

In press at: *Royal Society Open Science*

Format: *Research Article*

Running head: *Screening samples for prosopagnosia*

Robust associations between the 20-Item Prosopagnosia Index and the Cambridge Face Memory Test in the general population

Katie L.H. Gray¹ Geoffrey Bird^{2,3}, and Richard Cook^{4*}

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

²Experimental Psychology Department, University of Oxford, Oxford, U.K.

³MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry,
Psychology, and Neuroscience, King's College London, London, U.K.

⁴Department of Psychology, City, University of London, London, U.K.

*Corresponding author:

Richard.Cook.1@city.ac.uk

Department of Psychology
City, University of London
Whiskin Street
London, U.K.
EC1R 0JD

Abstract

Developmental prosopagnosia (DP) is a neurodevelopmental condition, characterised by lifelong face recognition deficits. Leading research groups diagnose the condition using complementary computer-based tasks and self-report measures. In an attempt to standardise the reporting of self-report evidence, we recently developed the 20-Item Prosopagnosia Index (PI20), a short questionnaire measure of prosopagnosic traits suitable for screening adult samples for DP. Strong correlations between scores on the PI20 and performance on the Cambridge Face Memory Test (CFMT) appeared to confirm that individuals possess sufficient insight into their face recognition ability to complete a self-report measure of prosopagnosic traits. However, the extent to which people have insight into their face recognition abilities remains contentious. A lingering concern is that feedback from formal testing, received prior to administration of the PI20, may have augmented the self-insight of some respondents in the original validation study. To determine whether the significant correlation with the CFMT was an artefact of previously delivered feedback, we sought to replicate the validation study in individuals with no history of formal testing. We report highly significant correlations in two independent samples drawn from the general population, confirming i) that a significant relationship exists between PI20 scores and performance on the CFMT, and ii) that this is not dependent on the inclusion of individuals who have previously received feedback. These findings support the view that people have sufficient insight into their face recognition abilities to complete a self-report measure of prosopagnosic traits.

1. Introduction

Developmental prosopagnosia (DP) is a neurodevelopmental condition, characterised by lifelong deficits in facial identity recognition, despite normal intelligence, typical low-level vision, and no history of brain damage [1-4]. Individuals with DP typically utilise non-face cues including voice, gait, and hairstyle to recognise others. Consequently, they often experience great difficulties when non-face cues are unavailable or changed, or when familiar people are encountered out of context. DP is known to be a heterogeneous condition; for example, some individuals appear to perceive facial expressions normally [5], whereas others exhibit problems with facial expression perception [6]. Similarly, some individuals with DP recognize objects normally [7, 8], while others exhibit broader object recognition deficits [9, 10]. DP can be a socially debilitating condition often associated with social isolation, depression and anxiety, and reduced employment opportunities [11, 12].

DP is not listed in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [13] and currently no formal diagnostic criteria exist. Leading research groups therefore diagnose DP through the accumulation of convergent diagnostic evidence. Computer-based tests of face recognition ability, including the Cambridge Face Memory Test (CFMT [14]) and the Cambridge Face Perception Test (CFPT [10]) form a key part of most diagnostic batteries. Many authors also report performance on famous face recognition tests [e.g., 7, 15, 16]. In addition to scores on computer-based tests, however, self-report measures provide a complementary source of diagnostic evidence. For example, research groups routinely conduct diagnostic interviews and administer questionnaire measures that enquire about the face recognition experience of potential DPs [e.g., 17]. Where objective computer-based measures and subjective self-report measures provide convergent evidence of impairment, researchers can be confident about diagnosis and classification [18].

Historically, different research groups have employed bespoke self-report procedures, hampering the description and comparison of self-report data. In an attempt to standardise the reporting of self-report evidence, the *troublewithfaces.org* team recently published the 20-Item Prosopagnosia Index (PI20), a short questionnaire measure of prosopagnosic traits suitable for screening adult samples for DP [19]. Respondents indicate the extent to which twenty statements describe their face recognition abilities and experiences. Agreement is rated on a five-point scale yielding scores ranging from 20 to 100. Sample items include: *I*

often mistake people I have met before for strangers; Without hearing people's voices I struggle to recognise them; I sometimes find movies hard to follow because of difficulties recognising characters. Scores on the questionnaire have been shown to effectively distinguish previously classified DPs from typical observers falling within the normal range of abilities [6, 19, 20].

As part of the validation procedures, the original PI20 paper described a highly significant correlation ($r = -.68, p < .001$) between PI20 scores and performance on the CFMT [14]; a leading standardized measure of face recognition ability, employing a three-alternative-forced-choice match-to-sample design (see Validation Study 4 [19]). These results helped to confirm a key premise underlying the logic of the PI20; that individuals have sufficient insight into their face recognition ability to complete a self-report measure of prosopagnosic traits. However, the validation studies included a number of previously diagnosed DPs in the sample (~21%). Some of these known DPs have been involved in previous research and had therefore received feedback from formal testing prior to administration of the PI20 questionnaire. It is conceivable that this feedback may have augmented their self-insight and thereby influenced how they completed the scale.

The possibility that the highly significant correlation described in the original PI20 paper [19] is an artefact of previously delivered feedback (e.g., results from formal testing), casts doubt on the crucial self-insight premise on which the PI20 is predicated, and potentially undermines the value of the scale as an independent source of diagnostic evidence. Here, we present novel data addressing this concern. We confirm that a significant relationship exists between PI20 scores and performance on the CFMT, and that this is not dependent on the inclusion of individuals who have previously received feedback from formal testing. We focus on the relationship with the CFMT as this is widely regarded as the most telling source of diagnostic evidence; whilst high scores on the CFMT typically exclude a diagnosis, members of DP samples sometimes score within the normal range on the CFPT [e.g., 15] and on famous face recognition tests [e.g., 16].

2. Method and results

Data was collected from two independent samples. None of the participants had completed formal testing of their face recognition ability. Self-reported face recognition ability played

no role in the recruitment or selection of participants. The first sample ($N = 142$) was collected at City, University of London, and comprised adults recruited from the local subject pool ($M_{\text{age}} = 29.23$, $SD_{\text{age}} = 11.91$, 56 males). Individuals were paid a small honorarium in return for their participation. The second sample ($N = 283$) was collected by undergraduate students at the University of Reading ($M_{\text{age}} = 26.64$, $SD_{\text{age}} = 13.16$, 106 males). All participants completed both the PI20 and the CFMT. The first sample completed the PI20 before the CFMT; the second sample completed the CFMT, then the PI20. Participants were debriefed and given feedback only once both tasks had been completed. Ethical clearance was granted by the local ethics committees. The study was conducted in line with the ethical guidelines laid down in the 6th (2008) Declaration of Helsinki. All participants gave informed consent.

The first sample ($N = 142$) scored between 23 and 68 on the PI20 ($M = 40.10$; $SD = 9.58$) and between 45.8% and 100% on the CFMT ($M = 80.65$; $SD = 12.79$). The second sample ($N = 283$) scored between 20 and 74 on the PI20 ($M = 41.70$; $SD = 10.10$) and between 47.2% and 100% on the CFMT ($M = 76.80$; $SD = 12.90$). Three participants from the first sample, and nine from the second, yielded PI20 scores that exceeded the diagnostic cut-off (≥ 65) suggested in the original study [19]. Crucially, we found highly significant correlations between participants' scores on the PI20 and CFMT in both the first sample, $r = -.394$, $p < .001$ (Figure 1a), and in the second sample, $r = -.390$, $p < .001$ (Figure 1b). The participants in these samples had no opportunity to use feedback from formal testing to inform their responses. These findings therefore lend further support for the view that people have sufficient insight into their face recognition abilities to complete a self-report measure of prosopagnosic traits.

3. Discussion

The correlations presented here represent important additions to the literature on the PI20 insofar as they estimate the relationship seen between PI20 scores and CFMT performance in the general population. Of the 110 observers who took part in the original validation study [19], 23 (21%) were known or suspected DPs. In contrast, the incidence of DP in the general population is thought to be ~2% [21, 22]. A substantial number of DPs were included in the original sample in order to document the relationship between PI20 scores and performance on the CFMT across the entire range of abilities (i.e. normal and impaired), and thereby

confirm the utility of the PI20 as a diagnostic tool. Recently, the aim of the original study has been misunderstood; some authors have implied this correlation estimates the relationship between PI20 scores and performance on the CFMT in the general population [23-25]. However, this was *not* the aim of the original study; rather it sought to validate the PI20 as a diagnostic instrument [19]. It is very clear that the incidence of DP in the general population is much lower than 21% [21, 22].

As expected, the correlations observed in the present datasets ($\sim -.39$) were weaker than seen in the original validation study. Two factors are likely to contribute to this disparity. First, the range of abilities in the present samples is narrower than was employed in the original validation study. The variability within to-be-correlated variables will inevitably influence the strength of any correlation observed. By way of analogy, one may expect a weaker relationship between IQ and school achievement in samples of university students, than in samples drawn from the general population [26]. Consistent with this observation, a weaker correlation is also seen when the correlational analysis described in the original validation study ($N = 110$, $r = -.68$) is restricted to those participants who did not describe face recognition problems ($N = 87$, $r = -.32$). The strength of correlation seen in small samples drawn from the general population may be quite variable as it is influenced by the number of potential prosopagnosics identified. With larger samples, the correlation estimates are likely to stabilise.

Second, PI20 scores are ill-suited for estimating individual differences within the normal or superior range of abilities. Despite the correlations observed here, it is important to recognise that the PI20 is a measure of prosopagnosic traits, not a measure of face recognition ability *per se*. For example, observers in the 45th and 55th percentile of the general population will likely respond in very similar ways to items such as “*Anxiety about face recognition has led me to avoid certain social or professional situations.*” Only people with very bad face recognition are likely to recognise such experiences; the rest of the population will not, irrespective of whether they have adequate, good, or excellent face recognition.

Unsurprisingly, PI20 scores from the typical population therefore exhibit some positive skewing (see Figure 1a and 1b), suggestive of asymmetric sensitivity. This feature is seen in several popular instruments used to screen for neurodevelopmental disorders [e.g., 27].

Whether or not people have insight into their face recognition ability is a deceptively complex question; findings will likely depend on how estimates of self-reported ability are elicited and who is asked. Because the PI20 uses a number of concrete statements and easy-to-recognise anecdotes, respondents can interpret items even if they have had little cause to reflect on their ability. Estimating self-reported ability using abstract single-item measures (e.g. asking participants to rate their face recognition ability “compared with the average person”) may not be a fruitful approach [19]. Nevertheless, we note that self-report scores elicited using abstract one-shot measures *do* correlate significantly with objective measures of face recognition ability [24, 25, 28]. Individuals with extremely good or extremely bad face recognition ability (so-called ‘super-recognisers’ [29] and DPs, respectively) are also more likely to encounter situations in their daily lives that illustrate that face recognition is a distributed ability, and suggest where they might fall within that distribution. Unnuanced assertions that people lack insight in to their ability [e.g., 18] are therefore overly-simplistic.

Cases of DP should not be diagnosed based solely on self-report evidence. However, when used properly, the PI20 provides independent diagnostic evidence that complements scores from objective computer-based tasks. There is a multitude of reasons why participants with typical face perception may score badly on computer-based tests, including boredom and fatigue, a lack of motivation, prioritisation of response speed over accuracy, test anxiety, manual and technical difficulties [19]. When tested on the CFMT, large undergraduate samples routinely yield numerous scores in the DP range [28]. However, in the absence of convergent self-report evidence, such scores should be treated with caution; the embarrassing social consequences of poor face recognition ensure that genuine sufferers are usually aware of their issue. The inclusion of self-report measures in diagnostic batteries also ensures that novel forms of DP do not go undetected. For example, selective difficulties processing dynamic faces, or problems learning faces from multiple encounters, will not be picked up by leading computer-based tests which assess perception of static unfamiliar faces only [10, 14].

Ethics

Both studies were granted ethical clearance by the local ethics committees (University of Reading and City, University London) and were conducted in line with the ethical guidelines laid down in the 6th (2008) Declaration of Helsinki. All participants gave informed consent.

Data accessibility

The datasets supporting this article are available as electronic supplementary material.

Competing interests

The authors have no competing interests.

Authors' contributions

KG contributed to the study design, supervised data collection, conducted analysis, and helped draft the manuscript for publication. GB contributed to the study design and helped draft the manuscript for publication. RC contributed to the study design, supervised data collection, and helped draft the manuscript for publication.

Funding

The authors received no funding for this research.

References

1. Behrmann M, Avidan G. 2005 Congenital prosopagnosia: face-blind from birth. *Trends Cogn. Sci.* **9**, 180-187.
2. Cook R, Biotti F. 2016 Developmental prosopagnosia. *Curr. Biol.* **26**, R312-R313.
3. Duchaine B, Nakayama K. 2006 Developmental prosopagnosia: a window to content-specific face processing. *Curr. Opin. Neurobiol.* **16**, 166-173.
4. Susilo T, Duchaine B. 2013 Advances in developmental prosopagnosia research. *Curr. Opin. Neurobiol.* **23**, 423-429.
5. Duchaine B, Parker H, Nakayama K. 2003 Normal recognition of emotion in a prosopagnosic. *Perception* **32**, 827-838.
6. Biotti F, Cook R. 2016 Impaired perception of facial emotion in developmental prosopagnosia. *Cortex* **81**, 126-136.
7. Duchaine B, Nakayama K. 2005 Dissociations of face and object recognition in developmental prosopagnosia. *J. Cogn. Neurosci.* **17**, 249-261.
8. Duchaine B, Yovel G, Butterworth EJ, Nakayama K. 2006 Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. *Cogn. Neuropsychol.* **23**, 714-747.
9. Behrmann M, Avidan G, Marotta JJ, Kimchi R. 2005 Detailed exploration of face-related processing in congenital prosopagnosia: 1. Behavioral findings. *J. Cogn. Neurosci.* **17**, 1130-1149.
10. Duchaine B, Germine L, Nakayama K. 2007 Family resemblance: ten family members with prosopagnosia and within-class object agnosia. *Cogn. Neuropsychol.* **24**, 419-430.
11. Dalrymple KA, Fletcher K, Corrow S, das Nair R, Barton JJ, Yonas A, Duchaine B. 2014 "A room full of strangers every day": the psychosocial impact of developmental prosopagnosia on children and their families. *J. Psychosom. Res.* **77**, 144-150.
12. Yardley L, McDermott L, Pisarski S, Duchaine B, Nakayama K. 2008 Psychosocial consequences of developmental prosopagnosia: a problem of recognition. *J. Psychosom. Res.* **65**, 445-451.
13. American Psychiatric Association. 2013 *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition ed.). Washington, DC: American Psychiatric Association.
14. 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**, 576-585.
15. Bobak AK, Parris BA, Gregory NJ, Bennetts RJ, Bate S. 2017 Eye-movement strategies in developmental prosopagnosia and super-face recognition. *Q. J. Exp. Psychol.* **70**, 201-217.

16. Liu TT, Behrmann M. 2014 Impaired holistic processing of left-right composite faces in congenital prosopagnosia. *Front. Hum. Neurosci.* **8**, 750.
17. Grüter T, Grüter M, Carbon CC. 2011 Congenital prosopagnosia. Diagnosis and mental imagery: Commentary on “Tree JJ, and Wilkie J. Face and object imagery in congenital prosopagnosia: a case series”. *Cortex* **47**, 511-513.
18. Dalrymple KA, Palermo R. 2016 Guidelines for studying developmental prosopagnosia in adults and children. *Wiley Interdiscip. Rev. Cogn. Sci.* **7**, 73-87.
19. 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. *R. Soc. Open Sci.* **2**, 140343.
20. Rubino C, Corrow SL, Corrow JC, Duchaine B, Barton JJ. 2016 Word and text processing in developmental prosopagnosia. *Cogn. Neuropsychol.* **33**, 315-328.
21. Kennerknecht I, Grüter T, Welling B, Wentzek S, Horst J, Edwards S, Grüter M. 2006 First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). *Am. J. Med. Genet. A* **140**, 1617-1622.
22. Kennerknecht I, Ho NY, Wong VCN. 2008 Prevalence of hereditary prosopagnosia (HPA) in Hong Kong Chinese population. *Am. J. Med. Genet. A* **146**, 2863-2870.
23. Bate S, Tree JJ. 2017 The definition and diagnosis of developmental prosopagnosia. *Q. J. Exp. Psychol.* **70**, 193-200.
24. Bobak AK, Pampoulov P, Bate S. 2016 Detecting superior face recognition skills in a large sample of young British adults. *Front. Psychol.* **7**, 1378.
25. Palermo R, Rossion B, Rhodes G, Laguesse R, Tez T, Hall B, Albonico A, Malaspina M, Daini R, Irons J. 2017 Do people have insight into their face recognition abilities? *Q. J. Exp. Psychol.* **70**, 218-233.
26. Howitt D, Cramer D. 2009 *Introduction to Research Methods in Psychology*. Harlow: Pearson Prentice Hall.
27. Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E. 2001 The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* **31**, 5-17.
28. Bowles DC, McKone E, Dawel A, Duchaine B, Palermo R, Schmalzl L, Rivolta D, Wilson CE, 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**, 423-455.
29. Russell R, Duchaine B, Nakayama K. 2009 Super-recognizers: People with extraordinary face recognition ability. *Psychon. Bull. Rev.* **16**, 252-257.

Figure

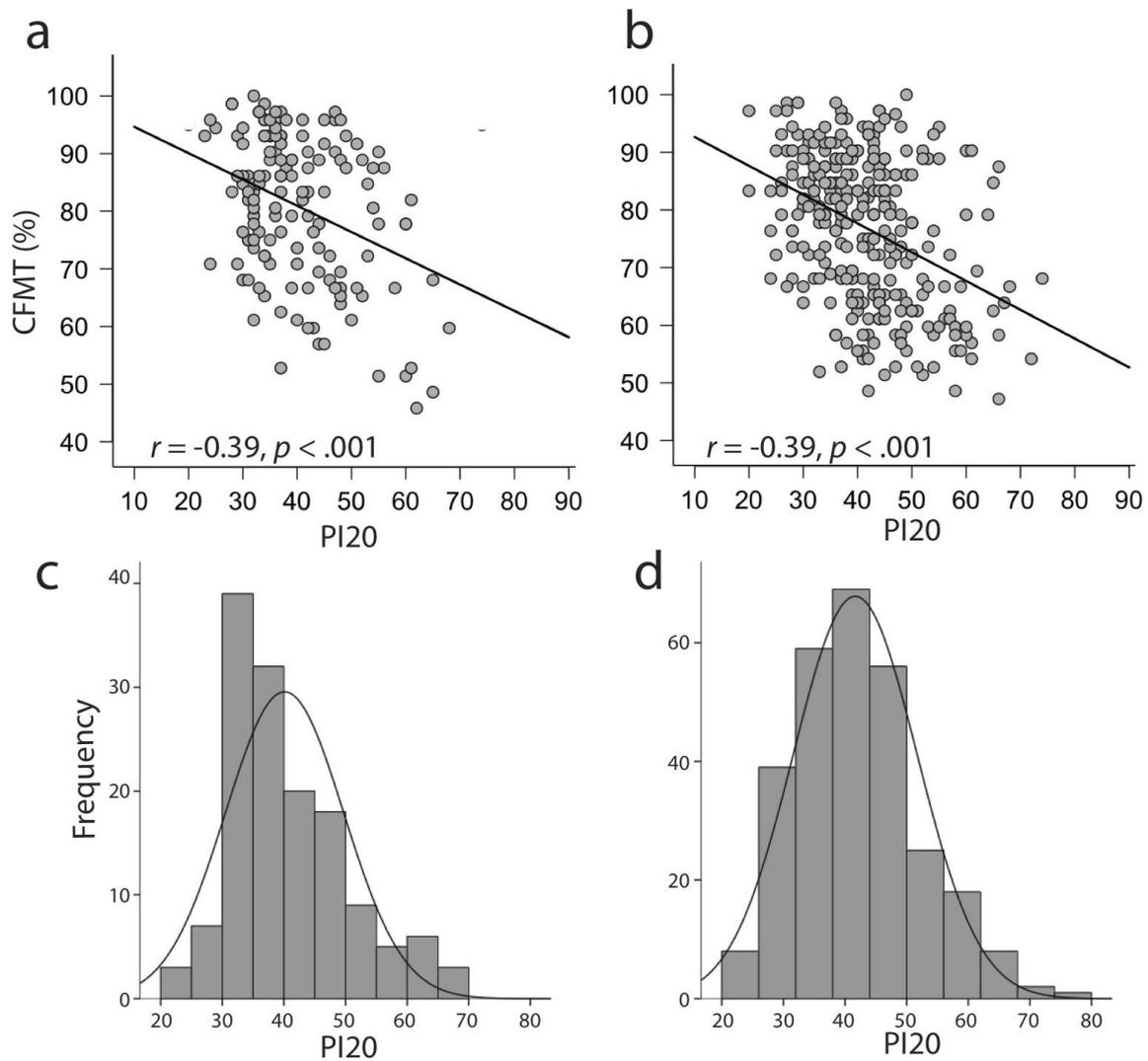


Figure 1. Associations between PI20 scores and performance on the CFMT for (a) the first sample collected at City, University of London ($N = 142$) and (b) the second sample collected through the University of Reading ($N = 283$). Both the first (a) and second (b) samples exhibited some indication of positive skewing. Note the frequency values differ between (c) and (d).