City Research Online

City St George's, University of London

Citation: 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(6), 140343. doi: 10.1098/rsos.140343

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/16738/

Link to published version: https://doi.org/10.1098/rsos.140343

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

ROYAL SOCIETY OPEN SCIENCE

rsos.royalsocietypublishing.org

Brief report

Cite this article: 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. http://dx.doi.org/10.1098/rsos.140343

Received: 29 September 2014
Accepted: 26 May 2015

Subject Category:
Psychology and cognitive neuroscience

Subject Areas:
cognition/neuroscience

Keywords:
developmental prosopagnosia, self-report, questionnaire, face blindness, congenital prosopagnosia, face perception

Author for correspondence:
Richard Cook
e-mail: richard.cook.1@city.ac.uk

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsos.140343 or via http://rsos.royalsocietypublishing.org.

THE ROYAL SOCIETY PUBLISHING

The 20-item prosopagnosia index (PI20): a self-report instrument for identifying developmental prosopagnosia

Punit Shah[1,2], Anne Gaule[3], Sophie Sowden[1], Geoffrey Bird[1,4] and Richard Cook[2]

[1]Social, Genetic and Developmental Psychiatry Centre (MRC), Institute of Psychiatry, Psychology and Neuroscience, King's College London, University of London, London, UK
[2]Department of Psychology, City University London, London, UK
[3]Division of Psychology and Language Sciences, University College London, Bedford Way, London, UK
[4]Institute of Cognitive Neuroscience, University College London, London, UK

PS, 0000-0001-5497-4765

Self-report plays a key role in the identification of developmental prosopagnosia (DP), providing complementary evidence to computer-based tests of face recognition ability, aiding interpretation of scores. However, the lack of standardized self-report instruments has contributed to heterogeneous reporting standards for self-report evidence in DP research. The lack of standardization prevents comparison across samples and limits investigation of the relationship between objective tests of face processing and self-report measures. To address these issues, this paper introduces the PI20; a 20-item self-report measure for quantifying prosopagnosic traits. The new instrument successfully distinguishes suspected prosopagnosics from typically developed adults. Strong correlations were also observed between PI20 scores and performance on objective tests of familiar and unfamiliar face recognition ability, confirming that people have the necessary insight into their own face recognition ability required by a self-report instrument. Importantly, PI20 scores did not correlate with recognition of non-face objects, indicating that the instrument measures face recognition, and not a general perceptual impairment. These results suggest that the PI20 can play a valuable role in identifying DP. A freely available self-report instrument will permit more effective description of self-report diagnostic evidence, thereby facilitating greater comparison of prosopagnosic samples, and more reliable classification.

© 2015 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

1. Introduction

Developmental prosopagnosia[1] (DP) is a neurodevelopmental condition characterized by face recognition difficulties, despite normal intelligence, typical visual acuity and intact socio-cognitive abilities [1–6]. Unlike cases of acquired prosopagnosia [7], deficits are seen in the absence of brain injury. In the majority of cases, ability to judge facial emotion [8,9], and make other social attributions [10,11], is unaffected. DP often runs in families, suggestive of a genetic component [12–15]. Individuals with DP often avoid social situations, experiencing feelings of guilt and shame about actual or imagined offence caused to others [16]. Long-term consequences can include a reduced social circle, loss of self-confidence and limited work opportunities [17–19].

Studies have estimated that DP may affect between 1.9% [20] and 2.5% [21] of the population. Crucially, however, identifying cases is not straightforward [22]: DP is not listed in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5 [23]) as a psychiatric disorder and no formal diagnostic criteria exist. Current approaches to 'diagnosis' emphasize performance on objective, computer-based tests of face recognition ability, including the Benton Facial Recognition Test [24], different versions of the Famous Faces Recognition Test (FFRT [25]), the Cambridge Face Memory Test (CFMT [26]) and the Cambridge Face Perception Test (CFPT [13]). The current consensus is that an individual should demonstrate substantial impairment (e.g. a score of 2 s.d. below the mean of a matched control sample) on either the CFPT or the CFMT—regarded as the leading objective tests of face recognition—to be diagnosed with DP [27].

Few would disagree that objective tests have been crucial for establishing our current knowledge of DP and must continue to play a key role in identifying individuals with the disorder [4,27]. However, relying exclusively on these tasks for diagnostic evidence may not always produce reliable classification. Whether a given score falls within the prosopagnosic range—beyond 2 s.d. of the control mean—is sensitive to the composition of the control group, and may be affected by control-participants' age, gender, IQ and ethnic origin [27–29]. Equally, cut-offs are affected by the spread of control scores and the treatment of outliers. Further factors, unrelated to the presence of DP, may also cause someone to perform badly, including poor motivation, misinterpretation of instructions (e.g. prioritizing response speed over accuracy), poor mouse control and test anxiety [30]. For example, administering the CFMT and CFPT to undergraduate samples routinely produces scores within the prosopagnosic range [27], however the nature of these outlying scores remains unclear; many of the individuals identified report entirely normal face-recognition in their daily lives.

Conversely, individuals may have genuinely impaired face perception but perform within the normal range on computer-based tests. For example, prosopagnosics may develop compensatory strategies to negotiate laboratory tasks or have perceptual issues extending beyond static greyscale faces. Given the three-alternative-forced-choice (3AFC) format of the CFMT, some respondents will gain a non-trivial number of hits by chance. It is worth noting an asymmetry in the standard of evidence provided by the CFMT and related tests employing this format: whereas errors are unlikely to be due to chance (where correct responses are known, participants rarely enter incorrect responses erroneously), correct answers may often reflect lucky guesses (a third of guesses will be correct in a 3AFC task). Consequently, scores within the typical range may not always be strong evidence of intact perception.

In recognition of the foregoing issues, few researchers attempt to characterize DP using a single diagnostic instrument; rather, leading research groups use a number of tests to develop a perceptual profile, and diagnose DP only where convergent evidence accumulates. In practice, self-report forms a key part of this profile. When contacted by suspected prosopagnosics, researchers routinely conduct interviews (e.g. [31,32]) or administer bespoke questionnaires. This self-report evidence is valuable when interpreting performance on computer-based tasks. For example, a low score on a computer-based test in the absence of self-reported problems (or vice versa) may be regarded with caution and additional testing conducted. However, the absence of standardized instruments has led to poor reporting of self-report evidence. Articles are often vague with respect to the methods and results of self-report assessments. Moreover, the use of different instruments by different research groups has prevented comparison of self-reported problems across samples.

To help address these issues, we present a short, validated self-report questionnaire for assessing prosopagnosic traits. Self-report questionnaires have been crucial in the study of other neurodevelopmental disorders (e.g. [33]) and often prove significant in the development of formal diagnostic criteria [34]. The only existing self-report questionnaire for prosopagnosic traits [21] has been

[1]We use the term developmental prosopagnosia in preference to congenital prosopagnosia to reflect the possibility that the condition emerges during development, and may not necessarily be present from birth.

criticized on the grounds that it correlates poorly with objective measures of face recognition ability [35]. Published correlations between scores on this scale and objective tests of face recognition ability range from $r = 0.20$ [15] to 0.55 [36]. The weak relationship observed probably reflects the inclusion of items pertaining to navigation deficits, the presence of face recognition difficulties in the respondents' wider family, and ability to judge facial attractiveness, facial emotion and facial gender: issues that are not reliable features of DP.[2]

2. The 20-item prosopagnosia index

The 20-item prosopagnosia index (PI20) is a self-report instrument assessing the presence of prosopagnosic traits. Respondents indicate the extent to which 20 statements describe their face recognition experiences (table 1). Agreement is scored on a five-point scale (strongly agree to strongly disagree). Fifteen statements are scored positively, whereby strongly agree is scored '5' and strongly disagree is scored '1'. Five items are reverse scored (strongly agree is scored '1' and strongly disagree is scored '5').

Items were generated following review of the qualitative [16,18,19,37–39] and quantitative literature (e.g. [1,2]) on DP, and through discussions with DPs. Items require no previous knowledge about DP, permitting the identification of sufferers who are unaware they have the condition. No items relate to emotion recognition, navigation difficulties, problems judging facial attractiveness and facial gender. Furthermore, no items were included on the presence of DP in the respondents' wider family, ensuring that the PI20 can be used in quantitative genetic studies estimating the heritability of the condition (defining DP using such criteria renders any conclusion about heritability circular).

3. Validation Study 1

The PI20 purports to measure prosopagnosic traits. A key indicator of its construct validity is therefore its ability to distinguish known or suspected DPs from the wider population. To determine whether the PI20 satisfies this fundamental criterion, the questionnaire was administered remotely via the Internet, to a sample of suspected DPs and typically developed (TD) controls.

3.1. Participants and methods

Three-hundred-and-nineteen adults aged between 18 and 74 years participated in Validation Study 1: 242 TD ($M_{age} = 29.8$ years; 87 males) and 77 suspected DPs ($M_{age} = 43.0$ years; 30 males). All participants reported normal or corrected-to-normal vision. TD participants were recruited using a local participant database. Suspected DPs contacted the authors via www.troublewithfaces.org complaining of face recognition difficulties, or were recruited via online communities for individuals with DP. Importantly, these individuals identified themselves as suspected prosopagnosics *before* administration of the PI20 questionnaire. Typically, the suspected DPs had heard about the condition through friends, family or the popular media. Having sought further information from a variety of sources, and recognized the features and anecdotes described, individuals made themselves known to the authors.

3.2. Results

The mean PI20 score of the suspected DPs ($M = 82.01$, s.d. $= 9.34$; figure 1) far exceeded that of the controls ($M = 38.90$, s.d. $= 10.88$), $t_{317} = 31.29$, $p < 0.001$. Of the 77 suspected DPs, 74 (96.1%) scored more than 2.5 s.d. above the control mean (more than or equal to 66), and 67 (87.0%) scored more than 3 s.d. above the control mean (more than or equal to 72). The mean responses given by the two samples differed significantly on all 20 items (all t's > 4.47, all p's < 0.001). The 319 responses yielded a Chronbach's α of 0.96, indicating that the 20 items have high internal consistency. Exploratory factor analysis with Varimax rotation also suggested a strong single factor structure accounting for 61% of the variance in responses.[3]

[2]By including these items, the authors hoped to improve the specificity of the instrument; in other words, to exclude individuals from their DP research who report such problems as they are likely to suffer from a co-occurring condition [32]. However, this means that the instrument is not specifically designed for, nor is suitable for, the quantification of prosopagnosic traits.

[3]The PI20 exhibited strong internal consistency when reliability analyses were restricted to the sample of suspected DPs (Chronbach's $\alpha = 0.84$). Similarly, a clear single factor structure was also present.

Table 1. The 20 statements comprising the PI20, shown with the mean scores for the TD controls and suspected prosopagnosics for each item.

		controls	troubled
1	My face recognition ability is worse than most people	1.88 (*0.89*)	4.66 (*0.51*)
2	I have always had a bad memory for faces	1.88 (*0.96*)	4.55 (*0.74*)
3	I find it notably easier to recognize people who have distinctive facial features	3.69 (*1.11*)	4.31 (*1.00*)
4	I often mistake people I have met before for strangers	1.90 (*1.04*)	4.54 (*0.70*)
5	When I was at school I struggled to recognize my classmates	1.34 (*0.72*)	3.43 (*1.25*)
6	When people change their hairstyle, or wear hats, I have problems recognizing them	1.86 (*0.95*)	4.33 (*0.86*)
7	I sometimes have to warn new people I meet that I am 'bad with faces'	1.47 (*0.85*)	4.12 (*1.07*)
8*	I find it easy to picture individual faces in my mind	2.38 (*1.16*)	4.25 (*0.96*)
9*	I am better than most people at putting a 'name to a face'	2.76 (*1.17*)	4.55 (*0.80*)
10	Without hearing people's voices, I struggle to recognize them	1.66 (*0.87*)	3.78 (*1.03*)
11	Anxiety about face recognition has led me to avoid certain social or professional situations	1.36 (*0.79*)	3.76 (*1.27*)
12	I have to try harder than other people to memorize faces	1.84 (*1.03*)	4.43 (*0.76*)
13*	I am very confident in my ability to recognize myself in photographs	1.42 (*0.87*)	2.58 (*1.28*)
14	I sometimes find movies hard to follow because of difficulties recognizing characters	1.73 (*1.08*)	4.52 (*0.75*)
15	My friends and family think I have bad face recognition or bad face memory	1.50 (*0.87*)	4.16 (*0.83*)
16	I feel like I frequently offend people by not recognizing who they are	1.69 (*0.98*)	4.25 (*0.79*)
17*	It is easy for me to recognize individuals in situations that require people to wear similar clothes (e.g. suits, uniforms and swimwear)	2.55 (*1.24*)	4.25 (*1.05*)
18	At family gatherings, I sometimes confuse individual family members	1.36 (*0.83*)	2.67 (*1.25*)
19*	I find it easy to recognize celebrities in 'before-they-were-famous' photos, even if they have changed considerably	2.79 (*1.23*)	4.61 (*0.72*)
20	It is hard to recognize familiar people when I meet them out of context (e.g. meeting a work colleague unexpectedly while shopping)	2.01 (*1.11*)	4.49 (*0.66*)

Standard deviations are shown in italics inside parentheses. Items marked with asterisks (*) are reverse scored.

4. Validation Study 2

Many of the respondents described in Validation study 1 have never undergone formal testing of their face recognition ability. It is therefore likely that both groups are heterogeneous: some self-diagnosed DPs may in fact fall within the typical range of face recognition ability and some of the individuals identified as TD may exhibit some prosopagnosic traits. Validation Study 2 therefore sought to confirm that the PI20 distinguishes between prosopagnosics whose deficits have been verified by formal testing and age-matched controls.

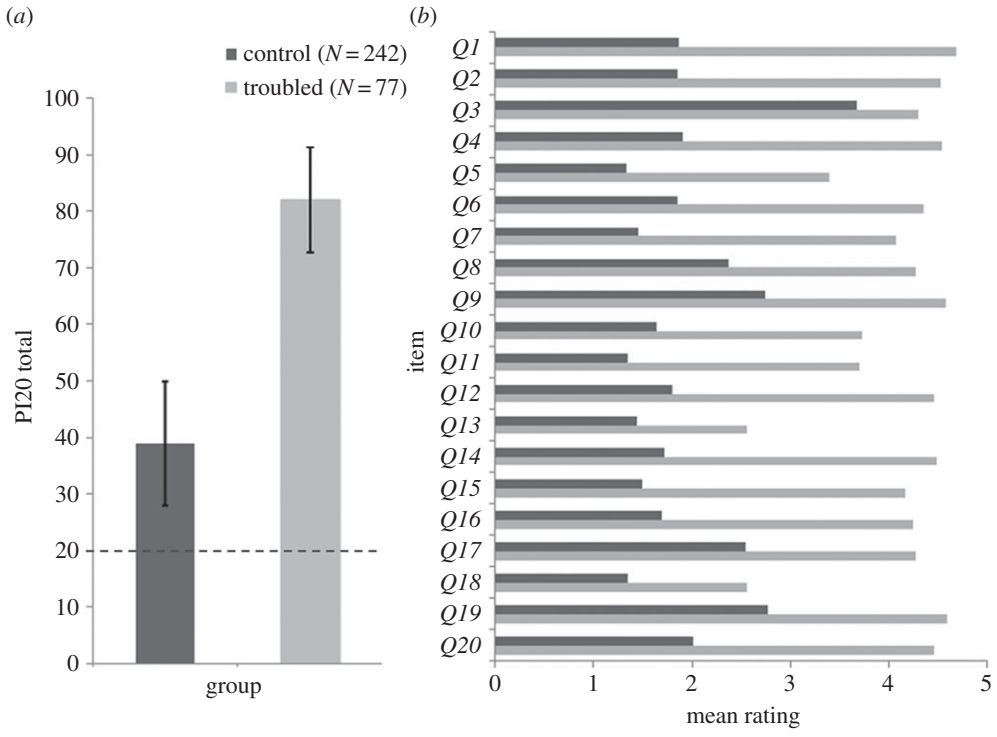

Figure 1. (*a*) Mean PI20 totals observed for the suspected prosopagnosics and TD controls in Validation Study 1. Error bars represent ±1 s.d. The dashed line at 20 represents the minimum possible score. (*b*) Mean responses given by the suspected prosopagnosics and TD controls for each item.

4.1. Participants and methods

Thirty-six participants aged between 20 and 74 years completed Validation Study 2: 18 suspected prosopagnosics ($M_{age} = 46.7$; 12 males), recruited via www.troublewithfaces.org, and 18 matched TD controls ($M_{age} = 43.5$; 12 males), recruited via the local participant database. All participants completed a range of computer-based tests in our laboratory facilities to assess their face and object recognition ability, including the CFMT [26], the CFPT [13], a version of the FFRT [25] and the CCMT [40]. The scores of the DP group and the control group are shown in table 2. Comparing performance on this battery of tasks is commonly used to diagnose DP. All members of the DP group exhibited evidence of impairment on convergent objective tests of face recognition ability.

4.2. Results and discussion

The mean PI20 score of the suspected DPs ($M = 81.22$, s.d. = 9.47) again exceeded that of the controls ($M = 41.67$, s.d. = 12.10), $t_{34} = 10.92$, $p < 0.001$. Of the 18 suspected DPs, 13 scored more than 2.5 s.d. above the control mean (more than or equal to 72), and 12 scored more than 3 s.d. above the control mean (more than or equal to 78). The mean PI20 score of the DP and TD groups in Validation Study 1 and 2 corresponded closely. Importantly, these results confirm that the PI20 distinguishes between prosopagnosics whose deficits have been verified by formal testing and age-matched controls. The results of the Validation Studies 1 and 2 indicate that the PI20 measures the DP construct as it is currently understood. Suspected DPs recognize the experiences and anecdotes contained within the PI20 and the resulting scores afford classification convergent with existing diagnostic procedures. In Validation Studies 3–5, we assessed the relationship between PI20 scores and objective measures of face and object recognition in more detail.

5. Validation Study 3

In our third validation study, we sought to determine whether PI20 scores correlate with respondents' ability to recognize famous faces. Any suggestion that self-report measures can contribute to the

Table 2. The scores achieved on PI20 and other diagnostic tests by each member of the prosopagnosic sample. Scores on the Famous Faces Recognition Test (FFRT), the Cambridge Face Memory Test (CFMT) and Cambridge Car Memory Test (CCMT) reflect % correct. Scores on the Cambridge Face Perception Test (CFPT) reflect total deviation errors.

case	age	gender	PI20	FFRT (%)	CFMT (%)	CCMT (%)	CFPT upright	CFPT inverted
1	20	M	84***	52*	60**	82	50*	60
2	24	M	73**	59	60**	50*	60**	78
3	41	F	73**	73	68*	85	48*	64
4	31	M	71**	31**	56**	65	30	70
5	33	F	87***	46*	60**	74	78***	84*
6	36	M	87***	24**	57**	56*	42*	60
7	42	M	78***	44*	58**	93	52**	50
8	45	M	92***	15***	51***	94	86***	54
9	48	F	78***	30**	58**	86	34	52
10	51	F	85***	42*	46***	64	74***	94*
11	57	M	69**	48*	61**	53*	32	52
12	59	M	97***	3***	49***	82	56**	64
13	67	M	92***	10***	28***	47**	92***	78
14	69	F	95***	32**	36***	76	100***	92*
15	74	M	82***	34**	60**	67	42*	70
16	73	M	68**	30**	53***	72	84***	90*
17	42	M	83***	27**	56**	90	44*	100**
18	29	F	68**	48*	61**	64	32	58
control mean			41.7	74.9	84.3	77.2	29.4	63.3
control s.d.			12.1	17.5	9.9	15.0	10.9	15.6
best control			23	100	99	100	10	36
worst control			65	34	67	47	52	84

Asterisks denote: single (*), differs from control mean 1 s.d.; double (**), differs from control mean 2 s.d.; triple (***), differs from control mean 3 s.d.

classification of DP rests on the assumption that people have insight into their own face recognition ability. If respondents are poor judges of their face recognition ability, high self-report scores may simply reflect respondents' personality; for example, some individuals are known to underestimate their cognitive abilities [41]. Alternatively, strong correlation between PI20 scores and objective measures of face recognition ability would confirm that individuals *do* have insight into their face recognition ability.

5.1. Participants and methods

One-hundred-and-seventy-three of the respondents from Validation Study 1, aged between 18 and 74 years, including 100 TD ($M_{age} = 30.0$; 27 males; 84 UK-based) and 73 suspected DPs ($M_{age} = 42.9$; 28 males; 39 UK-based) participated in Validation Study 3. This sample included participants from Validation Study 2 (see the electronic supplementary material). All participants completed an Internet-based version of the FFRT [25] remotely, during which they had to identify 34 international celebrities (actors, singers, sports stars and politicians), from cropped photographic images, by providing their name or other identifying information. Faces were visible until participants responded. Scores reflect the number of correct identifications expressed as a percentage of the number of celebrities with whom respondents were familiar.

5.2. Results

The suspected DPs identified significantly fewer celebrities ($M = 41.73\%$, s.d. $= 16.58\%$) than controls ($M = 78.28\%$, s.d. $= 14.51\%$), $t_{171} = 15.41$, $p < 0.001$. Importantly, PI20 scores correlated closely with

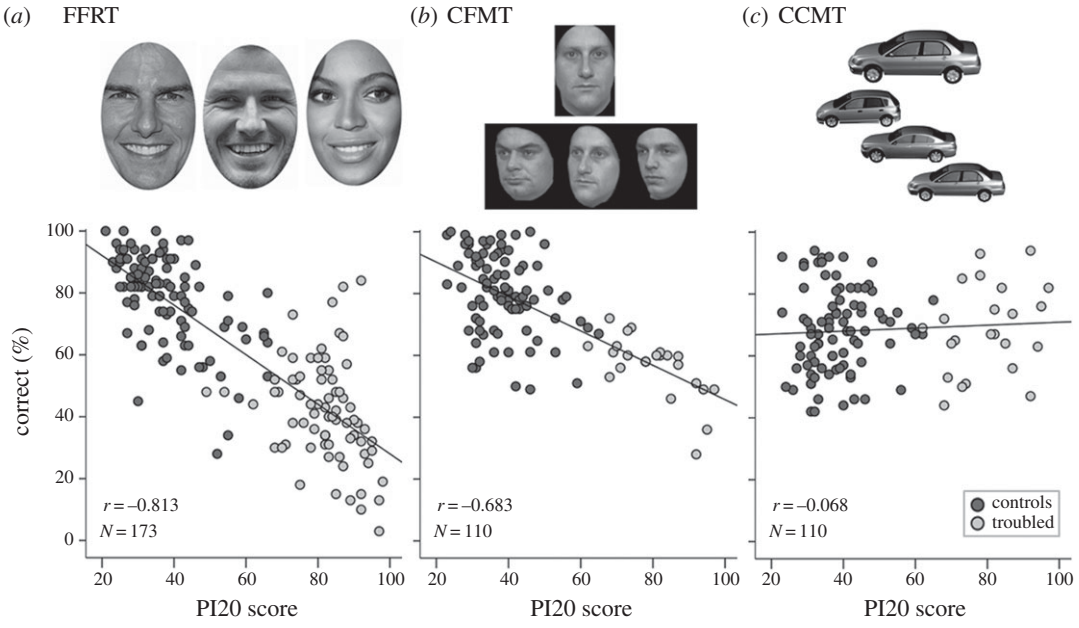

Figure 2. (*a*) The simple correlation observed between PI20 scores and performance on the FFRT in Validation Study 3. (*b*) The simple correlation observed between PI20 scores and performance on the CFMT observed in Validation Study 4. (*c*) The simple correlation observed between PI20 scores and performance on the CCMT observed in Validation Study 5.

participants' famous face recognition ($r = -0.81$, $p < 0.001$; figure 2*a*). To control for the influence of demographic factors, participant age (years), gender (1 = male; 2 = female) and location (1 = based in UK, 2 = based outside UK), were entered into the first step of a hierarchical regression, with PI20 scores entered second. Participant age ($\beta = -0.42$, $p < 0.001$), gender ($\beta = 0.16$, $p = 0.021$) and location ($\beta = -0.17$, $p = 0.015$) were all significant predictors, together accounting for 28.8% of the variance. Crucially, PI20 score remained highly predictive ($\beta = -0.76$, $p < 0.001$), accounting for a further 40.8% of unique variance. These results indicate that PI20 scores correlate with ability to recognize familiar faces.

6. Validation Study 4

Next, we sought to determine whether PI20 scores predict performance on the CFMT [26]. Whereas the FFRT used in Validation Study 3 assesses ability to recognize familiar faces, the CFMT measures ability to match unfamiliar faces, thought to depend on different neurocognitive mechanisms [42–44]. A correlation between PI20 scores and CFMT performance would confirm that respondents have insight into both their familiar and unfamiliar face recognition ability.

6.1. Participants and methods

One-hundred-and-ten participants from Validation Study 1, aged between 18 and 74 years, including 87 TD ($M_{age} = 28.6$ years; 30 males) and 23 suspected prosopagnosics ($M_{age} = 45.8$ years; 15 males) participated in Validation Study 4. A subset of the sample also participated in Validation Studies 2 and 3 (see the electronic supplementary material). All participants were living in the UK at the time of testing. The DP sample contacted the authors via www.troublewithfaces.org. TD participants were recruited through the local participant database. All participants completed the CFMT in our laboratory facilities. The test comprises 72 trials and employs a 3AFC match-to-sample design. Participants first learn a target face in left three-quarters-profile view, frontal view and right three-quarters-profile view. During a subsequent recall phase, participants are required to identify the target in a 3AFC procedure [26].

6.2. Results

The suspected DPs were significantly impaired on the CFMT ($M = 56.42\%$, s.d. = 10.04%) relative to controls ($M = 79.98\%$, s.d. = 13.00%), $t_{108} = 8.07$, $p < 0.001$. Crucially, PI20 scores correlated closely with

participants' CFMT scores ($r = -0.68$, $p < 0.001$; figure 2*b*). Additional hierarchical regression analysis was conducted to control for the influence of participant age (years) and gender (1 = male; 2 = female). When entered in the first step of the model, participant age was predictive ($\beta = -0.40$, $p < 0.001$), but participant gender was not predictive ($\beta = 0.10$, $p = 0.30$), of CFMT performance. Together, these factors accounted for 19.1% of the variance. Importantly, when added to the model, PI20 scores remained highly predictive of CFMT score ($\beta = -0.65$, $p < 0.001$), accounting for a further 27.8% of unique variance. The results of Validation Study 3 indicate that PI20 scores correlate with unfamiliar face recognition. Together, Validation Studies 3 and 4 confirm that people have insight into their face recognition ability.

7. Validation Study 5

The results of Validation Studies 1–4 confirm that PI20 scores correlate with familiar and unfamiliar face recognition. One account of this relationship is that the PI20 measures a relatively specific construct—face recognition ability. However, a second possibility is that the PI20 measures a broader construct (e.g. general memory ability). If the PI20 is measuring a general factor, scores should also correlate with performance on the Cambridge Car Memory Test (CCMT [40]), a well-validated test of non-face object recognition employing an identical format to the CFMT.

7.1. Participants, methods and results

The CCMT was administered to the 110 respondents who participated in Validation Study 4. The car recognition ability of the suspected DPs ($M = 68.64\%$, s.d. = 14.34%) was very similar to that of the TD controls ($M = 68.29\%$, s.d. = 13.56%), $t_{108} = 0.11$, $p = 0.91$. No correlation was observed between participants' PI20 scores and performance on the CCMT ($r = 0.07$; figure 2*c*). Hierarchical regression was conducted to determine whether PI20 scores were predictive of CCMT performance once individual differences in age and gender were controlled for. When entered in the first step, participant age was not predictive of CCMT scores ($\beta = 0.08$, $p = 0.44$), but, in line with previous findings [40], respondent gender was a significant predictor ($\beta = -0.20$, $p = 0.044$), together accounting for 5.6% of the variance. When subsequently added to the regression model, PI20 scores were not predictive ($\beta = -0.04$, $p = 0.71$), accounting for a further 0.1% of unique variance. These results suggest that the PI20 is measuring face recognition ability, and not a general factor.

8. Discussion

This paper introduces the PI20, a 20-item self-report measure of prosopagnosic traits. The new instrument successfully distinguishes suspected DPs from TD adults (Validation Studies 1 and 2). Strong correlations were observed between PI20 scores and performance on objective tests of familiar (Validation Study 3) and unfamiliar face recognition ability (Validation Study 4). Importantly, PI20 scores do not correlate with non-face object recognition (Validation Study 5), indicating that the instrument measures face recognition ability, not a general factor (e.g. wider memory ability).

The results of Validation Studies 3 and 4 confirm that people have the necessary insight into their face recognition ability, required by a self-report instrument. These findings contradict previous suggestions that adults lack insight into their own face recognition ability [27,35]. For example, having asked undergraduates to rate their ability to recognize faces in everyday life 'compared with the average person', Bowles *et al*. [27] found only weak correlations between self-rated ability and performance on the CFMT and CFPT. However, ratings derived from a single question are likely to provide noisy estimates, making weak correlations unsurprising. In addition, items asking about tangible experiences, such as those included in the PI20 (e.g. 'I sometimes find movies hard to follow because of difficulties recognizing characters'), may be less ambiguous than abstract questions about average face recognition ability.

The foregoing results suggest that the PI20 can play a valuable role in the identification of DP, permitting better description of self-report evidence, thereby facilitating greater comparison of prosopagnosic samples. Based on the relationships observed between PI20 scores and performance on objective tests of face recognition, PI20 scores in the ranges 65–74, 75–84, 85–100 may be broadly indicative of mild, moderate and severe DP, respectively. To be clear, we are not suggesting that the PI20 should replace objective tests of face recognition ability; rather we intend the PI20 to be used as a complementary diagnostic instrument. Where PI20 and computer-based tests provide convergent

evidence of impairment, authors can be confident in the composition of prosopagnosic samples. Conversely, where there is discrepancy between objective and self-report evidence, further testing can be undertaken, for example, to determine whether an individual has a severe, lifelong perceptual impairment that they are unaware of, or whether they have simply under-performed on a given task. The use of convergent tasks and complementary paradigms is likely to result in more reliable classification [34].

DP is a heterogeneous condition [4,36,45,46]. The inclusion of self-report measures in diagnostic batteries guards against the possibility that new sub-groups of the DP population go undetected because current computer-based tests are insensitive to their characteristic deficits. For example, current tests require participants to judge static facial images. However, the faces we encounter outside of the laboratory are dynamic [47,48]. Should prosopagnosics exist who have selective problems processing facial motion [49], they may perform within the normal range on current diagnostic tests, despite experiencing face recognition difficulties in their daily lives.

There have been calls for a quick, easy-to-administer instrument for the purposes of screening populations for DP [38]. Computer-based tests are unsuitable for screening large populations. For example, batteries of computerized tasks frequently exceed 45 min in duration and require control groups for interpretation. These factors, together with the expertise and equipment required to administer computer-based tests, limit their clinical and practical utility [30]. Conversely, the PI20 can be completed very quickly in the absence of a computer, by clinicians (e.g. to clients or patients), employers (e.g. to prospective police or border control officers) and judiciary (to eyewitnesses, jurors), to screen populations for DP. Future research directly assessing the PI20's utility as a screening tool in applied contexts will prove informative.

In academic contexts, the instrument may be used both to identify individuals with DP for inclusion in research samples (e.g. screening undergraduate cohorts) and to exclude individuals exhibiting prosopagnosic traits from studies addressing normative face perception. As we have noted, the PI20 is also well suited for use in genetic studies estimating the heritability of the condition. In addition, the availability of a validated self-report measure also permits systematic investigation of the relationship between self-report and objective measures of face recognition ability. For example, future studies might try to better understand which observers are likely to over- or underestimate their actual face recognition ability. Longitudinal studies might also address whether this variability is systematically related to changes in face recognition ability over time.

An increasing number of authors are taking advantage of online platforms to collect behavioural data remotely. While these methods facilitate the collection of large datasets, this trend has provoked considerable discussion about the quality of the data collected [50]. Some readers might therefore query our decision to collect the data reported in Validation Studies 1 and 3 online. A degree of caution is justifiable insofar as current understanding of research conducted online remains relatively limited and best practice continues to evolve. Importantly, however, the findings from these studies were replicated in Validation Studies 2 and 4, respectively, using controlled experimental procedures conducted in the laboratory. Not only do these results confirm that the PI20 can be administered effectively via the Internet—further underscoring its potential value as a screening instrument—but they support the view that online data collection has a valuable role to play in contemporary social perception research [51].

Discussion of the complementary roles played by self-report and objective tests in identification and diagnosis raises fundamental questions about the future of the DP construct. Although DP is not listed as a psychiatric disorder in DSM-5 [23], prosopagnosia is recognized by the World Health Organization [52], and potentially meets the criteria for a mental disability [53]; i.e. a physical or mental impairment that has a 'substantial' and 'long-term' negative effect on one's ability to do normal daily activities. Crucially, however, the extent to which DP impairs *normal daily activities* is not easily assessed with face recognition tests in the laboratory. Any attempt to move DP into mainstream psychiatry may therefore necessitate a broader approach to diagnosis, encompassing self-report.

Ethics. Ethical clearance was granted by the local ethics committee and the study was conducted in accordance with the ethical standards laid down in the 2008 (6th) Declaration of Helsinki. Informed consent was obtained from all participants.

Data accessibility. The dataset supporting this article are available as electronic supplementary material.

Authors' contributions. P.S. and R.C. designed the PI20 and conceived the validation studies, P.S., A.G. and S.S. participated in data collection, P.S., A.G., S.S. and R.C. conducted all statistical analyses, P.S., G.B. and R.C. drafted the manuscript. All authors gave final approval for publication.

Competing interests. We have no competing interests.

Funding. P.S. and S.S. were supported by doctoral studentships from the Medical Research Council.

References

1. Behrmann M, Avidan G. 2005 Congenital prosopagnosia: face-blind from birth. *Trends Cogn. Sci.* **9**, 180–187. (doi:10.1016/j.tics.2005.02.011)
2. Duchaine BC, Nakayama K. 2006 Developmental prosopagnosia: a window to content-specific face processing. *Curr. Opin. Neurobiol.* **16**, 166–173. (doi:10.1016/j.conb.2006.03.003)
3. McConachie HR. 1976 Developmental prosopagnosia. A single case report. *Cortex* **12**, 76–82. (doi:10.1016/S0010-9452(76)80033-0)
4. Susilo T, Duchaine B. 2013 Advances in developmental prosopagnosia research. *Curr. Opin. Neurobiol.* **23**, 423–429. (doi:10.1016/j.conb.2012.12.011)
5. Dalrymple KA, Corrow S, Yonas A, Duchaine B. 2012 Developmental prosopagnosia in childhood. *Cogn. Neuropsychol.* **29**, 393–418. (doi:10.1080/02643294.2012.722547)
6. Kress T, Daum I. 2003 Developmental prosopagnosia: a review. *Behav. Neurol.* **14**, 109–121. (doi:10.1155/2003/520476)
7. Barton JJS. 2008 Structure and function in acquired prosopagnosia: lessons from a series of 10 patients with brain damage. *J. Neuropsychol.* **2**, 197–225. (doi:10.1348/174866407X214172)
8. Duchaine BC, Parker H, Nakayama K. 2003 Normal recognition of emotion in a prosopagnosic. *Perception* **32**, 827–838. (doi:10.1068/p5067)
9. Humphreys K, Avidan G, Behrmann M. 2007 A detailed investigation of facial expression processing in congenital prosopagnosia as compared to acquired prosopagnosia. *Exp. Brain Res.* **176**, 356–373. (doi:10.1007/s00221-006-0621-5)
10. Duchaine B, Murray H, Turner M, White S, Garrido L. 2009 Normal social cognition in developmental prosopagnosia. *Cogn. Neuropsychol.* **26**, 620–634. (doi:10.1080/02643291003616145)
11. Todorov A, Duchaine B. 2008 Reading trustworthiness in faces without recognizing faces. *Cogn. Neuropsychol.* **25**, 395–410. (doi:10.1080/02643290802044996)
12. Dobel C, Bölte J, Aicher M, Schweinberger SR. 2007 Prosopagnosia without apparent cause: overview and diagnosis of six cases. *Cortex* **43**, 718–733. (doi:10.1016/S0010-9452(08)70501-X)
13. Duchaine B, Germine L, Nakayama K. 2007 Family resemblance: ten family members with prosopagnosia and within-class object agnosia. *Cogn. Neuropsychol.* **24**, 419–430. (doi:10.1080/02643290701380491)
14. Lee Y, Duchaine B, Wilson HR, Nakayama K. 2010 Three cases of developmental prosopagnosia from one family: detailed neuropsychological and psychophysical investigation of face processing. *Cortex* **46**, 949–964. (doi:10.1016/j.cortex.2009.07.012)
15. Johnen A, Schmukle SC, Hüttenbrink J, Kischka C, Kennerknecht I, Dobel C. 2014 A family at risk: congenital prosopagnosia, poor face recognition and visuoperceptual deficits within one family. *Neuropsychologia* **58**, 52–63. (doi:10.1016/j.neuropsychologia.2014.03.013)
16. Dalrymple KA, Fletcher K, Corrow S, das Nair R, Barton JJS, 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. (doi:10.1016/j.jpsychores.2014.06.001)
17. Davis JM, McKone E, Dennett H, O'Connor KB, O'Kearney R, Palermo R. 2011 Individual differences in the ability to recognise facial identity are associated with social anxiety. *PLoS ONE* **6**, e28800. (doi:10.1371/journal.pone.0028800)
18. Fine DR. 2012 A life with prosopagnosia. *Cogn. Neuropsychol.* **29**, 354–359. (doi:10.1080/02643294.2012.736377)
19. 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. (doi:10.1016/j.jpsychores.2008.03.013)
20. Kennerknecht I, Nga YH, Wong VCN. 2008 Prevalence of hereditary prosopagnosia (HPA) in Hong Kong Chinese population. *Am. J. Med. Genet. A* **146**, 2863–2870. (doi:10.1002/ajmg.a.32552)
21. 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**, 1617–1622. (doi:10.1002/ajmg.a.31343)
22. Gainotti G. 2010 Not all patients labeled as 'prosopagnosia' have a real prosopagnosia. *J. Clin. Exp. Neuropsychol.* **32**, 763–766. (doi:10.1080/13803390903512686)
23. American Psychiatric Association. 2013 *Diagnostic and statistical manual of mental disorders* (*DSM -5*). Washington, DC: American Psychiatric Association.
24. Duchaine BC, Nakayama K. 2004 Developmental prosopagnosia and the Benton Facial Recognition Test. *Neurology* **62**, 1219–1220. (doi:10.1212/01.WNL.0000118297.03161.B3)
25. Sanders HI, Warrington EK. 1971 Memory for remote events in amnesic patients. *Brain* **94**, 661–668. (doi:10.1093/brain/94.4.661)
26. 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. (doi:10.1016/j.neuropsychologia.2005.07.001)
27. 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. (doi:10.1080/02643290903343149)
28. Germine LT, Duchaine B, Nakayama K. 2011 Where cognitive development and aging meet: face learning ability peaks after age 30. *Cognition* **118**, 201–210. (doi:10.1016/j.cognition.2010.11.002)
29. McKone E, Hall A, Pidcock M, Palermo R, Wilkinson RB, Rivolta D, Yovel G, Davis JM, O'Connor KB. 2011 Face ethnicity and measurement reliability affect face recognition performance in developmental prosopagnosia: evidence from the Cambridge Face Memory Test–Australian. *Cogn. Neuropsychol.* **28**, 109–146. (doi:10.1080/02643294.2011.616880)
30. Butcher JN, Perry JN, Atlis MM. 2000 Validity and utility of computer-based test interpretation. *Psychol. Assess.* **12**, 6–18. (doi:10.1037/1040-3590.12.1.6)
31. Grüter T, Grüter M, Bell V, Carbon CC. 2009 Visual mental imagery in congenital prosopagnosia. *Neurosci. Lett.* **453**, 135–140. (doi:10.1016/j.neulet.2009.02.021)
32. 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. (doi:10.1016/j.cortex.2010.08.005)
33. 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. (doi:10.1023/A:1005653411471)
34. Morey LC. 2013 Measuring personality and psychopathology. In *Handbook of psychology, volume 2, research methods in psychology* (eds IB Weiner, JA Schinka, W Velicer), pp. 395–427, 2nd edn. Hoboken, NJ: Wiley.
35. Laguesse R *et al*. 2013 Subjective self-assessment of face recognition ability is only weakly related to objective measures of face recognition performance. *J. Vis.* **13**, 979. (doi:10.1167/13.9.979)
36. Stollhoff R, Jost J, Elze T, Kennerknecht I. 2011 Deficits in long-term recognition memory reveal dissociated subtypes in congenital prosopagnosia. *PLoS ONE* **6**, e15702. (doi:10.1371/journal.pone.0015702)
37. Della Sala S, Schweinberger SR. 2013 Face blindness and person misidentification in non-scientific parlance. *Cortex* **49**, 2276–2280. (doi:10.1016/j.cortex.2013.05.001)
38. Fine DR. 2011 We need a simple test for prosopagnosia. *BMJ* **342**, d1736. (doi:10.1136/bmj.d1736)
39. Rivolta D. 2014 *Prosopagnosia*. Berlin, Germany: Springer.
40. Dennett HW, 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**, 587–605. (doi:10.3758/s13428-011-0160-2)
41. Furnham A. 2001 Self-estimates of intelligence: culture and gender difference in self and other estimates of both general (g) and multiple intelligences. *Pers. Indiv. Diff.* **31**, 1381–1405. (doi:10.1016/S0191-8869(00)00232-4)
42. Bruce V, Young A. 1986 Understanding face recognition. *Br. J. Psychol.* **3**, 305–327. (doi:10.1227/01.NEU.0000334050.45750.C9)
43. Jenkins R, Burton AM. 2011 Stable face representations. *Phil. Trans. R. Soc. B* **366**, 1671–1683. (doi:10.1098/rstb.2010.0379)
44. Shah P, Gaule A, Gaigg SB, Bird G, Cook R. 2015 Probing short-term face memory in developmental prosopagnosia. *Cortex* **64**, 115–122. (doi:10.1016/j.cortex.2014.10.006)
45. Cohan S, DeGutis J. 2014 A proposal for developmental prosopagnosia 'sub-types' based on

differential face perception and face memory performance. *J. Vis.* **14**, 1433–1433. (doi:10.1167/14.10.1433)

46. Le Grand R, Cooper PA, Mondloch CJ, Lewis TL, Sagiv N, de Gelder B, Maurer D. 2006 What aspects of face processing are impaired in developmental prosopagnosia? *Brain Cogn.* **61**, 139–158. (doi:10.1016/j.bandc.2005.11.005)
47. Hill H, Johnston A. 2001 Categorizing sex and identity from the biological motion of faces. *Curr. Biol.* **11**, 880–885. (doi:10.1016/S0960-9822(01)00243-3)
48. O'Toole AJ, Roark DA, Abdi H. 2002 Recognizing moving faces: a psychological and neural synthesis. *Trends Cogn. Sci.* **6**, 261–266. (doi:10.1016/S1364-6613(02)01908-3)
49. Longmore CA, Tree JJ. 2013 Motion as a cue to face recognition: evidence from congenital prosopagnosia. *Neuropsychologia* **51**, 864–875. (doi:10.1016/j.neuropsychologia.2013.01.022)
50. Gosling SD, Vazire S, Srivastava S, John OP. 2004 Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *Am. Psychol.* **59**, 93–104. (doi:10.1037/0003-066X.59.2.93)
51. Germine L, Nakayama K, Duchaine BC, Chabris CF, Chatterjee G, Wilmer JB. 2012 Is the web as good as the lab? Comparable performance from web and lab in cognitive/perceptual experiments. *Psychon. Bull. Rev.* **19**, 847–857. (doi:10.3758/s13423-012-0296-9)
52. World Health Organization. 1992 *The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research*. Geneva, Switzerland: World Health Organization.
53. Lockwood G, Henderson C, Thornicroft G. 2012 The Equality Act 2010 and mental health. *Br. J. Psychiatry* **200**, 182–183. (doi:10.1192/bjp.bp.111.097790)