAS	AQ
F33	23.08	**-3.47**	38	**-2.09**	51	-0.92	66	**-2.86**	63	1.76	46	22	11
F40	35.29	**-2.51**	37	**-2.16**	45	-1.73	68	**-2.95**	41	-1.32	45	20	14
F47	6.25	**-4.00**	39	*-1.81*	41	**-2.28**	52	-1.41	45	-0.76	55	52	30
M57	15.00	**-3.14**	28	**-2.83**	45	*-1.73*	70	*-1.93*	35	**-2.70**	29	29	20
M60	45.00	-1.26	30	**-2.49**	40	**-2.42**	48	-0.03	-	-	46	60	28
F23	25.00	**-3.70**	52	-0.40	46	-1.60	50	-1.45	54	0.50	41	49	21
F27	68.42	-0.97	41	*-1.73*	-	-	52	*-1.66*	39	-1.60	53	27	9
F33	40.00	**-2.46**	39	**-1.98**	-	-	40	-0.64	59	1.20	62	55	26
F42	16.67	**-3.55**	42	-1.55	38	**-2.69**	26	0.68	65	2.04	37	15	8
F44	11.76	**-3.78**	33	**-2.58**	-	-	54	*-1.67*	51	0.08	47	21	18
F46	13.33	**-3.62**	29	**-3.01**	-	-	62	**-2.29**	44	-0.90	51	12	9
M54	16.67	**-3.15**	29	**-2.81**	-	-	64	-1.62	48	-1.13	46	33	20
M59	30.77	**-2.14**	35	*-1.94*			38	0.69	67	1.15	49	36	34
Mean (SD)	26.71 (17.12)	-2.90	36.31 (6.75)	-2.11	43.71 (4.39)	-1.91	53.08 (13.05)	-1.32	53.38 (9.87)	0.17	46.69 (8.22)	33.15 (16.05)	19.08 (8.63)

Notes: MFFT-2008 - Macquarie Famous Face Test 2008 (Palermo et al., 2011, with z-score age-adjustment); CFMT - Cambridge Face Memory Test (Duchaine & Nakayama, 2006; which includes age-adjusted z-scores from Bowles et al., 2009; CFMT-Aus, (Cambridge Face Memory Test-Australian, McKone et al., 2011); CFPT – Cambridge Face Perception Test (Duchaine et al., 2007, which includes sex- and age-adjusted reversed z-scores from Bowles et al., 2009); CCMT – Cambridge Car Memory Test (Dennett et al., 2012), Kennerknecht et al. questionnaire (Kennerknecht et al., 2008); SIAS – Social Interaction Anxiety Scale (Mattick & Clarke, 1998); Autism Quotient (AQ) Baron-Cohen et al. 2001. Z-scores 2 standard deviations (SD) below controls are in bold and those -1.7 below are in italics.