University of Wollongong

Research Online

Faculty of Science, Medicine and Health - Papers: Part B

Faculty of Science, Medicine and Health

1-1-2020

Likeness, Familiarity, and the Ambient Portrait Average

Susan Hayes
University of Wollongong, suehayes@uow.edu.au

Peter Caputi
University of Wollongong, pcaputi@uow.edu.au

T Zaracostas

Maggie Henderson

Julie Telenta

See next page for additional authors

Follow this and additional works at: https://ro.uow.edu.au/smhpapers1

Publication Details Citation

Hayes, S., Caputi, P., Zaracostas, T., Henderson, M., Telenta, J., McCombe, E., Christopher, K., Calvert, E., Abbati, D., Smith, O., Medwell, E., & Wilcock, J. (2020). Likeness, Familiarity, and the Ambient Portrait Average. Faculty of Science, Medicine and Health - Papers: Part B. Retrieved from https://ro.uow.edu.au/smhpapers1/1289

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Likeness, Familiarity, and the Ambient Portrait Average

Abstract

© The Author(s) 2020. This artist-led research project involved 10 visual artists producing 10 ambient portraits and a portrait average of a locally familiar Sitter, and 10 ambient portraits and a portrait average of a less locally familiar Sitter. All were then assessed for likeness by more than 150 members of the general public attending an exhibition during Australia's 2018 National Science Week. The results of this study are that portrait averages can be highly shape accurate and tend to be seen as a good likeness by all viewers. However, the portrait average is not necessarily the best likeness. Extending and validating our previous findings regarding the relationship of likeness, familiarity, and shape accuracy (as measured using geometric morphometrics) in portraiture, unfamiliar viewers favouring shape accurate depictions of a Sitter attained statistical significance. Familiar viewers, however, although also tending to view shape accurate depictions a good to very good likeness, were shown to have a stronger preference for portraits that exaggerate a Sitter's facial distinctiveness, including an exaggeration of their head pose, providing such exaggerations are in approximate proportional agreement.

Publication Details

Hayes, S., Caputi, P., Zaracostas, T., Henderson, M., Telenta, J., McCombe, E., Christopher, K., Calvert, E., Abbati, D., Smith, O., Medwell, E. & Wilcock, J. (2020). Likeness, Familiarity, and the Ambient Portrait Average. Perception,

Authors

Susan Hayes, Peter Caputi, T Zaracostas, Maggie Henderson, Julie Telenta, Elspeth McCombe, Kim Christopher, Emma Calvert, Donna Abbati, Odette Smith, Emma Medwell, and Joyce Wilcock

This journal article is available at Research Online: https://ro.uow.edu.au/smhpapers1/1289

Perception

https://journals.sagepub.com/doi/10.1177/0301006620905420

Likeness, Familiarity and the Ambient Portrait Average

Susan Hayes[1,2*], Peter Caputi[3], T.S. Zaracostas[2], Maggie Henderson[2], Julie Telenta[2], Elspeth McCombe[2], Kim Christopher[2], Emma Calvert[2], Donna Abbati[2], Odette Smith[2], Emma Medwell[2], Joyce Wilcock[2]

1. *Centre for Archaeological Science, School of Atmospheric, Earth and Environmental Sciences, SMAH, University of Wollongong Australia*
2. *Red Point Artists' Association, Port Kembla, Australia*
3. *School of Psychology, Faculty of Social Sciences, University of Wollongong Australia*

*corresponding author

Abstract

This artist-led research project involved 10 visual artists producing 10 ambient portraits and a portrait average of a locally familiar Sitter, and 10 ambient portraits and a portrait average of a less locally familiar Sitter. All were then assessed for likeness by more than 150 members of the general public attending an exhibition during Australia's 2018 *National Science Week.* The results of this study are that portrait averages can be highly shape accurate, and tend to be seen as a good likeness by all viewers. However, the portrait average is not necessarily the best likeness. Extending and validating our previous findings regarding the relationship of likeness, familiarity and shape accuracy (as measured using geometric morphometrics) in portraiture, unfamiliar viewers favouring shape accurate depictions of a Sitter attained statistical significance. Familiar viewers, however, although also tending to view shape accurate depictions a good to very good likeness, were shown to have a stronger preference for portraits that exaggerate a Sitter's facial distinctiveness, including an exaggeration of their head pose, providing such exaggerations are in approximate proportional agreement.

Keywords:

face perception, recognition, likeness, distinctiveness, portraiture, geometric morphometrics

Accepted Manuscript

Introduction

During Australia's 2017 *National Science Week* (12-19 August), we undertook a pilot study involving an exhibition of 12 portraits depicting the same Sitter which were assessed for likeness by 108 volunteers. The results were promising regarding the role of facial shape depiction, abstraction and exaggeration in likeness assessments of portraits produced under relatively naturally occurring (ambient) conditions (Hayes et al., 2018). To assess the accuracy of the 12 portraits we applied geometric morphometrics (statistical shape analysis) which resulted in the statistical average of the 12 portraits being the most accurate for depicting the Sitter's facial shapes. However, because this average portrait was produced after the exhibition had closed, we had no opportunity to gather objective verification or refutation that it also appeared to be a very good portrait likeness.

Studies involving naturally occurring and diverse photographic portraits have found that averaging these ambient images results in comparatively low likeness/recognisability ratings from familiar assessors (Ritchie, Kramer, & Burton, 2018), even though they tend to increase the speed of familiar face recognition (for reviews, see Jenkins & Burton, 2011; Young & Burton, 2017). Averaging multiple photographs of the same individual has also been found to increase the accuracy of unfamiliar face matching (White, Burton, Jenkins, & Kemp, 2014), which the authors suggest is likely due the averaging process reducing the idiosyncratic 'noise' of individual photographs (e.g. variation in head pose and lighting conditions). Similarly, averaging low resolution images mimicking poor quality of CCTV stills enhances the accuracy of face matching with both unfamiliar assessors and automated face matching software (Ritchie, White, et al., 2018). Facial averaging has also included laboratory-based studies involving averages created from multiple facial composites (i.e. 'fotofits' of selected facial parts) depicting the same target face. These studies have shown a clear advantage of the average over the likeness of individual facial composites, whether this involves the recognition of target faces of familiar celebrities (Bruce, Ness, Hancock, Newman, & Rarity, 2002) or unfamiliar faces following a one minute exposure to a photograph (Hasel & Wells, 2007).

During Australia's 2018 *National Science Week* (11-19 August), we held a follow-up exhibition to both extend and validate our previous study while addressing the likeness and shape accuracy of a portrait average. In 2017, our analyses of the relationship between

Accepted Manuscript

portrait likeness and shape accuracy was dominated by unfamiliar face-matching. This arose because we had failed to take into account that while the Sitter, Nick Rheinberger, is a well-known local ABC Illawarra Radio *Mornings* presenter, for most of the assessors their prior familiarity did not extend to his facial appearance. For this iteration of the study, therefore, the exhibition contained portraits depicting two Sitters: the Wollongong Lord Mayor, Gordon Bradbery AM, who is a very well-known face in the Illawarra region of New South Wales, and Professor Gordon Wallace, New South Wales' Scientist of the Year (2017-2018) and a research participant in our 2017 study. The exhibition was promoted as *Portraits of Gordon*^{2} (see Figure 1), and involved 10 Artists each producing two portraits – one of each Sitter – resulting in 20 portraits (Figures 2-3) and two portrait averages (Figure 4). Because all of the exhibited portraits were – as in 2017 – anonymous as to who produced them, each of the portrait averages was able to be exhibited as a portrait in its own right, alongside the 10 individual artworks that comprised them.

Figure 1. **Promotion of the *Portraits of Gordon*^{2} *exhibition***. Image adapted from Professor Wallace's life sitting session. Photograph inserts are the two reference photographs used for the portrait production.

Portraits of Gordon^2 attracted over 200 visitors to the Red Point Artists' Association (RPAA) Gallery. Of these, 153 visitors were actively willing to assess the likeness of both sets of exhibited portraits, with 111 reporting prior familiarity with the Lord Mayor's facial appearance. Of these, 60% were moderate to highly familiar, enabling a more nuanced and statistically sound study of the impact of different levels of familiarity on portrait likeness assessments (i.e. no familiarity, some familiarity, moderate-high familiarity). In contrast, only

Accepted Manuscript

36 assessors reported at least some familiarity with Professor Gordon Wallace's facial appearance prior to attending the exhibition, which is similar to the 2017 visitor cohort and enables comparison with, and validation of, the 2017 study findings. Although we again focus exclusively on portrait shape accuracy in this follow up study (an experimental analysis of texture is forthcoming), some modifications were instituted to improve both the process and the outcomes. These changes are presented in more detail the Methods, and include likeness assessments of the reference photographs used for the creation of the portraits, additional analyses regarding head pose, and more detail concerning the Artists' production of the portraits.

To summarise, the overall aims of this study were to:

(i) examine the shape accuracy and likeness of ambient portrait averages,

(ii) see whether our earlier findings regarding portrait likeness assessments would differ with a more facially familiar Sitter (Mayor Bradbery), and/or

(iii) see if our findings would be replicated with a Sitter with similarly low levels of prior familiarity as the 2017 study (Professor Wallace).

Methods

Recruitment of Artist Collaborators

Once both Sitters had confirmed their willingness to participate in the research project, an EOI was forwarded via email to the RPAA membership. This resulted in 10 Artist research collaborators, all of whom are women. Three were Artist collaborators on the 2017 project, and as with our previous study, the Artists were active participants in, and aware of, the research aims, design and intended outcomes, though the anonymity regarding which Artist produced what portrait was maintained throughout the exhibition and subsequent analyses.

Production of the Portraits

In 2017, the portraits were produced in reference to a life-size monochromatic 3D print of the Sitter's head and face, which is not how portraits are typically produced under ambient conditions (see for example the process described in Wisely & Fine, 1997). For this study the portrait production commenced with life sittings to meet the Sitters and undertake preliminary sketches. Each Sitter attended the RPAA Gallery on different days for approximately 1.5 hours, and each were seated approximately 2m from the Artists, who were

Accepted Manuscript

arrayed in a semi-circle (see Figure 1). To further clarify the visibility of the Sitters' facial features, each Sitter had their head and face side-lit with tungsten lights. The Sitters initially talked about their respective professional roles to enable the Artists to become familiar with the Sitters' animated faces, and to allow the Sitters to become familiar with the close visual scrutiny of the Artists surrounding them. Posed drawing sessions were for approximately 10 minutes, followed by informal interactions while the Artists rotated their easel positions.

Digital photographs of the Sitter were taken during each life sitting, including a series of frontal photographs approximately 2m from the Sitters. The Sitters assumed their own head pose, which is understood within portraiture to be characteristic of a person's depicted likeness (Faigin, 1990; Maughan, 2004; Speed, 1917). The Sitters were, however, requested to avoid open-mouthed smiles as the depiction of teeth is notoriously difficult for even experienced portrait artists. At the end of the life sitting each Sitter reviewed the series of frontally orientated images and selected those they most preferred/least disliked. From this selection, and in keeping with research that has found people are poor selectors of their own photographic likenesses (White, Burton, & Kemp, 2016), the photograph that became the reference for the portraits was the one which the Artists considered to be the most representative of the Sitter's facial appearance.

In Mayor Bradbery's selected reference photograph the head is depicted slightly turned and canted (tilted) towards the left shoulder. Professor Wallace is also shown with a slight head turn to the left, but canted towards the right shoulder and with an upwards head pitch (see insert photographs, Figure 1). A left head turn contracts the horizontal widths of the left hemiface and expands those on the right, while head canting has relatively minimal impact on facial shapes. An upwards head pitch, however, widens and lengthens the jaw, lowers the mouth corners, presents an upturned nose, and contracts both the upper face and eye spacing (Hayes, 2010; Hayes & Milne, 2011; Hayes & Tullberg, 2012).

All Artists based their portraits on the selected Sitter reference photographs, and all were aware that this was to reduce variance in their individual depictions of the Sitters' head orientations, which will typically dominate a geometric morphometric analysis of facial morphology (Hayes & Milne, 2011; Hayes & Tullberg, 2012). The reference photographs were distributed to the Artists both electronically and as photographic prints, together with a selection of images taken from different angles during each life sitting. All Artists also took

Accepted Manuscript

their own photographs during the life sittings, and all were able to access a wide range of additional photographs of both Sitters on the Internet. One Artist (portrait I) was unable to attend the life sitting for Mayor Bradbery, but was already familiar with his facial appearance in life, and another (portrait S) was unable to attend Professor Wallace's life sitting, and instead viewed interviews and presentations on the internet (e.g. www.youtube.com/watch?v=HUauycEDRQU) to supplement the still images.

Because our previous study had found no significant impact of hue, medium or scale on the likeness judgements of portraits (Hayes et al., 2018), the Artists completed their works using the materials, media and dimensions of their choosing. The Artists had four weeks after the life sittings to complete and submit both portraits to RPAA on a set day to DA, a member of the research team responsible for the collation of all Artist, portrait and likeness assessment data, but not an Artist participant. On receipt of the portraits, each was allocated a randomized numerical code that evaded any indication of which two portraits each Artist had produced, and each Artist's age (in decades), handedness, prior experience in portrait depiction (1 none – 5 very), the hours they spent producing each portrait, and their prior familiarity with each Sitter (1 none – 5 very) was recorded.

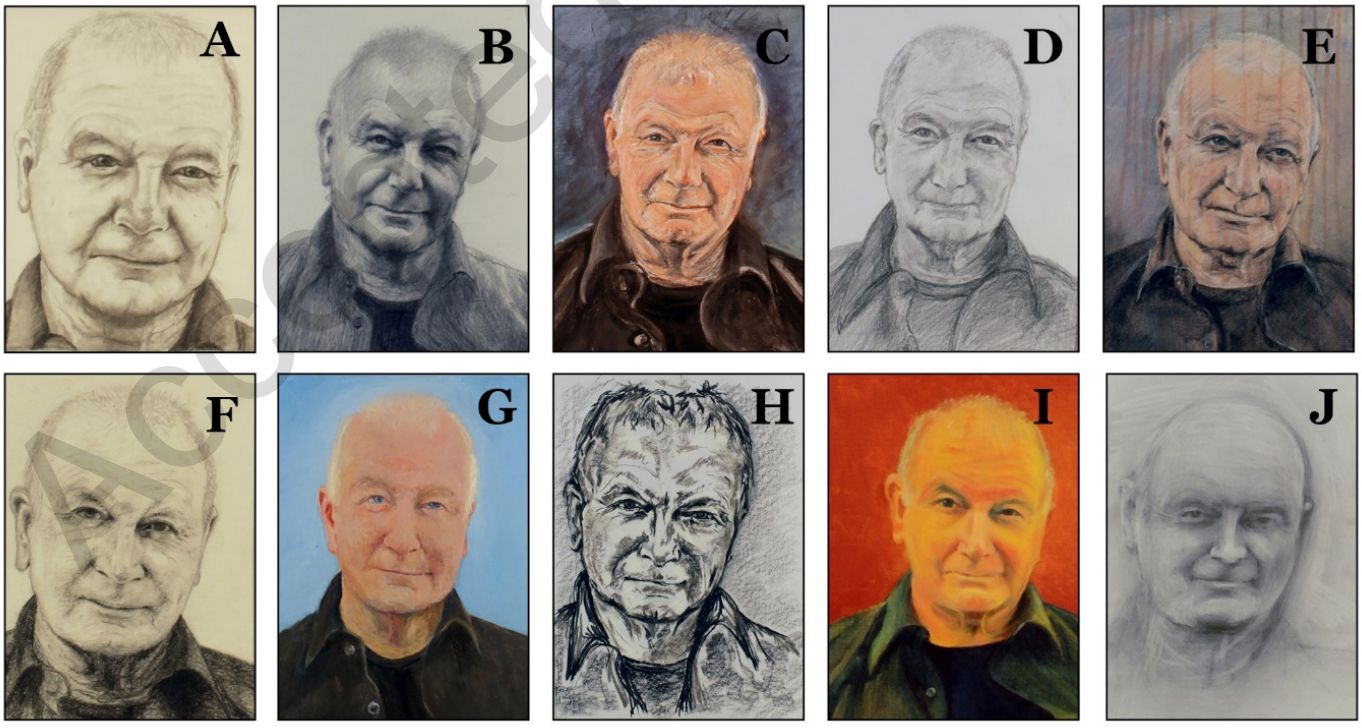


Figure 2. **Portraits of Mayor Gordon Bradbery**, in exhibition order A-J.

Accepted Manuscript

Once all of the portraits had been delivered and coded, each was photographed in the RPAA Gallery using a Canon EOS 60D mounted on a tripod at a distance of 1.5m, with the selion (the deepest part of the soft tissue nasal bridge) as the focal point. To facilitate the identification of homologous landmarks for the geometric morphometric analyses, each digital portrait image and the Sitters' reference photographs were entered into *Adobe Photoshop CC 2018*, and rotated until a horizontal guide intersected both the right and left exocanthi (outer eye corners). The portrait images were then scaled to be equidistant with the corresponding Sitter's reference photograph between the right exocanthion and the stomion (centre of where the lips of the mouth meet). All images were cropped to head and shoulders and output as high resolution TIFFS (no compression). Note that Figures 2-3 show each of the works in their original, unscaled orientations.

Figure 3. **Portraits of Professor Gordon Wallace**, in exhibition order M-V.

Accepted Manuscript

Production of the Portrait Averages

High resolution images of the portrait averages were achieved using the geometric morphometric thin plate spline (tps) software created by James Rohlf (Rohlf, 2015). Digitising the x,y landmark coordinates (*tpsDig v232*), checking the suitability of the overall

variance (*tpsSmall v32*) and creating the statistical average based on these landmark coordinates (*tpsSuper v32*) are described in our 2017 study (Hayes et al., 2018), though for this iteration involved a larger, and different, suite of landmark coordinates were applied (Supplementary Information, Figures 1-2, Tables 1-2). Furthermore, the two Sitters' image data sets required different landmark coordinates to adequately capture their depicted morphological variance.

Production of each of the portrait averages required the creation of two statistical averages – one following the 77 head and face landmark coordinates, and one additionally capturing the outer corners of each cropped portrait (landmarks 78-81). The latter were used to orient the statistical average of the head and face to the average of the portrait image outer edges, and to include an average background using *Adobe Photoshop CC 2018*. The resulting portrait averages were output as TIFFs (no compression), commercially printed as A3 colour photographs with a matt finish (see Figure 4) which were subsequently mounted on foam core by a local picture framer. The decision to print the averages as A3 was so that the dimensions fell within the range of the individual Artist portraits (~A5-A2).

Figure 4. **Average Portraits.**
Mayor Gordon Bradbery (K, left);
Professor Gordon Wallace (W, right).

Exhibition of the Portraits

Portraits of Gordon[2] was exhibited in the RPAA Gallery 10, which is a rectangular exhibition space. The portraits of Mayor Bradbery were displayed to the left, starting near the Gallery entry, while the portraits of Professor Wallace were on the right, commenced on the opposite wall and finished near the Gallery entry. Each portrait was exhibited so as to have the depicted selion at a height of approximately 1.5m, and spaced equidistantly. Both sets of Sitter portraits commenced with a framed 6" x 8" photographic print of the Sitter's reference photograph exhibited on a plinth. The individual portraits were then arranged in an aesthetic

Accepted Manuscript

sequence by an RPAA member highly experienced in curating exhibitions, and not a participant in the study. In both cases, however, the last image in each Sitter sequence was the Sitter's portrait average. Each portrait was given a unique letter displayed as a large label beneath the work, and following the curated sequence: A-K depicted the Mayor Bradbery; M-W Professor Wallace, with the two average portraits labelled K and W respectively (see Figures 2-4).

As in 2017, participation in the portrait assessments was voluntary, and part of a range of interactive *National Science Week* events. The exhibition officially opened on Saturday, with a preview the night before, and closed the following Sunday (i.e. 10 days). The majority of the 153 volunteer assessors attended during the preview and on the opening day (n=101), and the two Sitters carried out their own assessments (not included in the volunteer assessment data) during their private viewing. Images of the two Sitters attending their (separate) private viewings can be found at: www.illawarramercury.com.au/story/5586518/likeness-of-two-wollongong-gordons-are-cloned-for-scientific-art/.

Volunteer Likeness Assessments

After a short introduction to the purpose of the project, each volunteer recorded their age (in decades, 10s-90s, with 10s representing 18-19 years), sex (F/M), and level of experience as a visual artist (none/some/very). This resulted in the 2018 assessor cohort being very similar to the 2017 study. There were twice as many women (F = 104, M = 49), with the mean age for both sexes being 53 years, and skewed to the older decades (age range: 50s-80s n = 117). Artistic experience was fairly evenly spread, though as in 2017, more women reported being experienced visual artists (F = 74, M = 27) and of these, considerably more women reported being very experienced (F = 36, M = 5).

The design of the Likert scales for this iteration of the study was modified to improve the volunteer assessors' experience and to more accurately record their likeness assessments. Our 2017 study used a Likert scale represented by 5 tick marks on a continuous line, and many assessors carefully recorded their assessments at discrete locations in between the tick marks, which – of necessity – were all recorded as half-way scores. We retained the verbal cues (Very Low, Very High) but used a 7-point scale represented by open circles, and the assessments were subsequently recorded as 1-7 (very low – very high). The A4 printed assessor data and assessment recording sheets were double sited – one side for Mayor

Accepted Manuscript

Bradbery and the other for Professor Wallace, with each including a thumbnail of the respective reference photograph to assist the likeness assessments of unfamiliar viewers.

The visitors first recorded their prior familiarity with the Sitter, and those with at least some prior familiarity (and were therefore able to have an opinion regarding this), recorded their likeness assessment of the displayed Sitter reference photograph. To maintain a similarity of exhibition experience, all visitors were instructed to view and assess the portraits depicting Mayor Bradbury first. As in 2017, no interpretation was provided as to what constituted *likeness*, the visitors were encouraged to view all of the portraits of each Sitter prior to undertaking their assessments, and the order in which they undertook their assessments of each Sitter's portraits was not specified. Once the assessments were complete, the visitors posted their forms into a purpose-built box. Each day, after the exhibition had closed, the box was emptied and the completed assessments were allocated a unique numeric code (that included the date) and filed. Assessments that did not include both sets of portraits were not included in the analyses.

Analyses of the Likeness Assessments

The statistical picture of the cohort of volunteer assessors and their likeness judgements of the reference photographs and portraits (including the portrait averages) were compiled using the following functions from *PAST3 v3.21*: descriptive statistics (mean, standard deviation and standard error), bivariate correlation of the Likert scale and ranked data (Spearman's r_s), and Mann–Whitney pairwise post hoc tests with Bonferroni correction. In addition to calculating the overall means of the likeness assessments, the data were further analysed subdivided by the volunteer assessor's reported prior familiarity with the Sitters' facial appearance – not at all (Likert scale 1); somewhat (2-4); well to very well (5-7).

Reference Photograph Distinctiveness

Facial distinctiveness was estimated from the Sitter's reference photographs, allowing that the display of ambient head poses impacts facial dimensions. This involved calculating the scale-independent indices using the measurement tool in *Adobe Photoshop CC 2019,* and comparing the results to published averages that were also derived from photographs of adult males of European population ancestry (George, 2007).

Accepted Manuscript

Analyses of Shape Accuracy

In a geometric morphometric analysis there is a statistical requirement that the number of landmark coordinates is less than half the number of individuals in the sample in order to undertake multivariate regression analyses (Cardini & Elton, 2007; Webster & Sheets, 2010). As in 2017, the sample sizes for this study are small, and the necessity of applying different landmarks for the two sets of Sitter images meant each analysis could only contain 12 samples (1 Sitter photograph, 10 Artist portraits, 1 portrait average). However, as in 2017, the analyses were undertaken to attain the Procrustes chord Distance (PD) between the reference photograph and each portrait image, and to illustrate the output of the first two Principle Components (PC1 & PC2), which capture most of the shape variance within a sample. The analyses were undertaken twice for each Sitter: firstly with the full complement of head and face landmark coordinates used to create the facial averages (n=77, though differing in location for each set of Sitter images), and secondly with a subset of landmarks excluding the less accurately homologous landmarks that were used to capture scalp hair (see Supplementary Information, Figure SI1-SI2, Table SI1-SI2). The PD calculations and PC1/PC2 thin plate spline wireframes were undertaken in *morphologika v2.5* (O'Higgins & Jones, 1998, 2006), supplemented by a PCA of relative warps using the palaeontological statistics program, *PAST v3.21* (Hammer, Harper, & Ryan, 2001), which enables the identification of PC significance (Bootstrap 100, broken stick) independently of a multivariate regression.

Head Cant

Head canting (tilting the head towards one shoulder) is part of a Sitter's characteristic head pose, but is largely removed from a geometric morphometric analysis through rotation during Procrustes registration. Therefore, the degree of head cant expressed in the images was estimated as a horizontal deviation from the transverse plane between the outer eye corners (exocanthion-exocanthion) and as a vertical deviation from the sagittal midline between the brows and chin (glabella-menton) using the angle measurement tool in *Adobe Photoshop CC 2019.*

Accepted Manuscript

Results

The analyses were undertaken after the exhibition had closed and the likeness data collated. Only 3 individuals failed to assess all of the exhibited portraits, so each portrait and portrait average received ≥ 150 assessments for likeness. Although, as reported in the Methods, the cohort of volunteers is skewed, there was no statistically significant relationship between the likeness assessments of the images and the assessors' sex, age and artistic experience (Spearman's r_s with Bonferroni correction). For greater cohesion, the results are presented separately by Sitter.

Sitter 1: Mayor Bradbery

Likeness and Familiarity

Of the 153 volunteer assessors who indicated their prior familiarity with Mayor Bradbery's facial appearance, 42 had none (Likert scale 1), 41 had some (2-4) and 70 reported a high level of prior familiarity (5-7). All images depicting Mayor Bradbery received a range of assessments from low (1-2) to high (6-7). The overall mean likeness assessments, together with these assessments by level of prior familiarity, are illustrated in Figure 5 and listed, together with the variance and standard error (SE), in the Supplementary Information (Table SI3).

With regards to the reference photograph (as opposed to the portraits and portrait average) the assessors with the highest levels of prior familiarity tended to assess this image a better likeness, and with greater agreement (low variance), than those with only some familiarity, and the difference is significant ($p = 0.007$, Mann-Whitney pairwise post-hoc test with Bonferroni correction).

There is no significant correlation (Spearman's r_s) between the assessors' reported familiarity and their likeness assessments of the portraits and portrait average (A-K), and the most familiar assessors only rated portrait B a better likeness than those with less or no prior familiarity. Differences by familiarity level is only significant for portrait J, and only between those with some and no prior familiarity with Mayor Bradbery's facial appearance ($p = 0.02$, Mann-Whitney pairwise post-hoc test with Bonferroni correction). These results were

Accepted Manuscript

repeated when the Likert scale familiarity levels were re-expressed as some (2-5) and very (6-7).

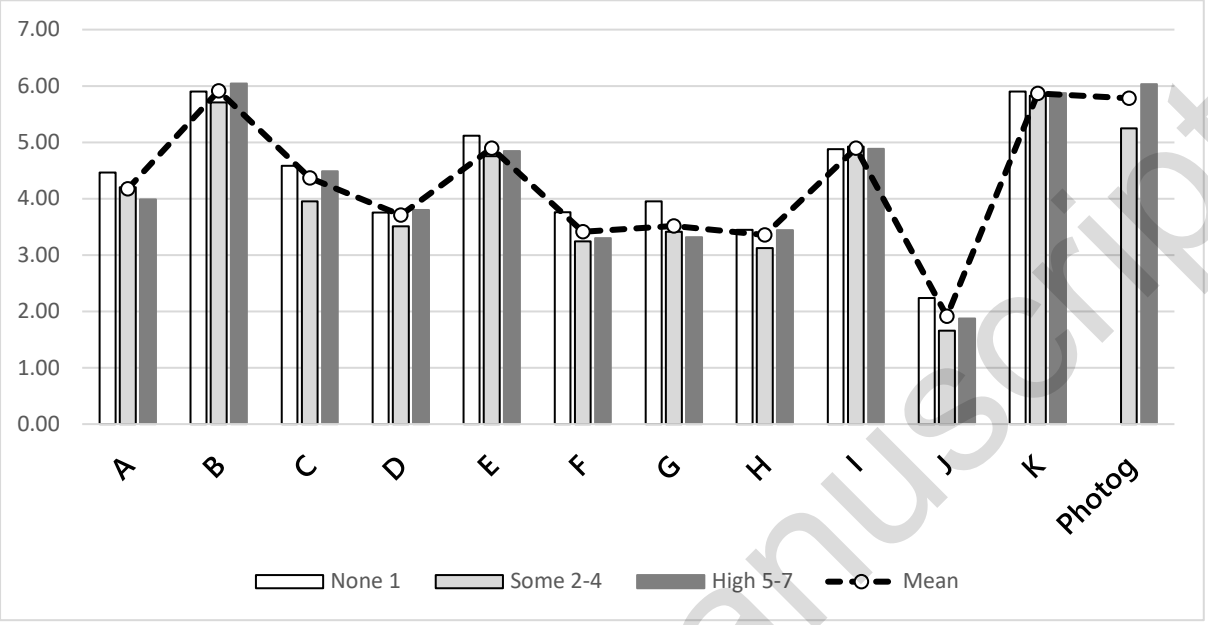


Figure 5. **Comparison of Mayor Bradbery's Portrait Likeness Assessments by Level of Familiarity** The dashed line are the mean likeness judgements (n=153), while the bars show the level of prior familiarity the assessors had with the Sitter's facial appearance: no familiarity (white), some familiarity (mid-grey), moderate to high familiarity (dark grey). Portrait K is the portrait average.

The portrait average (K) is highly ranked for likeness (> 5) across all levels of familiarity, and portrait E by those lacking prior familiarity. Portrait B has the highest overall likeness assessment, and the highest mean likeness rating (6.04) of all of the images depicting Mayor Bradbery, including the reference photograph. This high assessment is from those most familiar with Mayor Bradbery's facial appearance in life, and their assessments have a strong level of agreement (low variance and SE).

Mayor Bradbery assessed portrait B and the portrait average (K) as equally highest in self-likeness, and his assessments of all of the portraits depicting him significantly correlate with the visitor assessments across all levels of familiarity (familiar: $r_s = 0.75$, $p = 0.008$, some familiarity: $r_s = 0.74$, $p = 0.009$, no familiarity: $r_s = 0.79$, $p = 0.004$).

Reference Photograph Distinctiveness

The reference photograph depicts Mayor Bradbery's face with a slightly wider than average upper face width in relation to length, and a wider than average jaw in relation to upper face

Accepted Manuscript

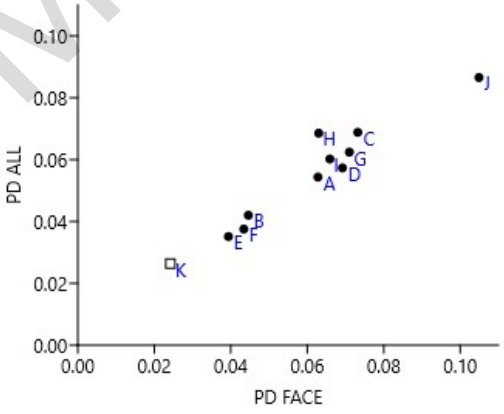
width. The nose is wide at the base in relation to inter-ocular distance, but not in relation to mouth width, while the mouth is wide relative to eye spacing. For all of the other facial indices the reference photograph falls within the average range.

Portrait Shape Accuracy and Variance

The Procrustes chord Distance (PD) of the portraits and portrait average from Mayor Bradbery's reference photograph resulted in the portrait average (K) being closest to the shape variance displayed in the reference photograph, followed by portraits E, F and B (Table 1). The most shape accurate portrait, portrait E, is considered a very good likeness by unfamiliar viewers, and portrait B, a very good likeness by all levels of familiarity, but especially by the most familiar viewers. Portrait F, which is more shape accurate than portrait B, is not highly ranked for likeness irrespective of prior familiarity.

Table 1. **Procrustes chord Distances (PD) resulting from Mayor Bradbery's images**. The most shape accurate cluster of portraits (K, followed by E, F and B) are shaded and in bold. The graph illustrates the agreement between the head and face PD scores. K is the portrait average.

	PD All (77 landmarks)	**PD Face** (68 landmarks)
A	0.0544	0.0628
B	**0.0420**	**0.0446**
C	0.0688	0.0732
D	0.0573	0.0692
E	**0.0351**	**0.0394**
F	**0.0375**	**0.0434**
G	0.0624	0.0710
H	0.0686	0.0630
I	0.0602	0.0659
J	0.0866	0.1049
K	**0.0264**	**0.0242**



The relationship between the results for the full set of landmarks and those focusing on the face is strong (Spearman's $r_s = 0.97$, $p < 0.001$), and therefore only the geometric morphometric analyses of the face (68 landmarks) is presented here.

Both PC1 and PC2 are significant (Bootstrap 100, broken stick), and account for 60% of the overall variance. When these results are plotted (Figure 6), it can be seen that the portrait

Accepted Manuscript

average is located close to the centre of the axes, while the reference photograph is in the positive values for PC1 and negative values for PC2.

PC1 (39.8% variance) is predominantly differentiating between wide-narrow faces, with the narrower faces (positive values) displaying a longer nose and larger eyes that are closer set at the outer eye corners, and a wider left jaw (Figure 6a). Relative to the reference photograph, these characteristics have been slightly exaggerated by the otherwise shape accurate portrait F, and slightly reduced by portraits E and B.

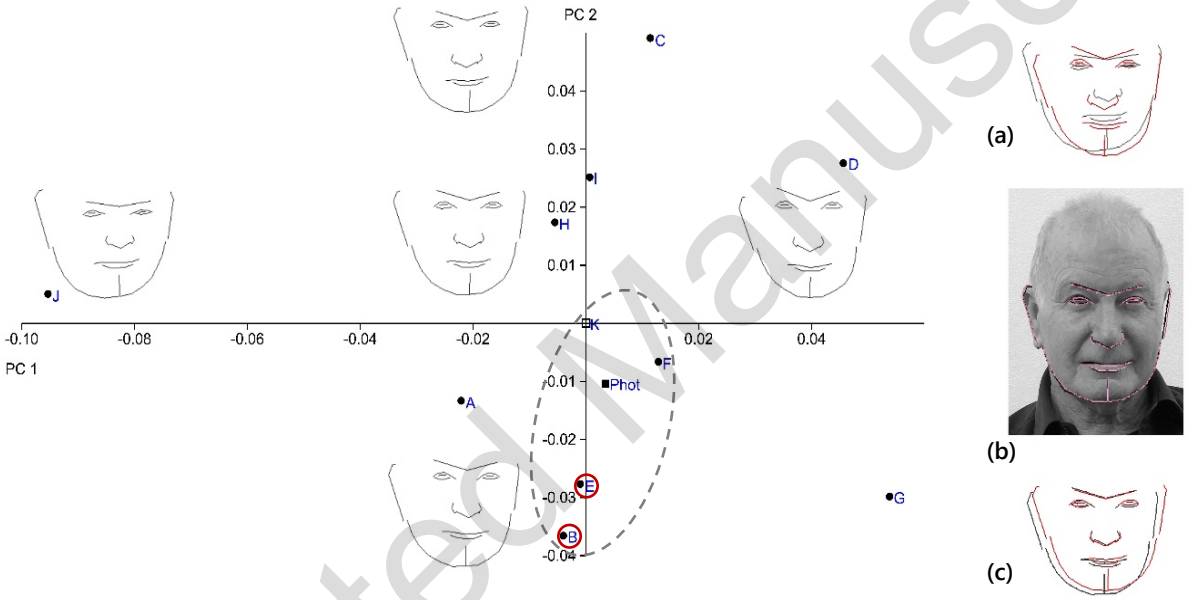


Figure 6. **PC1 and PC2: Morphological Variance of Mayor Bradbery's Images**. The wireframes at the terminal points of the axes are from the maximum values of the shape variance (PC1: -0.10, 0.06; PC2: -0.04, 0.05), and the mean wireframe is located at the centre of the plot (as is the portrait average). The reference photograph is indicated by a filled square, the average portrait by an open square. Inserts to the right are wireframe superimpositions arising from (a) PC1 and (c) PC2, with the positive values shown in red. The centre image (b) shows the Procrustes registered wireframes from the portrait average and Mayor Bradbery's reference photograph. The portraits with high mean likeness scores (> 5) are indicated by a red circle, and the most shape accurate by PD are enclosed in a dashed ellipsis.

PC2 (19.9% variance) is capturing head turn, with the negative values tending towards a more exaggerated head turn, a longer jaw, large eyes that are closer set medially, with the left eye slightly smaller and lower than the right, and a wider mouth (Figure 6c). These characteristics, which more closely overlap with the reference photograph's facial distinctiveness, are reduced by portrait F, exaggerated by portrait E and further exaggerated by portrait B.

Accepted Manuscript

Superimposition of the portrait average Procrustes registered wireframe over the wireframe associated with the reference photograph (Figure 6b) indicates there was a general tendency to depict Mayor Bradbery with a narrower lower vermillion height and a more gracile right jaw.

Likeness, Familiarity, Artistry and Shape Accuracy

Comparison of the likeness assessments with the PD scores (Spearman's r_s) results in a weakly significant overall correlation ($r_s = -0.60$, $p = 0.05$), and is only significant for viewers with some and no prior familiarity (some: $r_s = -0.65$, $p = 0.03$; none: $r_s = -0.63$, $p = 0.04$). The magnitude of the significance, however, is due to an outlier in shape accuracy (portrait J) and all levels of familiarity fail to attain significance when this is removed from the analysis (mean: $r_s = -0.46$, $p = 0.18$; some: $r_s = -0.52$, $p = 0.13$; none: $r_s = -0.52$, $p = 0.12$).

When the Artists' age, portraiture experience, time spent on the portrait's production and their prior familiarity with Mayor Bradbery's facial appearance are compared to the Procrustes chord Distances and Likeness data, none achieve a statistically significant correlation.

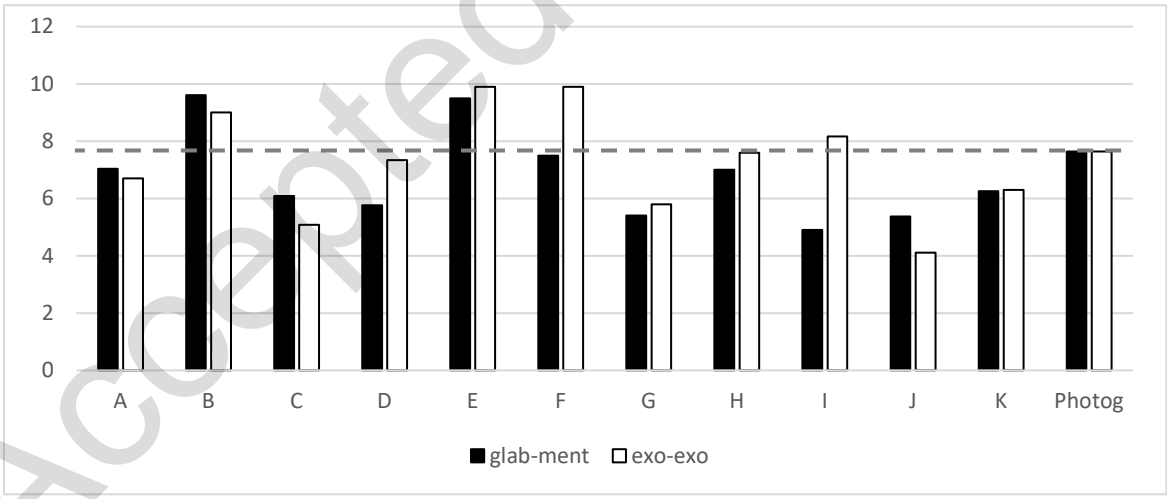


Figure 7. **Depiction of Mayor Bradbery's Head Cant**
The dark bars are the degree of vertical deviation from the midline (glab-ment), the light bars are the horizontal deviation of the eyes (exo-exo). The grey dashed line represents the head cant angles as depicted in the reference photograph (Photog), and the portrait average is portrait K.

Accepted Manuscript

Head Cant Depiction

In the reference photograph Mayor Bradbery's depicted head cant is approximately 7.5° for both the vertical (glabella-menton) and horizontal (exocanthion-exocanthion) planes, indicating the exocanthi are symmetrically located. As can be seen in Figure 7, the portrait average (K) has a reduced, but balanced representation of head cant, while portraits B and E have exaggerated the degree of depicted head cant, and fairly equally. The remaining portraits have tended to reduce the depicted head cant, and more so with the head (glabella-menton) than the outer eye corners (exocanthion-exocanthion). This includes the highly shape accurate portrait F, where only the angle of the eyes has been exaggerated.

Sitter 2: Professor Wallace

Likeness and Familiarity

Of the 127 viewers who recorded their prior familiarity with Professor Wallace's facial appearance, the majority (n=91) were unfamiliar with his facial appearance prior to attending the exhibition, 15 had some familiarity, and 21 reported high to very high familiarity. The reference photograph (on which the portraits were based) was considered by all assessors to be a good to very good likeness (4-7), while all other images attracted a range of assessments from low (1-2) to high (6-7). The mean likeness assessments are illustrated in Figure 8 and listed in the Supplementary Information (Table SI4).

Due to an oversight in the design of the exhibition space, only 32 familiarity-based likeness assessments were undertaken of Professor Wallace's reference photograph (12 some, 20 high familiarity). However, there is a significant difference between these assessors ($p = 0.02$, Mann-Whitney pairwise post-hoc test with Bonferroni correction), with the more familiar having a higher mean assessment of the reference photograph likeness, and a greater level of agreement (low variance and SE).

Across all three levels of prior familiarity the highest ranked for likeness is the portrait average (W), which attains the highest mean ranking (6.43) from those most familiar with Professor Wallace's facial appearance, and the level of agreement is strong (lowest variance of all assessments).

Accepted Manuscript

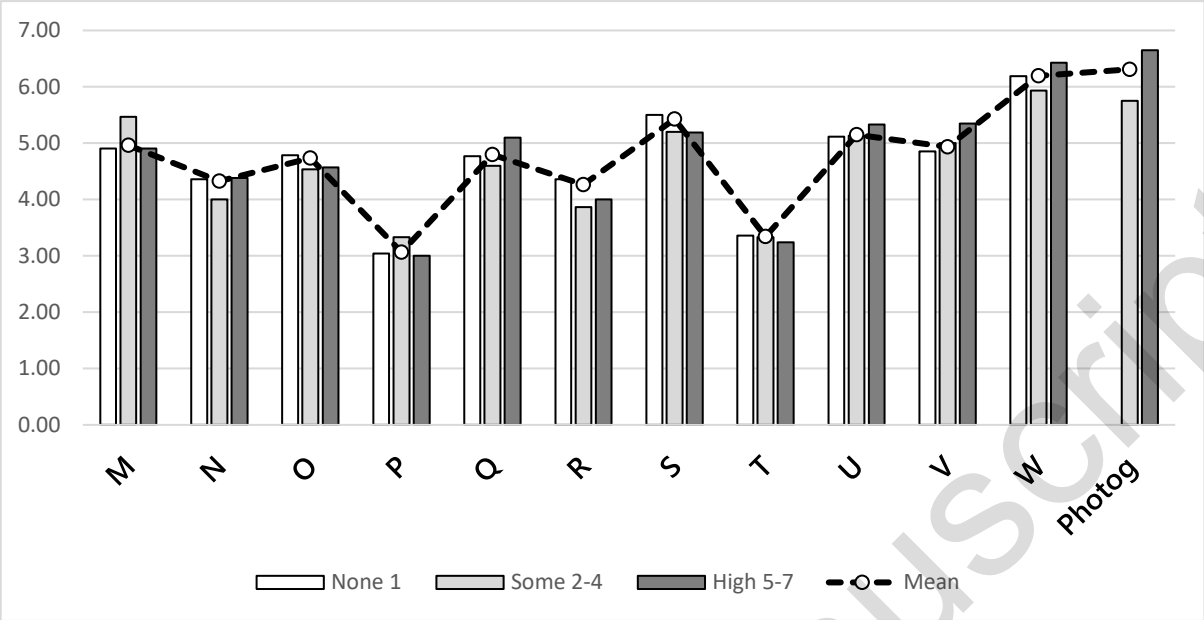


Figure 8. **Comparison of Professor Wallace's Portrait Likeness Assessments by Level of Familiarity**
The dashed line are the mean likeness judgements (n=153), while the bars show the level of prior familiarity the assessors had with the Sitter's facial appearance (n=127): no familiarity (white), some familiarity (mid-grey), moderate to high familiarity (dark grey). Portrait W is the portrait average.

Portraits M, Q, S, U and V all attained a high (> 5) mean likeness assessment. The assessments of portrait V and U increase with familiarity, and this is weakly significant for portrait V (Spearman's $r_s = 0.19$, $p = 0.03$). Portrait S attains a high overall likeness across all groups (5.43), and which is highest for the unfamiliar viewers (5.50). Portrait M is rated highly by those with some familiarity (5.47), and portrait Q is highly ranked by the most familiar (5.10). While the more familiar assessors tended to rate the portrait average and portraits Q, U and V more highly, there is no significant difference between any of the assessments by level of familiarity (Mann-Whitney pairwise post-hoc test with Bonferroni correction).

The self-assessments provided by Professor Wallace do not attain a significant correlation with any of the likeness data, principally because the assessments are very high (≥ 6) for all of the portraits depicting him. The highest ranked works for Professor Wallace's self-likeness are the portrait average (W), and portraits R and O. Portrait O is the most abstract of the portraits exhibited, and is the portrait Professor Wallace indicated he most preferred. Portrait V, which has a weak significant correlation between likeness with familiarity (Spearman's r_s

Accepted Manuscript

= 0.19, $p = 0.03$), was the most preferred by Professor Wallace's family, even though they considered the portrait average (W) to be a better likeness.

Reference Photograph Distinctiveness

Professor Wallace's reference photograph depicts a face that has a wider than average inter-ocular distance, and a wide mouth relative to eye spacing. For all other facial indices that could be estimated (i.e. not obscured by glasses frame or facial hair) the photographed features fall within the average range.

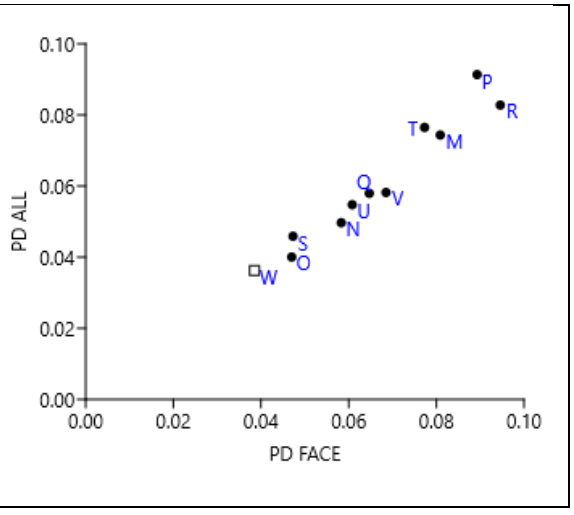
Portrait Shape Accuracy and Variance

The Procrustes chord Distance (PD) analyses of Professor Wallace's images (see Table 2) indicates the portrait average (W) is the most accurate shape, followed by portraits O and S, and there is strong agreement between the two sets of landmarks (Spearman's $r_s = 0.98$, $p < 0.001$). Therefore, only the geometric morphometric analyses of the face are presented here.

PC1 and PC2 account for 47.3% of the overall variance, and both have borderline significance (Bootstrap 100, broken stick). The reference photograph is located within the negative values of PC1 and positive values of PC2 (Figure 9), and the portrait average (W) is at the centre of the axes. Of the portraits with high (> 5) likeness assessments (M, Q, S, U, V), only portrait V shares the same quadrant as the reference photograph.

Table 2. **Procrustes chord Distances (PD) resulting from Professor Wallace's images**. The most shape accurate cluster of portraits (W, O and S) are shaded and in bold. The graph illustrates the agreement between the head and face PD scores. W is the portrait average.

	PD All (77 landmarks)	**PD Face** (67 landmarks)
M	0.0744	0.0809
N	0.0497	0.0583
O	**0.0401**	**0.0470**
P	0.0914	0.0893
Q	0.0580	0.0647
R	0.0828	0.0946
S	**0.0459**	**0.0473**
T	0.0765	0.0773
U	0.0548	0.0608
V	0.0582	0.0685
W	**0.0363**	**0.0385**



Accepted Manuscript

The PC1 thin plate spline wireframes (Figure 9a) indicate that PC1 is capturing variance in jaw shape in relation to head turn, the size of the internal features relative to face size, and overlaps with the facial distinctiveness able to be estimated from the reference photograph. The positive PC1 values depict a more frontally orientated and slightly wider face, larger eyes that are close-set medially, a larger left eye that is located higher on the face, a shorter nose height and a wider mouth on the right. These characteristics are exaggerated by portraits M, Q, S and U, relative to the reference photograph. Both portrait V and the reference photograph are located within the negative values of PC1, which is capturing a wide inter-ocular distance, a left eye that is smaller and located lower than the right, and a more turned head.

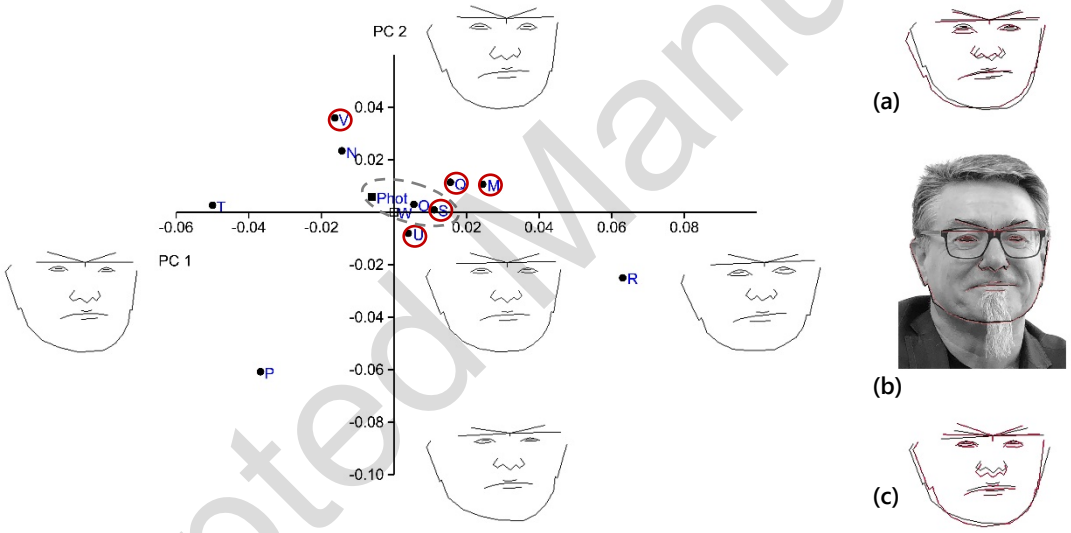


Figure 9. **PC1 and PC2: Morphological Variance of Professor Wallace's Images**. The wireframes at the terminal points of the axes are from the maximum values of the shape variance (PC1: -0.05, 0.07; PC2: -0.06, 0.04), and the mean wireframe is located at the centre of the plot (as is the portrait average). The reference photograph is indicated by a filled square, the average portrait by an open square. Inserts to the right are wireframe superimpositions arising from (a) PC1 and (c) PC2, with the positive values shown in red. The centre image (b) shows the Procrustes registered wireframes from the portrait average and Professor Wallace's reference photograph. The portraits with high mean likeness scores (> 5) are indicated by a red circle, and the most shape accurate by PD are enclosed in a dashed ellipsis.

PC2 is capturing variance in facial width, eye height, and nose width and length (Figure 9c). The positive values, where the reference photograph is located, enlarge the eyes, position the left eye lower on the face, narrow facial and mouth widths, and lengthen the nose. These

Accepted Manuscript

characteristics are most exaggerated by portrait V, slightly exaggerated by portraits M and Q, and slightly reduced by portrait S. Only portrait U is located within the negative PC2 values.

As indicated by the superimposition of the Procrustes registered portrait average (W) over the Professor Wallace's reference photograph (Figure 9b), there is a general tendency for the portraits to depict an exaggerated width of the mouth on the right, and present a more gracile right jaw.

Likeness, Familiarity, Artistry and Shape Accuracy

The mean likeness assessments of Professor Wallace's portraits and portrait average are significantly correlated with face shape accuracy (r_s = -0.62, p = 0.04), and there are no outliers in the PD scores. This significance is repeated by, and only holds for, those assessors with no prior familiarity (r_s = -0.62, p = 0.04).

When the Artists' age, portraiture experience, time spent on the portrait's production and their prior familiarity with Professor Wallace's facial appearance are compared to the Procrustes chord Distances (PD) and Likeness assessments, the Artist's experience as a portrait artist is significant for PD. The significance is strong, and holds for both the accuracy in the depiction of the face (r_s = -0.83, p = 0.003) and the head (r_s = -0.86, p = 0.001).

Head Cant Depiction

Professor Wallace is depicted in the reference photograph with a 14.6° head tilt to the right shoulder when estimated from the vertical, and 10.9° when estimated from the horizontal plane. This indicates an asymmetrical placement of the exocanthi, with the right outer eye corner located slightly higher than the left. No portrait replicates this relationship as depicted in the reference photograph (Figure 10), though portraits N and V retain the proportional differences in angulation while reducing the degree of depicted head cant.

Accepted Manuscript

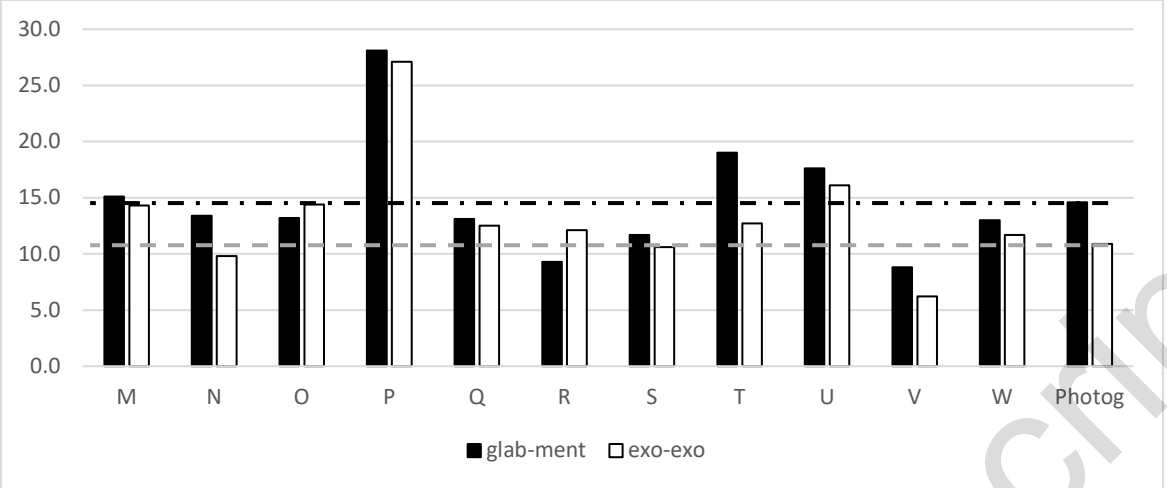


Figure 10. **Depictions of Professor Wallace's Head Cant**
The dark bars are the degree of vertical deviation from the midline (glab-ment), the light bars are the horizontal deviation of the eyes (exo-exo). The black dashed line is the midline cant and the grey dashed line the horizontal cant as depicted in the reference photograph (Photog). The portrait average is W.

Discussion

The aims of this study were to see if ambient portrait averages were both shape accurate and good likenesses, and whether our 2017 findings (Hayes et al., 2018) would differ and/or be replicated with a more facially familiar Sitter (Mayor Bradbery) and a Sitter with relatively low levels of prior familiarity (Professor Wallace).

With regards to the accuracy of ambient portrait averages, both of the statistical averages produced for this study are the most accurate representations of each Sitter's facial shapes and feature configurations, as measured using geometric morphometrics. These results repeat our findings from 2017, which was a study involving a different Sitter, different portrait production conditions, and a different composition of Artist participants. This current study also affirms what we were unable to test in 2017: portrait averages can be very good likenesses.

Although this is the first study (as far as we know) to include likeness assessments of portrait averages produced and exhibited under relatively ambient conditions, our findings agree with analogous research involving the average of laboratory produced facial composites depicting familiar and unfamiliar target faces (Bruce et al., 2002; Hasel & Wells, 2007), and unfamiliar

Accepted Manuscript

face matching involving the average of photographs simulating the low resolution output of CCTV stills (Ritchie, White, et al., 2018). Where our findings differ is that while ambient photograph averages have been found to attract low likeness/recognisability ratings from familiar viewers (Ritchie, Kramer, et al., 2018), our portrait averages attained equally high and higher likeness assessments than most of the individual portraits that comprised them, and this was the case across all levels of prior familiarity.

The tendency for the portrait averages to attract high likeness assessments may be due to the portrait averages being aesthetically pleasing. During the exhibition there were a few inquiries from the visitors (and two of the Artists) as to who created portraits K and W, and one visitor was overheard explaining the (manual) artistic techniques that were applied by the 'artist' to create the visual effects present in portrait average K. There is also the possibility that, as with our previous study (Hayes et al., 2018), the process of undertaking the assessments of the portraits likely provided otherwise unfamiliar visitors with a mental image of each Sitter's face, and this priming effect has been found to increase recognisability (and therefore likeness) when unfamiliar viewers are exposed to a range of ambient photographs depicting a target face (e.g. Andrews, Jenkins, Cursiter, & Burton, 2015; Ritchie & Burton, 2017). As described in the Methods, both portrait averages were exhibited last in each of the curated portrait sequences, and therefore, for those visitors who chose to assess each of the portraits following the exhibited sequence (and some were observed to do this), their assessment of the portrait average would likely have been influenced by their increased levels of familiarity. However, while both ambient portrait averages tended to be considered a very good portrait likeness by all viewers, the portrait average was not always considered the best likeness.

This iteration of our research validates our previous finding that shape accuracy enhances a portrait's likeness for all viewers, and particularly for unfamiliar viewers. While identified as a trend in 2017, in the current study shape accuracy attained statistical significance with the large cohort of viewers who were unfamiliar with Professor Wallace's facial appearance, and adds greater weight to our conclusion that unfamiliar viewers of portraits engage in face matching. Face matching involves a piecemeal comparison of individual features, rather than the gestalt that occurs with familiar face recognition (Young & Burton, 2017). It is likely that our inclusion of the Sitters' reference photographs – both as an exhibit and as an assessment

Accepted Manuscript

sheet thumbnail – better facilitated face matching than our 2017 study's exhibition of the Sitter's face as a life-size 3D monochromatic print.

Not all shape accuracy, however, was perceived equally. While we can add further support to the findings of Ostrofsky et al. (2014) that shape accurate portraits can be perceived as a good likeness, we found that the portraits that did not adhere to the proportions of a Sitters' facial distinctiveness, but were in all other respects highly shape accurate, were not considered good likenesses by either familiar or unfamiliar viewers. Both of the Sitters' reference photographs – on which the portraits were based – display distinctive proportional relationships related to eye spacing. Although slight errors in the depiction of interocular distance in proportion to the other facial features would have a minimal effect on overall shape accuracy, both familiar and unfamiliar viewers have been found to be highly sensitive to even slight manipulations of inter-ocular spacing in photographs (Brédart & Devue, 2006; Ge, Luo, Nishimura, & Lee, 2003).

This iteration of our research also affirms that relatively inaccurate portraits have enhanced likeness ratings if they include exaggeration of distinctive features, and particularly for the likeness ratings of familiar viewers. The only portrait in this study to attain a significant positive correlation between likeness and familiarity (portrait V) was not particularly shape accurate. But portrait V, which depicted the less well-known Sitter, Professor Wallace, was the only artwork to exaggerate this Sitter's facial distinctiveness, including the degree of head turn. A related finding, which was suspected in the 2017 study but not proven, was that exaggeration and shape accuracy in portraiture need not be mutually exclusive. Portrait B is one of the three highly shape accurate portraits depicting Mayor Bradbery. Portrait B is also the portrait that contained the greatest exaggeration of this Sitter's distinctive ocular index, and greatest exaggeration of the Sitter's depicted head turn. Mayor Bradbery is a well-known face in the local region, and the volunteer assessors were predominantly familiar with his facial appearance in life. Although portrait B was considered a very good likeness by all viewers, it was, on average, assessed by those most familiar as the best likeness – better than the portrait average, and slightly better than the Sitter's reference photograph. This effect, which has been observed with ambient photographs, links levels of likeness to recognisability, and recognisability to levels of familiarity (Ritchie, Kramer, et al., 2018). Unlike the 2017 study, however, for this study the higher likeness assessments by the most familiar were not statistically significant for any of the portraits or portrait averages. The

Accepted Manuscript

exception to the relative flatness of the likeness assessments was the reference photographs. Both reference photographs tended to be considered very good likeness by the less familiar, and both attained a significantly higher rating by the most familiar. This indicates that (i) the reference photographs were of an appropriate likeness/recognisability on which to base the portraits, and (ii) it was possible, but did not happen in this iteration of the study, for the more familiar viewers to express a statistically significant 'peak shift' (Ramachandran & Hirstein, 1999) in their likeness assessments of the portraits and portrait average.

The overall tendency for the portraits (and therefore portrait averages) to depict both Sitters with a more gracile right jaw could be because all of the Artists are women, and therefore tended to inadvertently feminise the lower face of both Sitters. However, studies suggest artists exaggerate sexually dimorphic features in sitters of a different sex (Ramachandran & Hirstein, 1999). Alternatively, and because this tendency towards gracility only occurs on one side of the face, this could be due to both Sitters exhibiting a left head turn, which has been exaggerated in some, but not all, of the portraits – and a left head turn may produce a more gracile appearance to the right jaw.

Related to head turn, a more detailed analysis of head pose was made possible because the Artists used reference photographs of relatively ambiently posed Sitters to produce the portraits. An earlier study of portrait shape accuracy (Hayes & Milne, 2011) found that a person's head pose tends to be perceived in photographs and depicted in portraits as more upright. Furthermore, this tendency was found to be greater for head pitch (up/down) and cant (head tilting) than head turn. Nearly all of the portraits produced for this study either reduced the head cant and turn displayed in the reference photographs and/or depicted the extent of cant and turn differently across the head and face, impacting on the interaction of all of the landmark coordinates. Professor Wallace's reference photograph depicts an upwards head pitch as well as a right cant and left turn, thereby displaying a more complex head pose than the reference photograph depicting Mayor Bradbery. For this iteration of the study we added the Artists' experience as a portrait artist and their prior familiarity with each Sitter to the Artist data, and although neither was significant for the accuracy and likeness of the portraits depicting Mayor Bradbery, a greater experience with portraiture significantly contributed to more shape accurate depictions of Professor Wallace.

Accepted Manuscript

Regarding the cohort of volunteer assessors, while we again, and as has been reported elsewhere (Kozbelt, Seidel, ElBassiouny, Mark, & Owen, 2010), found no impact of age, sex and experience as a visual artist on the likeness assessments, this result would have been stronger were it not for the marked similarity of the composition of this cohort to those who contributed to our 2017 study.

Finally, there is the shape accuracy and likeness of the abstract ambient portrait, which essentially repeats the findings from the 2017 study. Portrait O, which is the only abstract work, was produced by a different Artist. However, and as in 2017, the only abstract portrait was the most accurate portrait for shape depiction, did not tend to be seen as a very good likeness across all levels of prior familiarity, and was the portrait the Sitter most preferred.

To summarise, this study has both extended and validated our previous findings, including that unfamiliar viewers tend to assess shape accurate portraits as good likenesses, and this includes shape-accurate portrait averages. Familiar viewers, however, tend to assess relatively shape accurate portraits a better likeness if they contain some degree of congruent exaggeration of a Sitter's facial distinctiveness (which covers aspects of head pose as well as proportional relationships between facial features), and while shape-accurate abstract portraits tend to attract relatively low likeness ratings from most viewers, they appear to be the portrait most preferred by the people they depict.

Acknowledgements

We acknowledge the traditional owners of the land on which this research has taken place, the Wodi Wodi of the Dharawal Nation, and pay our respects to Elders past, present and emerging for their knowledge and care of Country. The authors also thank the following for their support and assistance: Lord Mayor Gordon Bradbery AM (Wollongong City Council), Professor Gordon Wallace (ARC Centre of Excellence for Electromaterials Science, IPRI, AIIM, University of Wollongong), Dulcie Dal Molin, Meagan Powley, Katherine Orton (Red Point Artists Association, Port Kembla), Gerrit van den Bergh (Centre for Archaeological Science, SEALS, SMAH, University of Wollongong), Paul Jones (Media and Public Relations, University of Wollongong), all of the visitors who chose to attend the *Portraits of Gordon*^{2} exhibition, and the input of provided by our two anonymous reviewers.

Accepted Manuscript

Declaration of Conflicting Interests: The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding: This project was self-funded by the RPAA membership.

References

Andrews, S., Jenkins, R., Cursiter, H., & Burton, A. M. (2015). Telling faces together: Learning new faces through exposure to multiple instances. *The Quarterly Journal of Experimental Psychology, 68*(10), 2041-2050.

Brédart, S., & Devue, C. (2006). The accuracy of memory for faces of personally known individuals. *Perception, 35*(1), 101-106.

Bruce, V., Ness, H., Hancock, P. J., Newman, C., & Rarity, J. (2002). Four heads are better than one: combining face composites yields improvements in face likeness. *Journal of Applied Psychology, 87*(5), 894-902.

Cardini, A., & Elton, S. (2007). Sample size and sampling error in geometric morphometric studies of size and shape. *Zoomorphology, 126*(2), 121-134.

Faigin, G. (1990). *The Artist's Complete Guide to Facial Expression*. New York: Watson-Guptill Publications.

Ge, L., Luo, J., Nishimura, M., & Lee, K. (2003). The lasting impression of Chairman Mao: hyperfidelity of familiar-face memory. *Perception, 32*, 601-614.

George, R. M. (2007). *Facial Geometry: Graphic Facial Analysis for Forensic Artists*. Springfield Illinois: Charles C Thomas.

Hammer, Ø., Harper, D., & Ryan, P. (2001). PAST-Palaeontological statistics. *www. uv. es/~ pardomv/pe/2001_1/past/pastprog/past. pdf, acessado em, 25*(07), 2009.

Hasel, L. E., & Wells, G. L. (2007). Catching the bad guy: Morphing composite faces helps. *Law and Human Behavior, 31*(2), 193-207.

Hayes, S. (2010). *Seeing and Measuring the 2D Face.* (PhD), PhD Thesis. University of Western Australia,

Hayes, S., & Milne, N. (2011). What's wrong with this picture? An experiment in quantifying accuracy in 2D portrait drawing. *Visual Communication, 10*(2), 149-174.

Hayes, S., Rheinberger, N., Powley, M., Rawnsley, T., Brown, L., Brown, M., . . . Wallace, G. (2018). Variation and Likeness in Ambient Artistic Portraiture. *Perception, 47*(6), 585-607. doi:10.1177/0301006618770347

Hayes, S., & Tullberg, C. (2012). Police witness identification images: a geometric morphometric analysis. *Journal of Forensic Sciences, 57*(6), 1487-1494. doi:10.1111/j.1556-4029.2012.02168.x

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1571), 1671-1683.

Kozbelt, A., Seidel, A., ElBassiouny, A., Mark, Y., & Owen, D. R. (2010). Visual selection contributes to artists' advantages in realistic drawing. *Psychology of Aesthetics, Creativity, and the Arts, 4*(2), 93.

Maughan, W. L. (2004). *The Artist's Complete Guide to Drawing the Head*. Watson-Guptill: New York.

O'Higgins, P., & Jones, N. (1998). Facial growth in Cercocebus torquatus: An application of three dimensional geometric morphometric techniques to the study of morphological variation. *Journal of Anatomy, 193*, 251-272.

Accepted Manuscript

O'Higgins, P., & Jones, N. (2006). Tools for statistical shape analysis. Retrieved from http://hyms.fme.googlepages.com/resources

Ostrofsky, J., Cohen, D. J., & Kozbelt, A. (2014). Objective versus subjective measures of face-drawing accuracy and their relations with perceptual constancies. *Psychology of Aesthetics, Creativity, and the Arts, 8*(4), 486.

Ramachandran, V. S., & Hirstein, W. (1999). The science of art: a neurological theory of aesthetic experience. *Journal of Consciousness Studies, 6*(6-7), 15-51.

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *The Quarterly Journal of Experimental Psychology, 70*(5), 897-905.

Ritchie, K. L., Kramer, R. S., & Burton, A. M. (2018). What makes a face photo a 'good likeness'? *Cognition, 170*, 1-8.

Ritchie, K. L., White, D., Kramer, R. S., Noyes, E., Jenkins, R., & Burton, A. M. (2018). Enhancing CCTV: Averages improve face identification from poor-quality images. *Applied Cognitive Psychology, 32*(6), 671-680.

Rohlf, F. J. (2015). The tps series of software. *Hystrix, the Italian Journal of Mammalogy, 26*(1), 9-12.

Speed, H. (1917). *The Practice and Science of Drawing* (3rd ed.). New York: Dover.

Webster, M., & Sheets, H. D. (2010). A practical introduction to landmark-based geometric morphometrics. *Quantitative Methods in Paleobiology, 16*, 168-188.

White, D., Burton, A. L., & Kemp, R. I. (2016). Not looking yourself: The cost of self-selecting photographs for identity verification. *British Journal of Psychology, 107*(2), 359-373.

White, D., Burton, A. M., Jenkins, R., & Kemp, R. I. (2014). Redesigning photo-ID to improve unfamiliar face matching performance. *Journal of Experimental Psychology: Applied, 20*(2), 166.

Wisely, N., & Fine, G. A. (1997). Making faces: portraiture as a negotiated worker-client relationship. *Work and Occupations, 24*(2), 164-188.

Young, A. W., & Burton, A. M. (2017). Recognizing faces. *Current Directions in Psychological Science, 26*(3), 212-217.

Accepted Manuscript

Supplementary Information

Table SI1: Mayor Bradbery Landmarks

	Right	Medial	Left	Definition
Eyes	1		2	Centre of iris
	3		6	Lateral border of iris
	4		5	Medial border of iris
		10		Between eyebrow heads near horizontal skin crease
	24		25	Most superior point of eyebrow, lateral to lateral iris
	51-55		56-60	Upper eyelid, endocanthion to exocanthion, derived from curves
	61-64		65-68	Lower eyelid, curruncular margin to interior exocanthion, derived from curves
Nose	15		23	Superior nasal wing
	16		22	Lateral nasal wing
	17		21	Lateral base of wing/nostril
	18		20	Lateral superior nostril
		19		Mid inferior septum
Mouth		39-45		Mouth crease, right to left cheilia, derived from curves
		46-50		Lower lip vermilion, right to left, derived from curves
Ears	7		77	Most superior point on pinna
	8		9	Most lateral point on pinna
	13		-	Most inferior point on ear lobe
	-		14	Most lateral point of lower lobe
Jaw		11		Mid-point of upper chin crease
		12		Edge of lower chin beneath mental crease
	27-32		33-38	Jaw, from lobe/cheek to lateral chin, derived from curves
Hair		26		Thickest part of hairline superior to the glabella
		70-76		Scalp hair following densest growth, derived from curves

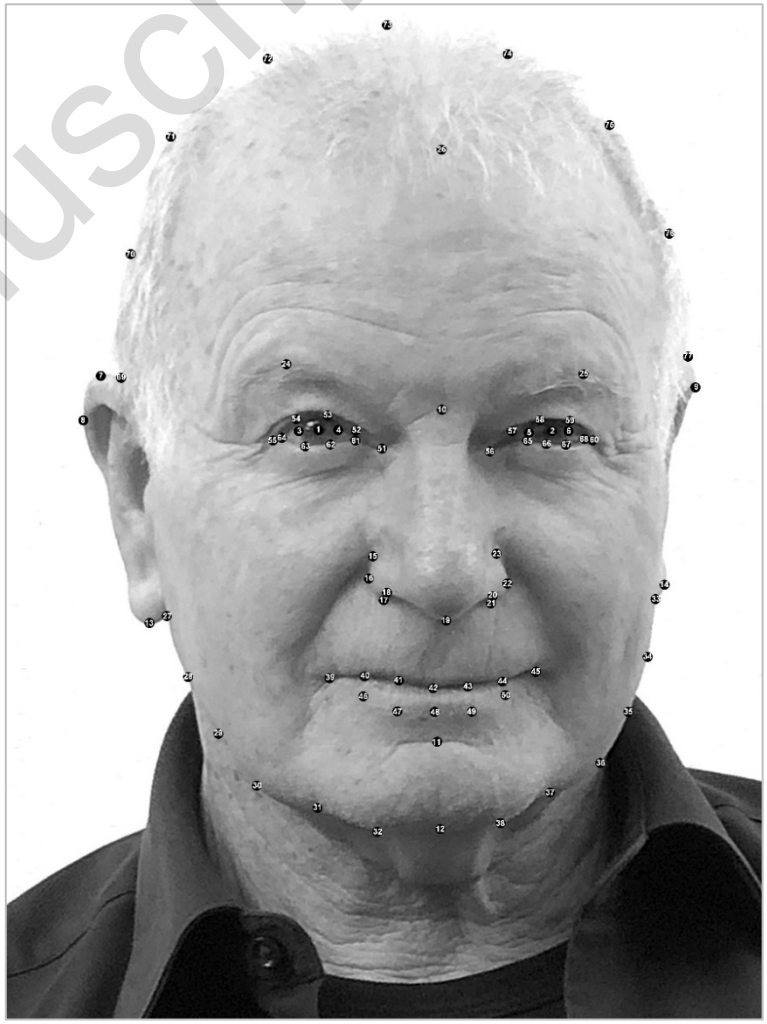


Figure SI1. **Mayor Bradbery Landmarks**

Accepted Manuscript

Table SI2: Professor Wallace Landmarks

	Right	Medial	Left	Definition
Eyes	1		2	Centre of iris
	3		6	Lateral border of iris
	4		5	Medial border of iris
	21		22	Most superior point of eyebrow, lateral to lateral iris
	48-52		53-57	Upper eyelid, endocanthion to exocanthion, derived from curves
	58-61		62-65	Lower eyelid, curruncular margin to interior exocanthion, derived from curves
Glasses frame	26		28	Lateral superior glasses frame
		27, 29		Medial glasses frame
Nose	12		20	Superior nasal wing
	13		19	Lateral nasal wing
	14		18	Lateral base of wing/nostril
	15		17	Lateral superior nostril
		16		Mid inferior septum
Mouth	75	76	77	Central upper lip/philtrum, derived from curves
	40	41-46	47	Mouth crease, right to left cheilia, derived from curves
	8	24	9	Lower lip vermilion, centre, derived from curves
Ears	66		-	Most superior point on pinna
	7		74	Most lateral point on pinna
	10		11	Most inferior point on ear lobe
Jaw	30-34		35-39	Jaw, from lobe/cheek to lateral goatee, derived from curves
Hair		23		Thickest part of hairline superior to the glabella
		67-73		Scalp hair following densest growth, derived from curves
		25		Lowest point of goatee with densest growth

Figure SI2. **Professor Wallace Landmarks**

Accepted Manuscript

Table SI3. Mayor Bradbery's Portrait Likeness Assessments: Overall and by Familiarity
The highest ranked images for mean likeness (> 5) are shaded.

	All (N = 153)	**Familiar** (n = 70)			**Some Familiarity** (n = 41)			**No Familiarity** (n = 42)		
	Mean	**Mean**	**Var.**	**SE**	**Mean**	**Var.**	**SE**	**Mean**	**Var.**	**SE**
Photog	5.78	6.03	1.53	0.15	5.25	2.26	1.50			
A	4.17	3.99	2.22	0.18	4.20	1.86	0.22	4.46	1.30	0.18
B	5.91	**6.04**	0.97	0.12	5.71	1.61	0.20	5.90	1.39	0.18
C	4.37	4.49	1.24	0.13	3.95	1.60	0.20	4.59	2.05	0.22
D	3.71	3.80	1.35	0.14	3.51	1.71	0.20	3.76	1.84	0.21
E	4.90	4.84	1.90	0.16	4.76	1.79	0.21	5.12	1.33	0.18
F	3.41	3.30	1.14	0.13	3.24	1.59	0.20	3.76	1.60	0.20
G	3.52	3.31	1.41	0.14	3.41	1.30	0.18	3.95	2.73	0.25
H	3.36	3.44	2.02	0.17	3.12	2.71	0.26	3.45	1.84	0.21
I	4.90	4.89	1.61	0.15	4.93	1.77	0.21	4.88	1.91	0.21
J	1.92	1.87	1.50	0.15	1.66	0.88	0.15	2.24	1.16	0.17
K	5.87	**5.87**	1.45	0.14	5.83	1.80	0.21	5.90	1.49	0.19

Table SI4. Professor Wallace's Portrait Likeness Assessments: Overall and by Familiarity
The highest ranked images for mean likeness (> 5) are shaded.

	All (N = 127)	**Familiar** (n = 21)			**Some Familiarity** (n = 15)			**No Familiarity** (n = 91)		
	Mean	**Mean**	**Var.**	**SE**	**Mean**	**Var.**	**SE**	**Mean**	**Var.**	**SE**
Photog	6.31	6.65	0.45	0.15	5.75	1.48	0.35			
M	4.96	4.90	1.89	0.30	5.47	0.70	0.22	4.91	1.51	0.11
N	4.33	4.38	1.95	0.30	4.00	0.86	0.24	4.36	1.99	0.13
O	4.73	4.57	2.16	0.32	4.53	0.98	0.26	4.79	1.84	0.13
P	3.07	3.00	2.00	0.31	3.33	2.10	0.37	3.04	1.54	0.11
Q	4.80	5.10	1.39	0.26	4.60	1.83	0.35	4.77	1.75	0.12
R	4.26	4.00	1.20	0.24	3.87	1.41	0.31	4.36	1.44	0.11
S	5.43	5.19	1.46	0.26	5.20	1.74	0.34	5.50	1.56	0.12
T	3.34	3.24	1.89	0.30	3.33	1.81	0.35	3.36	1.85	0.13
U	5.15	5.33	2.03	0.31	5.13	1.27	0.29	5.11	1.59	0.12
V	4.93	5.35	1.08	0.23	5.00	2.29	0.39	4.85	1.38	0.11
W	6.20	6.43	0.66	0.18	5.93	2.07	0.37	6.19	0.96	0.09

Accepted Manuscript

Accepted Manuscript