Faces of Merseyside: Exploring cognitive bias through facial averages

Caroline Wilkinson*, Stenton Mackenzie and Kathryn Smith. Email: c.m.wilkinson@ljmu.ac.uk.

Abstract
Faces of Merseyside is a gallery/online exhibition of digitally processed facial averages produced from Merseyside image collections by Face Lab, a research group in Liverpool School of Art & Design. The project sought to foreground the question of cognitive bias in relation to facial images that claim to represent particular communities, in the context of a resurgence of interest in physiognomic judgements and discrimination. By revisiting Galton's 19th century 'composite portraiture', as informed by current craniofacial research, *Faces of Merseyside* explores the claims advanced in relation to the representation of human diversity, and how they both inform and challenge social stereotyping.

Introduction

The 7th International Conference on The Image was hosted in Liverpool in 2016, under the theme of *Face Value: Personification and Identity in a Post-Digital Age*. In response to this theme, Face Lab produced *Faces of Merseyside (FoM)*, a public exhibition of digital facial averages constructed from photographic collections sourced within the Merseyside region.

FoM addressed the importance of the face in identity, character and social groups from a creative research perspective. The project was conceived as a polemic to engage conversation around the *perceived* differences and similarities between people at a time when implicit and explicit discrimination was finding expression in relation to the global 'migrant crisis' more broadly, and the Brexit referendum more locally. Our objective was to explore attitudes towards social cohesion in relation to concepts of national identity and internationalism.

As a contextual frame for the exhibition, this article considers the controversial historical context in which average/composite faces were produced to demonstrate or support ideas regarding human character, and ways in which such images have been employed more recently, informed by cognitive psychology research, revealing face perception bias. Finally, we consider the FoM visitor feedback data and what facial averaging may reveal about our contemporary attitudes towards individual and collective identities, and how such images reflect or contradict popular perception.

Face Research

The human face possesses special significance in terms of its biological and social roles [1] with classification categories at different levels of identity specificity, such as sex/gender, ethnicity, age, attractiveness and distinctiveness. Much of our interest in human faces is centred upon our own perception of them as 'beautiful' [2]; faces perceived as more beautiful tend towards symmetry and less symmetrical faces are often associated with undesirable character traits.

Recognising the wide range of abilities associated with human face processing – the so-called 'super recognisers' and prosopagnosics representing two extremes of a spectrum – face perception remains a primary entry point to human social interaction. Research in the field of psychology has found facial averaging to be a particularly useful tool, allowing testing of observer responses to different appearances, providing a comparative approach to how our minds work through behaviour and causality, and lends itself to analysing how we respond to faces as 'perceptual gestalts' [3]. However, difficulties arise when our tendency to interpret this information transcends the face and we attempt to link behavioural traits and affiliations to appearance, particularly when this is influenced by ideological agendas.

The tendency to judge human character via the head and its attributes, facial appearance included, can be traced back to Ancient Greece, Rome and China [4] and became consolidated as physiognomy, a highly popular and influential 18th century doctrine championed by Johann Kaspar Lavater (1741-1801). Following this now discredited exploration of the relationship between form and appearance, the English polymath Francis Galton (1822–1911) [5], utilised photography to explore human variation, a project which led him down the path of eugenics. Galton devoted many years to the use of 'composite portraiture', discovering that a minimum of 20 facial images produce a 'true' average [6], a technique he applied in his search to represent a 'type' of person by superimposing multiple photographs to produce a single blended visual average [7] (see Fig. 1).

Figure 1: Francis Galton's composite photographs of three sisters, C19th
https://commons.wikimedia.org/wiki/File:Galton_composite_photography_of_sisters.jpg
Image courtesy of http://galton.org/editor.html

Galton developed the technique using thousands of photographs of criminals [8]. Influenced perhaps by contemporary ideas of criminal atavism and diatheses [9], he was interested in

whether criminality could be identified from facial features. Yet Galton noted that, in his facial composites, "*the special villainous irregularities have disappeared and the common humanity that underlies them has prevailed. They represent, not the criminal, but the man that is liable to fall into crime*" [10;11]. A number of other researchers have attempted similar criminal averages with the same limited success [12]. Despite the discrediting of physiognomy and Galton's ideological desires [13], recent media reports have resurrected these interests, with some technology companies offering commercial face analysis services for business and industry [14]. Why does our interest in associating certain character traits with facial appearance persist?

Unexpectedly, the scientific value of some of Galton's ideas have been verified more than a century later. Psychologists have long known that when presented with novel faces, our brains seem to automatically extract the typical face, which we conceptualise as a prototype for a set of faces we have seen [15]. Galton anticipated this, believing *"that all of humanity's general notions of type issue from such a process, the impression together of individual instances of the same sort in the mind",* a phenomenon which he sought to demonstrate with the photographic composite [16]. Further, psychology researchers have demonstrated enhanced recognition when averaged faces are utilised for passports [17], age progressions and eye witness descriptions [18]. Galton also discovered that the resulting face was more attractive than any of the contributing individual faces, a finding that has been repeated many times [19]. Average faces have been used by modern psychologists to study a number of human characteristics including attractiveness [20], gender [21], facial preference [22] and elements of personality, such as trustworthiness [23;24;25].

Artists and scientists have both exploited these technologies to study globalisation and socioeconomic disparities. Burson used the technique to demonstrate the structural biases of social and political power [26] while Kimyonghun [27] reflected on cinematic 'identity' as embodied by a film's tonal values and central character. One web-based project *The Face of Tomorrow* used facial averages to present "an open-source web-based exploration of human identity as affected by the forces of globalization" [28], which inspired the University of Glasgow to extend this work, using data from forty-one countries in their attempt to capture current global ethnic group demographics [29]. Badino created average faces from US sports leagues [30] revealing the consistency across different teams in the same sport, but also tacitly demonstrating the appeal and/or opportunity in different sports across ethnic groups. Badino also created averages of Democrats and Republicans from the US House of Representatives to demonstrate the gender bias in US politics along with the racial disparity between political parties at that time.

Ellenbogen has suggested that facial composites provide a pictorial synthesis of data in contrast to the abstracted representations of graphs and similar diagrammatic abstractions [31]. This ability to 'embody' a data set within a single, relatable image is perhaps why scientists have found this method so attractive, securing the continuing application of the technique in modern research. These images also demonstrate cognitive bias, confirming or refuting commonly-held beliefs, echoing Abraham Lincoln's famous quip, *"The Lord prefers common looking people. That is the reason he made so many of them.'"* [32]

We evaluate faces we are presented with in as little as 38 milliseconds, processing multiple qualitative traits [33;34;35], even when exposed to an emotionally neutral face. This is barely enough time to register the face, yet these rapid judgements might have significant social consequences, ranging from electoral success [36;37;38] to legal decisions [39;40]. This is cognitive bias in action. Research also highlights cognitive bias by the public in relation to immigration [41;42], and the public and police in relation to drug use and other crimes [43;44;45;46]. These studies suggest that cognitive bias will play a part in our views relating to local demographics, culture, interests and occupations.

Facial averaging has been used to explore the cues that produce character-based impressions, with the results unfortunately reinforcing some social and gender-based stereotypes – a neutral male face is deemed less trustworthy than a smiling female face, whereas a smiling female is less dominant than middle-aged neutral male face [47]. The signals our faces project function to attract or discourage others, and in order to do this the cues we use to infer intentions and abilities must be culturally consistent [48]. However, as Todorov reminds us, reading masculinity as a universal signifier of dominance was taken for granted until it was recently shown to be unambiguous only in industrialised cultures [49], suggesting that there is *"little support for evolved honest signals of character in the face"* [50].

Benson and Perrett [51] note,

> *"Formation of category prototypes has been shown to be useful for recognising objects at the psychological level [...] by providing a level of abstraction which seeks to maximise the amount of information about a category with the least cognitive effort. In a sense the averaging process .... also maximises the information about the category to which the face belongs."*

In theory, this process facilitates decisions based upon perceived normality/typicality, and captures the subtleties of age, sex/gender, attractiveness, ethnicity or health. Average faces should represent the relative distinctiveness of groups within and across populations and by contrast, such abstractions should allow predictions to be made about the appearance and veracity of characteristics belonging to one or another classification.

However, researchers are unable to discover the cues informing these rapid visual judgements if they do not vary in the faces used to construct the experimental models designed to test them [52]. Might we then infer that it is impossible to glean anything truly meaningful from average faces, with their generic features, ageless and flattened appearance?

Faces of Merseyside

Contemporary and archival images were gathered from groups representing different occupations and interests, including university students and staff, performers, health practitioners, sports players, politicians, 1960s passport applicants, criminals and police officers, as well as the average male and female adult and child from Liverpool (see Figs 2-4).

Figure 2: Averages faces of academics and soccer players. Left to right: Female and male academics, male Everton FC, male Liverpool FC. (© Caroline Wilkinson)

For each average, a minimum of twenty faces, standardised for size and resolution, were 'annotated' within Abrosoft Fantamorph5, delineating face shape and internal features in order, to orientate each face in relation to the next. This process eliminates asymmetry and other irregularities and blemishes, thus producing smooth, homogenous faces that are utterly benign. Since beauty and symmetry are strongly correlated, these faces project a sense of aesthetic balance. In contrast, external features and any accoutrements are rendered as misty, swirling layers that radiate from the edges of these beatific countenances. The overall effect is spectral, with a concentration of detail centrally and a fading of detail peripherally.

Figure 3: Averages faces of police officers and criminals. Left to right: Female and male police officers, female and male criminals. (© Caroline Wilkinson)

Figure 4: Averages faces of Liverpool. Left to right: Female adult, male adult, female child, male child. (© Caroline Wilkinson)

The exhibition included a visitor feedback opportunity in the form of a questionnaire (this could also be accessed at the online version of the exhibition) designed to collect the audience response to the images in relation to preconceptions and representation. Analysis of visitor responses revealed that visitors found the faces more similar than they expected, but a significant number of people did not believe that our faces reflect our occupations. These results appear contradictory and may highlight the dichotomy between logical thought and cognitive bias [53]; for example, even a very basic scientific understanding of identical twins will lead to the conclusion that facial appearance is primarily influenced by genetic inheritance, this is contradicted by our emotional attachment to the concept that it might be possible to predict personality from facial appearance. Alternatively, this may indicate awareness that certain occupations/interests are more accessible to some ethnic groups than others, creating the expectation that average faces may differ between occupations/interests due to environmental factors.

Conclusion

FoM represents a hybrid form of social research that attempts to understand whether we are able to recognise differences and similarities between people via digital facial averages. Considered alongside the resurrection of this eighteenth-century pseudo-science in the contemporary media, does this appear to encourage potential opportunities for prejudice?

We continue to insist on the face having a compelling and meaningful status within our social lives. The very principles of physiognomy should be refuted by the simple fact that we process faces as perceptual gestalts, none of our impressions are reducible to single features. Yet facial profiling is given a *"veneer of legitimacy"* because we have a natural propensity to form character impressions when exposed to faces; impressions on which we share some agreement [54]. Todorov productively reframes the conversation about binary differences in favour of a more nuanced, culturally relative model, such as 'global face properties' with both empirical and less empirical aspects of our appearance in relation to our 'mental face sphere' (the memory of the faces we have been exposed to during our lives). This is the 'visual facial diet' that shapes what we perceive as typical or atypical and which, in turn, shapes our impressions [55]. Conceivably, those with a limited 'mental face sphere' may be more susceptible to negative biases of the cross-race/cultural effect.

Todorov distinguishes between hypothesis-led, theory-driven and data-driven approaches [56] and highlights the ability of one particular data-driven technique to *"…make invisible mental representations visible"* under controlled experimental conditions [57]. This 'suspicious perception' can demonstrate hidden cognitive biases by progressively altering neutral 'base face' images with noise masks, and asking participants to classify the resulting faces based on their perceptions and associations. Such experiments have demonstrated that *"members of different European nations see the 'typical' European face as more similar to the typical face of their own nation. […] Our biases shape the pictures in our heads"* [58]. This observation has particular salience here.

All told, the results reveal more about the expectations of those interpreting the images than about the data itself. First impressions are constructed from multiple and highly variable visual

cues that are more or less active in different contexts. The methods available to study these need to be understood for their darker legacies and limits, but also their opportunities to hold up a mirror to what contemporary communities 'look like', and how this might affect our participation across different social, professional and cultural communities.

Acknowledgments

References and Notes

1. S. Buchala, N. Davey, T.M. Gale and R.J. Frank, "Principal component analysis of gender, ethnicity, age, and identity of face images," Proceedings of the IEEE International Conference on Multimodal Interaction 7 (2005).

2. F.B. Naini. *Facial aesthetics: concepts & clinical diagnosis* (UK: Wiley-Blackwell, 2011).

3. R.E. Nisbett and T.D. Wilson, "Telling more than we can know: Verbal Reports on Mental Processes," Psychological Review Vol. 84, No. 3, 231-259 (1977).

4. D. McNeill. *The face: a natural history* (New York: Back Bay Books, 2000).

5. N.W. Gillham. *A Life of Sir Francis Galton: from African exploration to the birth of eugenics* (New York: Oxford University Press, 2001).

6. F. Galton, "Composite Portraits," Nature May 23, 97-100 (1878).

7. J. Ellenbogen. *Reasoned and Unreasoned Images: The Photography of Bertillon, Galton and Marey* (Pennsylvania: Pennsylvania State University Press, 2012) p.8.

8. A. Todorov. *Face Value: The Irresistible Influence of First Impressions* (Princeton and Oxford: Princeton University Press, 2017) p. 23, pp.150-151.

9. C. Lombroso [1878]. *Criminal Man* (New York: GP Putnam, 1911)

10. Todorov [8] p. 25.

11. F. Galton, "Eugenics: Its definition, scope and aims," American Journal of Sociology Vol. 10, No.1, 1-25 (July 1904).

12. *Drug Mugs: The Average Faces of Drug Abuse*, http://www.recovery.org/learn/average-faces-of-drug-abuse/#, accessed 31 October 2017.

13. A.A. Oosterhof and A. Todorov, "The functional basis of face evaluation," Proceedings of the National Academy of Sciences Vol. 105, No. 32, 11087-11092 (2008).

14. Faception – Face Personality Analytics https://www.faception.com, accessed on 2 November 2017

15. J.H. Langlois and L.A. Roggman, "Attractive faces are only average," Psychological Science Vol. 1, No.2, 115-121 (1990).

16. Ellenbogen [7] p. 8.

17. A.M. Burton, R. Jenkins, P.J. Hancock and D. White, D. "Robust representations for face recognition: The power of averages," Cognitive Psychology Vol. 51, No. 3, 256-284 (2005) and R. Jenkins and A.M. Burton, "100% Accuracy in Automatic Face Recognition," Science Vol. 319, No. 5862, 435-435 (2008).

18. V. Bruce, H. Ness, P.J. Hancock, C. Newman and J. Rarity, "Four Heads Are Better Than One: Combining Face Composites Yields Improvements in Face Likeness," Journal of Applied Psychology, Vol. 87, No. 5, 894 (2002).

19. Langlois and Roggman [15]

20. G. Rhodes and T. Tremewan, "Averageness, Exaggeration and Facial Attractiveness," Psychological Science 7; 105-110 (1996).

21. P.J. Benson & D.I. Perrett, "Extracting prototypical facial images from exemplars," Perception Vol. 22, No. 3, 257-262 (1993).

22. L. Fruhen, C.D. Watkins and B.C. Jones, "Perceptions of Facial Attractiveness, Dominance and Trustworthiness Predict Managerial Pay Awarded in Experimental Tasks," Leadership Quarterly (2015).

23. R. J. Evenhouse, M. Rasmussen and L. L. Sadler, "Computer-Aided Forensic Facial Reconstruction", Biostereometric Technology and Applications, Vol. 1380, 147-157, International Society for Optics and Photonics (April 1991).

24. A. Todorov, S.G. Baron, & N.N. Oosterhof. Evaluating face trustworthiness: A model based approach. Social Cognitive & Affective Neuroscience, 3(2), 119-127 (2008)

25. A. Todorov. Evaluating faces on trustworthiness. An extension of systems for recognition of emotions signaling approach/avoidance behaviors. Annual N.Y. Academy of Science, 1124, 208-224 (2008)

26. N. Burson (artist website) http://nancyburson.com/composite-silver-prints, accessed 10 October 2017.

27. Shinseungback Kimyonghun (artist website) http://ssbkyh.com/works/portrait, accessed 31 October 2017.

28. M. Mike. The Face of Tomorrow http://leoalmanac.org/gallery/gx/face/index.htm, accessed 31 October 2017.

29. Edmonds, L. Meet the world's Mrs Averages: Scientists blend thousands of faces together to reveal what the typical woman's face looks like in 41 different countries from around the globe. Mail Online, 27 September 2013 <http://www.dailymail.co.uk/news/article-2435688/The-average-woman-revealed-Study-blends-thousands-faces-worlds-women-look-like.html>, accessed 31 October 2017.

30 H. Badino. The Average Face, http://www.lelaps.de/projects/avgface.html, accessed 31 October 2017

31. Ellenbogen [7], p.9

32. T. Dennett. *Lincoln and the Civil War in the diaries and letters of John Hay* (New York, Da Capo Press, 1988) p.143.

33. J. Willis and A. Todorov, "First impressions: Making up your mind after 100ms exposure to a face," Psychological Science Vol. 17, No. 7, 592–598 (2006).

34. M. Bar, M. Neta and H. Linz, "Very first impressions," Emotion Vol. 6, No.2, 269–278 (2006).

35. C.C. Carbon. The first 100 milliseconds of a face: On the microgenesis of early face processing. Perceptual and Motor Skills, 113(3), 859-874. doi:10.2466/07.17.22.Pms.113.6.859-874 (2011)

36. A. Todorov, A.N. Mandisodza, A. Goren and C.C. Hall, "Inferences of competence from faces predict election outcomes," Science Vol. 308, No. 5728, 1623-1626 (2005).

37. C.C. Ballew and A. Todorov, "Predicting political elections from rapid and unreflective face judgments," Proceedings of the National Academy of Sciences 104(46), 17948-17953 (2007).

38. A.C. Little, R.P. Burriss, B.C. Jones, S.C. Roberts, "Facial appearance affects voting decisions," Evolution and Human Behavior 28(1), 18-27 (2007).

39. I.V. Blair, C.M. Judd and K.M. Chapleau, "The influence of Afrocentric facial features in criminal sentencing," Psychological Science Vol. 15, No. 10, 674–679 (2004).

40. J.L Eberhardt, P.G. Davies, V.J. Purdie-Vaughns and S.L. Johnson, "Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes," Psychological Science Vol. 17, No. 5: 382–386 (2006).

41. S. Blinder, M. Ruhs and C. Vargas-Silva. "Thinking behind the numbers: Understanding public opinion on immigration in Britain," Migration Observatory report, COMPAS, University of Oxford (2011) 2011.

42. S. Bechhofer & D. McCrone, "Being British: A crisis of identity?" The Political Quarterly, Vol. 78, No. 2, 251-260 (2007).

43. Criminal Justice Act 1991, Section 95
http://www.legislation.gov.uk/ukpga/1991/53/section/95, accessed 27 September 2010

44 N. Eastwood, M. Shiner and D. Bear, "The numbers in black and white: ethnic disparities in the policing and prosecution of drug offences in England and Wales," Drugs, The Law & Human Rights (2013)

45. C. Sharp and T. Budd. "Minority ethnic groups and crime: findings from the Offending, Crime and Justice Survey 2003," Home Office (2005), accessed 7 October 2010.

46 J. Hales; C. Nevill; S. Pudney and S. Tipping, "Longitudinal analysis of the Offending, Crime and Justice Survey 2003–06," Home Office (November 2009), accessed 27 September 2010.

47 Todorov [8] p.128

48 Todorov [8] p.187

49 Todorov [8] p.143

50 Todorov [8] p.187

51 P.J. Benson and I.D. Perrett, "Extracting prototypical facial images from exemplars," Perception, Vol. 22, No. 3, 257-262 (1993).

52 Todorov [8] p.129

53 J.S.Lerner, Y. Li, P. Valdesolo and K.S. Kassam, "Emotion and Decision Making," Annual Review of Psychology Vol. 66 (2015)

54 Todorov [8] p.27

55 Todorov [8] p.133

56 Todorov [8] p.103

57 Todorov [8] p.111

58 Todorov [8] p.107