



## King's Research Portal

DOI:

[10.1098/rsos.150305](https://doi.org/10.1098/rsos.150305)

*Document Version*

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Shah, P., Sowden, S., Gaule, A., Catmur, C., & Bird, G. (2015). The 20 item prosopagnosia index (PI20): relationship with the Glasgow face-matching test. *Royal Society open science*, 2(11), 150305.  
<https://doi.org/10.1098/rsos.150305>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



**Cite this article:** Shah P, Sowden S, Gaule A, Catmur C, Bird G. 2015 The 20 item prosopagnosia index (PI20): relationship with the Glasgow face-matching test. *R. Soc. open sci.* **2**: 150305.  
<http://dx.doi.org/10.1098/rsos.150305>

Received: 30 June 2015

Accepted: 8 October 2015

**Subject Category:**

Psychology and cognitive neuroscience

**Subject Areas:**

psychology/neuroscience/cognition

**Keywords:**

developmental prosopagnosia, self-report, Glasgow face-matching test, face matching, prosopagnosia, face perception

**Author for correspondence:**

Punit Shah

e-mail: [punit.shah@kcl.ac.uk](mailto:punit.shah@kcl.ac.uk)

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsos.150305> or via <http://rsos.royalsocietypublishing.org>.

# The 20 item prosopagnosia index (PI20): relationship with the Glasgow face-matching test

Punit Shah<sup>1,3</sup>, Sophie Sowden<sup>1</sup>, Anne Gaule<sup>4</sup>,  
Caroline Catmur<sup>2,6</sup> and Geoffrey Bird<sup>1,5</sup>


<sup>1</sup>MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

<sup>2</sup>Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, University of London, London, UK

<sup>3</sup>Department of Psychological Sciences, Birkbeck College, London, UK

<sup>4</sup>Institute of Neurology and <sup>5</sup>Institute of Cognitive Neuroscience, University College London, London, UK

<sup>6</sup>Department of Psychology, University of Surrey, Surrey, UK

 CC, 0000-0001-6996-4295; PS, 0000-0001-5497-4765

The 20 item prosopagnosia index (PI20) was recently developed to identify individuals with developmental prosopagnosia. While the PI20's principal purpose is to aid researchers and clinicians, it was suggested that it may serve as a useful screening tool to identify people with face recognition difficulties in applied settings where face matching is a critical part of their occupation. Although the PI20 has been validated using behavioural measures of face *recognition*, it has yet to be validated against a measure of face-*matching* ability that is more representative of applied settings. In this study, the PI20 was therefore administered with the Glasgow face-matching test (GFMT). A strong correlation was observed between PI20 and GFMT scores, providing further validation for the PI20, indicating that it is likely to be of value in applied settings.

## 1. Introduction

Face recognition is critical for social interaction, and has therefore been the subject of intensive scientific investigation. Studying individuals who have a selective impairment in face recognition in the absence of brain injury—those with developmental prosopagnosia (DP)—has contributed to understanding the mechanisms underlying typical face recognition (see [1]). Less attention, however, has been directed to studying DP within the context of individual differences in face recognition

ability and the practical implications of this condition on wider society. This is partly because identifying DP and quantification of its severity has not been standardized in the way in which it has for other developmental disorders (e.g. autism—autism diagnostic observation schedule; [2]). This lack of standardization is an obstacle to the development of formal diagnostic criteria, and until DP is better recognized among clinicians it is less likely to be considered in non-clinical settings [3].

To begin addressing these issues, the 20 item prosopagnosia index (PI20 [4]) was developed; a self-report measure designed to help identify adults with DP and, more generally, to quantify face recognition difficulties in the general population. The PI20's psychometric properties, as previously reported, are good, and it has good construct validity. It is therefore able to successfully distinguish individuals known to have DP from typical individuals. Furthermore, scores on the PI20 predict face recognition (e.g. Cambridge face memory test, CFMT [5]) abilities, but are not related to performance on the Cambridge car memory test [6]. This pattern of results suggests that the PI20 indexes face recognition ability specifically, rather than providing a measure of general object recognition ability. The reliability and validity of the PI20 suggest that it may benefit researchers interested in face perception, as well as clinicians who encounter DP or co-occurring DP in their practice. In a research context, the PI20 will be useful for classifying people with DP for inclusion in research samples and to exclude such individuals from studies of typical face perception. In clinical contexts, the PI20 may prove to be particularly useful during differential diagnosis in tertiary psychiatric units where co-occurring prosopagnosia is most likely to be reported [7].

The PI20 can be administered quickly—without specialist equipment or training—but its potential for use in applied settings is currently unknown. Although validated against the CFMT, widely regarded as the 'gold-standard' measure of face recognition ability within DP research, the CFMT may not measure the type of face processing skills required by, for example, passport control officials and police officers. First, the CFMT has been argued to conflate measurement of face memory and face recognition [8–10]. Second, the test format is—by design—not reflective of real-world situations; it includes only male faces from which hairlines have been cropped. Third, it is not representative of occupational settings where face *matching* is required, such as comparing two images taken on different cameras (e.g. photo-to-photo matching from surveillance footage and police booking photographs) or the photo-to-person matching required at passport control, for example.

In recognition of the foregoing issues (see also [11–14]), Burton *et al.* devised the Glasgow face-matching test (GFMT [9]); a psychometric instrument designed for applied face-matching research, with a view to using the measure in security settings (see Methods for detailed information). The GFMT has been used in several applied studies (see [15]), for example with passport control officers [16]. To investigate whether the PI20 is suitable for use in applied settings, this study, therefore, sought to investigate the relationship between scores on the PI20 and performance on the GFMT.

## 2. Method

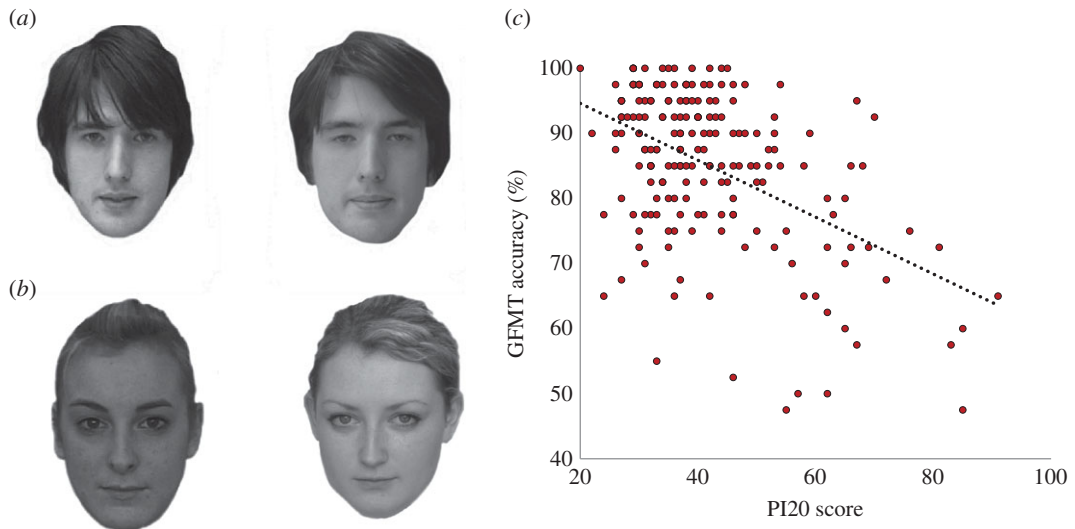
### 2.1. Participants

One hundred and ninety adults (59 males; 15 left-handed) aged between 18 and 75 years ( $M_{\text{age}} = 27.72$  years,  $s.d. = 10.66$ ) participated in the study. The sample size was chosen to match that used by Burton *et al.* [9] when they constructed the GFMT. In addition, power analysis [17] indicated that this sample size would be sufficient to detect a moderate sized correlation (where  $r = 0.30$ , power = 0.95,  $\alpha = 0.05$ ) between the GFMT and PI20. Participants only reported their age, sex, handedness and all had normal or corrected-to-normal vision. Participants were recruited using a local participant database and gave informed consent prior to their participation. Ethical clearance was granted by the local ethics committee and this study will be reported in line with recommendations from Simmons *et al.* [18].

### 2.2. Materials and procedure

The PI20 [4] (see the electronic supplementary material) required participants to indicate the extent to which 20 statements described their face recognition difficulties on a five-point Likert scale (strongly agree to strongly disagree), from which a score between 20 (no face recognition difficulties) and 100 (severe face recognition difficulties) was derived.

The abbreviated version of the GFMT [9] was used. This comprised 40 pairs of faces. All faces were of neutral expression and comprised high-resolution (1000 × 700) greyscale images. To reflect real-world face matching, hairlines were intact, 40% of face pairs were female, and photos within each



**Figure 1.** Examples of (a) same and (b) different-face trials from the short version of the Glasgow face-matching test (GFMT [9]). The GFMT is freely available from <http://www.facevar.com/downloads> and should be credited to Burton *et al.* [9] if it is used in published research. People who appear in the GFMT gave written consent for their images to be used in the context of academic research and publication. (b) The simple correlation between 20-item prosopagnosia index (PI20 [4]) scores and performance on the GFMT ( $r = -0.49$ ,  $p < 0.001$ ).

pair were taken using two different cameras. Of the 40 test pairs, 20 are same-face trials, in which two images of the same person are presented alongside one another (figure 1a). The other 20 pairs comprised different-face trials (figure 1b) of people that were similar in appearance. The two images are presented simultaneously in order to eliminate memory demands and participants were required to indicate whether the two images were the same or different. In accord with Burton *et al.*'s instructions, the GFMT was administered on a computer screen at 100% scale and self-paced, which took participants approximately 7 min ( $M = 7.6$  min,  $s.d. = 3.0$  min) to complete. The order of administration of the PI20 and GFMT was counterbalanced across participants.

### 3. Results

The mean PI20 score ( $M = 41.99$ ,  $s.d. = 13.40$ ) was consistent with the original validation study [4], the PI20's internal consistency (Chronbach's  $\alpha = 0.93$ ; Guttman split-half correlation = 0.91) was good, and its unifactorial structure was found to be robust. Performance on the GFMT ranged between 47.5% and 100.0% ( $M = 85.00$ ,  $s.d. = 11.92$ ). Signal detection analysis [19] was conducted to measure discrimination performance ( $d'$ ) independent of response bias. Reflecting percentage accuracy,  $d'$  scores ranged between  $-0.25$  and  $3.29$  ( $M = 2.27$ ,  $s.d. = 0.85$ ).

PI20 scores correlated significantly with percentage accuracy ( $r = 0.49$ ,  $p < 0.001$ ; figure 1c) and  $d'$  scores ( $r = 0.47$ ,  $p < 0.001$ ) on the GFMT. To control for the influence of demographic factors and time taken to complete the GFMT on face-matching performance, participant age (years), sex (1, male; 2, female) and completion time (seconds) were entered into the first step of a hierarchical regression, with PI20 scores entered second. Participant age ( $\beta = -0.14$ ,  $p = 0.051$ ), sex ( $\beta = 0.21$ ,  $p = 0.004$ ) and completion time ( $\beta = 0.23$ ,  $p = 0.001$ ) were all predictors of GFMT performance, together accounting for 12.8% of the variance. Female participants outperformed ( $M = 87.08$ ,  $s.d. = 10.92$ ) male participants ( $M = 80.38$ ,  $s.d. = 12.81$ ,  $t_{188} = 3.70$ ,  $p < 0.001$ ,  $d = 0.56$ ), whereas older age was weakly associated with less accurate GFMT performance ( $r = -0.15$ ,  $p = 0.040$ ). Importantly, however, the PI20 score remained the strongest predictor ( $\beta = -0.45$ ,  $p < 0.001$ ), accounting for a further 19.2% of unique variance. This pattern of significance was identical when  $d'$  scores were entered into analyses as the measure of GFMT performance.

## 4. Discussion

The PI20 was recently designed to support research and clinical practice involving DP, which guided its validation using well-established measures of face *recognition* ability. Because the PI20 also holds potential for use in applied settings, we sought to examine the relationship between the PI20 and the GFMT; a psychometric tool designed to measure *face-matching* ability in applied settings.

Following previous work [4], this paper demonstrates that the PI20 is a robust self-report measure as its psychometric properties hold when administered to a large sample of new participants. In addition, this study supports the claim that individuals have insights into their face perception ability. The literature investigating this question has produced equivocal findings to date, with some studies reporting correlations between confidence judgements and face-matching performance and others reporting no relationship [20–22]. It is possible that previous findings of a lack of a correlation between confidence and accuracy are due to the use of single-item questions designed to be answered with respect to a single time-point. These measures may produce noisier estimates of confidence than those obtained through administration of the PI20, whereby estimates are derived from responses to multiple questions, and participants are given the opportunity to estimate their average confidence based on a longer time period (i.e. their lifetime). It is also possible that within-task measures of confidence prompt participants to estimate their confidence on a given trial with respect to their own average performance, rather than with respect to other individuals as measured by the PI20, and that participants are able to do the latter but not the former. Both these possibilities are worthy of further investigation.

There was substantial inter-individual variation in GFMT performance, in line with that found by Burton *et al.* [9,16], supporting the suggestion that such tests should be administered when selecting personnel for occupations requiring face matching. As has been previously observed [23], including recently using a modified version of the GFMT [24], females outperformed males. The size of the sex effect was, however, small, particularly when compared with the large individual differences observed in face-matching performance in the sample as a whole. The relationship between age and GFMT performance was also small, though marginally significant, in line with a recent study on age-related decline in unfamiliar face matching [25]. Most importantly, with respect to the aim of this study was the finding that the PI20 was the strongest predictor of face-matching performance over and above that accounted for by demographic variables. Because a strong relationship was observed between PI20 score and performance on the GFMT, this provides further construct validation for the PI20 and indicates its potential for use in applied settings.

### 4.1. Outstanding issues, practical limitations and future directions

The majority of items on the PI20 enquire about *familiar* face recognition, yet scores on this measure predict performance on the GFMT, a measure of *unfamiliar* face matching. This is of interest owing to ongoing debate about the degree to which processing of familiar and unfamiliar faces recruit common or distinct neurocognitive mechanisms [26]. Behavioural research on this topic provides strong evidence that (upright) familiar faces recruit processes which are not recruited for unfamiliar faces (see [27] for an excellent overview of this evidence), although this research also demonstrates that familiar and unfamiliar faces may recruit some common, face-specific, processes [28]. Investigation of the neural correlates of face processing has provided mixed evidence for a dissociation between the processing of familiar and unfamiliar faces however, with many studies finding overlapping neural correlates of familiar and unfamiliar face processing [27,29]. The difference in the *magnitude* of the correlations between the PI20 and various objective measures of face perception—in this study and those reported in Shah *et al.* [4]—support the theoretical distinction between familiar and unfamiliar faces. As the PI20 focuses on recognition of familiar faces, PI20 scores show a greater ( $Z = 2.44$ ,  $p = 0.015$ ) correlation with scores on the famous face recognition test ( $r = -0.813$ ) than with scores on the CFMT ( $r = -0.683$ ), which assesses recognition of recently viewed novel faces. With respect to separable recognition and matching processes, the correlation between the PI20 and unfamiliar face matching (GFMT performance) is significantly weaker than that between PI20 scores and CFMT performance,  $Z = -2.45$ ,  $p = 0.014$ , and famous face recognition,  $Z = -5.65$ ,  $p < 0.001$ . While this pattern of results lends some support to the distinction between familiar and unfamiliar face perception, and between recognition and matching, it should be noted that this study was not designed to address this question and that comparisons are made across distinct participant samples.

Three factors should be considered before the use of an instrument such as the PI20 in applied settings. The first is that, in order to gain or retain employment, it is possible that individuals may complete self-report measures in a manner which they believe will enhance their prospects, rather than accurately reflecting their true ability. Second, the distribution of PI20 scores was—in accordance with Shah *et al.* [4]—skewed towards lower scores (figure 1c), reflective of low prosopagnosic traits in the population. The PI20 is therefore particularly sensitive at identifying individuals with DP, but relatively less able to identify individuals with exceptionally good face-matching ability—for which objective measures of face matching are more appropriate. Third, although the PI20 has good predictive validity at the group level, a non-trivial number of participants report good face processing ability on the PI20 yet perform poorly on the GFMT. This highlights the worrying tendency for some individuals to substantially overestimate their face-matching ability, which could, for example, have important implications in legal settings. The PI20 will therefore enable future research into why individuals under- or overestimate their face-matching ability, perhaps in relation to individual differences in personality (see [30]).

In the light of the above-mentioned issues, we suggest that the PI20 should *not* be considered as an alternative to objective measures of face-matching ability; rather it is hoped the PI20 will be used as a complementary instrument. Wherever possible, the PI20 should be administered with behavioural measures (e.g. GFMT) to form a detailed perceptual profile of an individual. This is important, because many factors—such as misinterpretation of instructions or test anxiety—may lead someone to perform badly on behavioural measures of face matching. This is particularly likely when they are not administered by trained experimenters in applied settings. Where PI20 and behavioural test scores converge, it is likely that users (e.g. employers) can be more confident of their validity. A direct test of the combined predictive power of the PI20 and objective face perception scores on ecologically valid face-matching tasks (e.g. photo-to-person tests [16]) is therefore a priority for future research.

## 5. Conclusion

In summary, this study indicates that the PI20 holds potential for use in applied settings, particularly when combined with an objective test of face matching such as the GFMT. It is hoped that research leading to the development of the PI20 will be extended by occupational and forensic research groups, using participants recruited within applied settings, and in field research involving photo-to-person matching.

**Ethics.** Ethical clearance was granted by the local ethics committee, all participants granted informed consent and the study was conducted in line with the ethical standards laid down in the declaration of Helsinki.

**Data accessibility.** The dataset supporting this article is available as electronic supplementary materials.

**Authors' contributions.** P.S. and S.S. conceived of the study and participated in data collection, P.S., S.S., C.C. and G.B. conducted statistical analyses, P.S., S.S., C.C. and G.B. drafted the manuscript. All authors gave final approval for publication and edited the manuscript.

**Competing interests.** We have no competing interests.

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

## References

- Susilo T, Duchaine B. 2013 Advances in developmental prosopagnosia research. *Curr. Opin. Neurobiol.* **23**, 423–429. (doi:10.1016/j.conb.2012.12.011)
- Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, Pickles A, Rutter M. 2000 The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J. Autism Dev. Disord.* **30**, 205–223. (doi:10.1023/A:1005592401947)
- Fine DR. 2011 We need a simple test for prosopagnosia. *Brit. Med. J.* **342**, d1736. (doi:10.1136/bmj.d1736)
- 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. (doi:10.1098/rso.s.140343)
- 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)
- 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)
- Kitamura H, Egawa J, Someya T. 2015 Developmental prosopagnosia referred to outpatient psychiatric service. *Psychiatry Clin. Neurosci.* **69**, 238. (doi:10.1111/pcn.12233)
- Jenkins R, Burton AM. 2011 Stable face representations. *Phil. Trans. R. Soc. B* **366**, 1671–1683. (doi:10.1098/rstb.2010.0379)
- Burton AM, White D, McNeill A. 2010 The Glasgow face matching test. *Behav. Res. Methods* **42**, 286–291. (doi:10.3758/BRM.42.1.286)
- Burton M. 2013 Why has research in face recognition progressed so slowly? The importance of variability. *Q. J. Exp. Psychol.* **66**, 1467–1485. (doi:10.1080/17470218.2013.800125)
- Megreya AM, Burton AM. 2006 Unfamiliar faces are not faces: evidence from a matching task. *Mem. Cognit.* **34**, 865–876. (doi:10.3758/BF03193433)
- Jenkins R, White D, Van Montfort X, Mike Burton A. 2011 Variability in photos of the same face. *Cognition*

- 121, 313–323. (doi:10.1016/j.cognition.2011.08.001)
13. Megreya AM, Burton AM. 2008 Matching faces to photographs: poor performance in eyewitness memory (without the memory). *J. Exp. Psychol. Appl.* **14**, 364–372. (doi:10.1037/a0013464)
  14. Bruce V, Henderson Z, Greenwood K, Hancock PJB, Burton AM, Miller P. 1999 Verification of face identities from images captured on video. *J. Exp. Psychol. Appl.* **5**, 339–360. (doi:10.1037/1076-898X.5.4.339)
  15. Valentine T, Davis JP (eds). 2015 *Forensic facial identification*. Chichester, UK: John Wiley & Sons, Ltd.
  16. White D, Kemp RI, Jenkins R, Matheson M, Burton AM. 2014 Passport officers' errors in face matching. *PLoS ONE* **9**, e103510. (doi:10.1371/journal.pone.0103510)
  17. Faul F, Erdfelder E, Lang A-G, Buchner A. 2007 G\* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191. (doi:10.3758/BF03193146)
  18. Simmons JP, Nelson LD, Simonsohn U. 2011 False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* **22**, 1359–1366. (doi:10.1177/0956797611417632)
  19. Green D, Swets A. 1966 *Signal detection theory*. New York, NY: Wiley.
  20. 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**, 979–979. (doi:10.1167/13.9.979)
  21. 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)
  22. Davis JP, Valentine T. 2009 CCTV on trial: matching video images with the defendant in the dock. *Appl. Cogn. Psychol.* **23**, 482–505. (doi:10.1002/acp.1490)
  23. Lewin C, Herlitz A. 2002 Sex differences in face recognition: women's faces make the difference. *Brain Cogn.* **50**, 121–128. (doi:10.1016/S0278-2626(02)00016-7)
  24. Megreya AM, Bindemann M, Hvard C. 2011 Sex differences in unfamiliar face identification: evidence from matching tasks. *Acta Psychol.* **137**, 83–89. (doi:10.1016/j.actpsy.2011.03.003)
  25. Megreya AM, Bindemann M. 2015 Developmental improvement and age-related decline in unfamiliar face matching. *Perception* **44**, 5–22. (doi:10.1068/p7825)
  26. Burton AM, Schweinberger SR, Jenkins R, Kaufmann JM. 2015 Arguments against a configural processing account of familiar face recognition. *Perspect. Psychol. Sci.* **10**, 482–496. (doi:10.1177/1745691615583129)
  27. Burton AM, Jenkins R. 2011 Unfamiliar face perception. In *The Oxford handbook of face perception* (eds AJ Calder, G Rhodes, JV Haxby, MJ Johnson), pp. 287–306. Oxford, UK: Oxford University Press.
  28. Johnston RA, Edmonds AJ. 2009 Familiar and unfamiliar face recognition: a review. *Memory* **17**, 577–596. (doi:10.1080/09658210902976969)
  29. Natu V, O'Toole AJ. 2011 The neural processing of familiar and unfamiliar faces: a review and synopsis. *Br. J. Psychol.* **102**, 726–747. (doi:10.1111/j.2044-8295.2011.02053.x)
  30. Megreya AM, Bindemann M. 2013 Individual differences in personality and face identification. *J. Cogn. Psychol.* **25**, 30–37. (doi:10.1080/20445911.2012.739153)