University of Wollongong

Research Online

Faculty of Science, Medicine and Health -
Papers: part A

Faculty of Science, Medicine and Health

1-1-2016

A geometric morphometric evaluation of the Belanglo 'Angel' facial approximation

Susan Hayes

University of Wollongong, suehayes@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/smhpapers>

Part of the [Medicine and Health Sciences Commons](#), and the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Hayes, Susan, "A geometric morphometric evaluation of the Belanglo 'Angel' facial approximation" (2016). *Faculty of Science, Medicine and Health - Papers: part A*. 4189.
<https://ro.uow.edu.au/smhpapers/4189>

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

A geometric morphometric evaluation of the Belanglo 'Angel' facial approximation

Abstract

In August 2011, a 2D facial approximation was undertaken of remains discovered in Australia's Belanglo State Forrest; in October 2015, the young woman was identified. Referencing three photographs of the young woman as she appeared in life and a database of 64 sex, age, head pose and population matched images, the facial approximation is evaluated for relative shape accuracy through the application of geometric morphometrics. The results are that the facial approximation is significantly similar to the images of the young woman in facial morphology ($p = 0.002$) when most of the variance due to depicted head pose is removed from the analyses. The geometric morphometric analyses, however, also highlight the facial approximation's face and feature discrepancies, some of which would have likely disrupted familiar face recognition. Although predominantly verified methods were applied in 2011, they are limited in their predictive accuracy, not every feature of the face has a verified method to apply, and practitioner errors as well as photographic distortions are apparent. Furthermore, an assumption that the verified methods require inter-feature agreement (in this instance eye spacing and mouth width) was proved to be false. Overall this study shows that it is possible to assess the morphological accuracy of a forensic facial approximation when a number of antemortem images are available, though the influence of photographic distortion within 2D photographs will always preclude a precise metric assessment.

Disciplines

Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Hayes, S. (2016). A geometric morphometric evaluation of the Belanglo 'Angel' facial approximation. *Forensic Science International*, 268 e1-e12.

Accepted Manuscript

Title: A geometric morphometric evaluation of the Belanglo 'Angel' facial approximation

Author: Susan Hayes

PII: S0379-0738(16)30415-7
DOI: <http://dx.doi.org/doi:10.1016/j.forsciint.2016.09.009>
Reference: FSI 8599

To appear in: *FSI*

Received date: 10-12-2015
Revised date: 24-3-2016
Accepted date: 12-9-2016

Please cite this article as: Susan Hayes, A geometric morphometric evaluation of the Belanglo 'Angel' facial approximation, Forensic Science International <http://dx.doi.org/10.1016/j.forsciint.2016.09.009>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A geometric morphometric evaluation of the Belanglo 'Angel' facial approximation

Susan Hayes, PhD¹

¹*Faculty of Science, Medicine and Health, University of Wollongong*

Address for correspondence:

Dr Susan Hayes
Centre for Archaeological Science
School of Earth and Environmental Sciences
Faculty of Science, Medicine and Health
University of Wollongong
New South Wales Australia 2522

E: susan_hayes@uow.edu.au

T: +61408784512

Please note that I am currently undertaking research in Indonesia. Therefore email is the preferred method of contact.

Highlights

- The first evaluation of a forensic facial approximation's accuracy since 2001
- Geometric morphometrics enables comparisons involving head pose variation
- The results are that the facial approximation is accurate ($p=0.002$)
- Although accurate, the inaccuracies would likely have disrupted recognition

Abstract

In August 2011, a 2D facial approximation was undertaken of remains discovered in Australia's Belanglo State Forrest; in October 2015, the young woman was identified. Referencing three photographs of the young woman as she appeared in life and a database of 64 sex, age, head pose and population matched images, the facial approximation is evaluated for relative shape accuracy through the application of geometric morphometrics. The results are that the facial approximation is significantly similar to the images of the young woman in facial morphology ($p = 0.002$) when most of the variance due to depicted head pose is removed from the analyses. The geometric morphometric analyses, however, also highlight the facial approximation's face and feature discrepancies, some of which would have likely disrupted familiar face recognition. Although predominantly verified methods were applied in 2011, they are limited in their predictive accuracy, not every feature of the face has a verified method to apply, and practitioner errors as well as photographic distortions are apparent. Furthermore, an assumption that the verified methods require inter-feature agreement (in this instance eye spacing and mouth width) was proved to be false. Overall this study shows that it is possible to assess the morphological accuracy of a forensic facial approximation when a number of antemortem images are available, though the influence of photographic distortion within 2D photographs will always preclude a precise metric assessment.

In August 2010, the skeletal remains of a young woman (15-25 years, European population affinity) were found in Australia's Belanglo State Forest. A year later, the New South Wales Police requested a facial approximation and the results were released to the national media in December 2011 (Figure 1a). Although the estimation of this young woman's facial appearance assisted in generating a substantial number of new leads, none resulted in a positive identification. In 2013, the methods used to estimate facial appearance formed part of a critical comparative review of the largely invalidated and popular forensic facial reconstruction approach [1]. In October 2015, the young woman, referred to in the Australian media as "Angel", was identified as Ms Karlie Jade Pearce-Stevenson, facilitated by a lead unconnected to those generated by the facial approximation. The facial approximation essentially failed in its primary purpose, which is assist in identification. However, an evaluation of the predictive accuracy of a facial approximation is part of the process when the face is estimated as applied research, and ideally all forensic estimations are evaluated, regardless of the outcome.

Some of the local media coverage included comments regarding the similarity between the facial approximation and the photographs of Ms Pearce-Stevenson which predominated in the media following her identification: e.g. "remarkably accurate" [2]; "amazing likeness" [3]. Such a *post hoc* response to a facial approximation is fairly commonplace [4, 5] and does not reflect the actual recognisability or accuracy of the results. A facial approximation is undertaken to stimulate leads to identification from colleagues, friends and family members [6, 7], but face recognition studies show that individuals who are unfamiliar with the person depicted utilise different neural mechanisms, attend to different aspects of the head and face [8], and essentially engage in face-matching [for reviews, see 9, 10, 11]. Face-matching is frequently confounded by variation in head pose even when the images are captured under

the same photographic conditions, and the most salient feature is the shape and hue of terminal head hair [8, 12-14]. In contrast, familiar face recognition is influenced by facial configuration rather than individual features, though this is a more complex neural process than past studies have indicated [15], is head-pose invariant, though this is shared to some extent with unfamiliar face recognition [16], very rapidly processed [17], and occurs even when highly degraded CCTV images are used [18].

Familiar assessment has been undertaken with a facial reconstruction created using CT scans of a colleague's skull and teeth [19], and there is a report of parents being unable to recognise the facial reconstruction of their child, even after being told this is who the results were intended to predict [20]. It is, however, generally agreed that a facial approximation undertaken with a forensic case cannot be easily, or ethically, tested for familiar face recognition [21]. As a consequence the recommended accuracy test of a facial approximation is visual assessment by unfamiliar participants [e.g. 21, 22, 23], which has been found to require a minimum of 115 assessors [21]. Metric assessments of accuracy occur within laboratory-based studies, but no actual forensic estimation has been formally evaluated for predictive accuracy since at least 2001. However, this is likely due to traditional anthropometric measures requiring standardised photographs of the identified individual to undertake a comparison.

Traditional anthropometrics is reliant on direct measurements taken from a person's head and face (i.e. using callipers) or scaled, orthogonal images (e.g. 3D scans). If anthropometric measures reference photographs, each individual needs to be photographed bearing the same head orientation and facial expression, and the camera angle and lens-subject distance needs to be standardised so as to reduce, but not entirely remove, the effect of perspective distortion

[see, for example, 24]. Images taken outside of the laboratory, including arrest photographs, tend not comply with these requirements [25]. Instead, most photographs typically differ in head pose, facial expression, camera make and camera angle as well as variation in lens-subject distance, and this holds true for the photographs of Ms Pearce-Stevenson published in the Australian media.

Geometric morphometrics is a tool for statistical shape analysis that compares the patterns of shape variance across a suite of homologous landmarks, and therefore accounts for the interaction of all landmarks simultaneously rather than piecemeal. Because of its capacity to process complex patterns of shape variance, previous research has found that geometric morphometrics can enable statistically meaningful comparisons between (i) highly disparate image types depicting the face of the same individual [25, 26], (ii) widely divergent palaeoart facial reconstructions of the same skull [27], and (iii) between images of different individuals displaying a range of habitual head pose variations [28]. It is because of this ability to identify shape variance due to head pose that geometric morphometrics has been applied for this evaluation of the facial approximation of Ms Pearce-Stevenson, though photographic distortion remains a confounding variable.

Materials

(i) Photographs of Ms Pearce-Stevenson and the facial approximation

Four photographs of Ms Pearce-Stevenson have appeared in the Australian media since her identification was announced [e.g. 29, 30, 31] and the three that most clearly display her face and features were used to estimate the morphological accuracy of the facial approximation. Permission was granted to reproduce Figure 1c (Ava Benny-Morrison, [30]), and this image,

together with the two remaining photographs are represented in Figure 1d-f by the landmark coordinate wireframes used in the geometric morphometric analyses.

The three photographs are unscaled and informally posed. Two depict Ms Pearce-Stevenson seated and gazing upwards (Figure 1d-e). While neither seems to display an upwards head orientation, both appear to have been taken from an angle above head height, which, on the basis of past forensic photographic comparisons using geometric morphometrics [e.g. 32, 33-35], results in a facial morphology that shares some of the shape changes resulting from a downwardly pitched head pose (the upper face is expanded and lengthened while the lower face is contracted and foreshortened [26]). One photograph is of comparatively low resolution and depicts Ms Pearce-Stevenson fairly frontally orientated with a relatively neutral/serious facial expression (Figure 1d), while in the higher resolution photograph her head is turned to the right shoulder (both eyes are visible but the right ear is obscured) and Ms Pearce-Stevenson is displaying a closed-mouth smile (Figure 1e). The third photograph also depicts Ms Pearce-Stevenson displaying a right head turn and closed-mouth smile, but in this image she appears to be standing and the photograph has been taken from below (Figure 1c, f), which may result in an upwards head pose facial morphology (the upper face is contracted and foreshortened while the lower face is expanded and lengthened [28]). The two seated photographs were taken in 2008 when Ms Pearce-Stevenson was ~ 20 years of age, while the third photograph is from when Ms Pearce-Stevenson was still attending secondary school (personal correspondence, Ava Benny-Morrison [30]).

All three photographs of Ms Pearce-Stevenson are used for this geometric morphometric analyses because no one image can be assumed to accurately represent Ms Pearce-Stevenson's facial morphology. All three are confounded by variation in head pose and

camera orientation, two images are of low resolution, two display facial expression, one is of Ms Pearce-Stevenson some years prior to her death, and all three can be assumed to contain varying elements of photographic distortion.

The facial approximation of Ms Pearce-Stevenson is frontally orientated, has a neutral facial expression, and also contains elements of photographic distortion. As described previously [1], the facial approximation was undertaken in reference to digital, scaled photographs of the skull and, due to the constraints of the assigned location, from a distance of 2 m from the nasion. Each of the resulting images were scaled to life size and checked against calliper measurements for overt levels of photographic distortion (heights and widths of the neurocranium, facial skeleton, orbits, nasal aperture, mandible and teeth). It took the better part of a day to achieve satisfactory results, and these were described as 'orthogonal' [1]. Photographic distortion, however, has a complex impact on the morphology of both the skull and the face. Early anthropometric research with film-based photographs found that a lens-subject distance of ~ 10-12 m was required before distortions were within acceptable measurement errors (1%) [36, 37]. A later study by Eliášová and Krsek using digital cameras found that perspective distortion in frontal views of the skull is ~ 3% when the lens-subject distance is 2.5 m; while at distances less than 2.5 m the lateral neurocranium and jaw noticeably recedes while the anterior frontal bone expands [38]. Similarly, and more recently, Stephan finds a distortion effect of ~ 5% at distances less than 2.5 m, and confirmed earlier studies that 12 m is required for photographic distortion to be negligible (< 1%) [39]. Therefore, although the images of the skull that were referenced for the facial approximation agreed with the calliper measures, it is highly likely that the lateral cranium, frontal bone and mandible contain elements of photographic distortion.

(ii) *Image composition of the database*

The database used for the geometric morphometric analyses consists of 64 photographs, which is comparable to, or larger than, previous geometric morphometric analyses of the face [40-42] and studies of facial averages [43, 44]. The photographs are of young women of recorded or apparent European population ancestry with a recorded or apparent age of 18-25 years, and only closed-mouth smiles were selected from the images displaying this facial expression. The database images were sourced as follows:

- (i) *PAL Face Database* (n=16): North American women recorded as “white”, aged between 19-22 years, displaying an informal frontal head orientation and with a very slight smile, neutral or serious facial expression. All images are sourced from the PAL Face Database [45], accessed on 2 August 2014 (<http://agingmind.utdallas.edu/facedb>);
- (ii) *Utrecht Face Database* (n=6): Frontally posed images of young women displaying a slight smile, serious or neutral facial expression (actual age and population ancestry unknown); one image was flipped along the horizontal plane to display a right head turn. All images were collected at the European Conference on Visual Perception in Utrecht (2008), and made available on the Stirling University website. All were accessed on 6 December 2015 (http://pics.stir.ac.uk/2D_face_sets.htm);
- (iii) *Pasco Sheriff Arrest Database* (n=24): Arrest images of young women recorded as “white”, born between 1994-1997 (i.e. 19-22 years of age on the day the arrest image was taken) displaying a slight smile, neutral or serious facial expression, with the face relatively unobscured. Four images were accessed on 17 January 2010 and the remainder during 4-13 February 2016, with the website limiting access to 100 arrest photographs taken over the preceding 60 days (<http://pascosheriff.com/arrests/>);
- (iv) *Internet images* (n=18): Young women of European facial appearance (actual age and population ancestry unknown). These images are predominantly of professional and

amateur models, and were primarily selected for their similarity in head pose to the photographs of Ms Pearce-Stevenson that occur less frequently in the Database sourced images (downwards and/or right head turn). One internet image was sourced on 18 October 2011 [46], 10 images were sourced in November, 2016 and 7 in February, 2016. Four of the internet images were flipped along the horizontal plane to display a right head turn (for the internet sources, refer [47]).

To avoid confusion, hereinafter all references to the PAL, Utrecht and/or Pasco Sheriff Arrest Databases will be capitalised, while the database of images composed of the three Database and internet photographs (n=64) will be lower case. In order to assess the degree to which this database of photographs deviates from an average face derived from a more homogenous source with regards to head pitch and turn, face shape and facial expression, a facial average was sourced as follows:

- (v) *UK facial average* (n=1): the average wireframe resulting from a study which applied 224 “feature points” (facial landmarks) to produce a facial average referencing photographs of 60 “caucasian” (British) young women (20-30 years) bearing a neutral facial expression [43].

Methods

The 64 database photographs, the three images of Ms Pearce-Stevenson, the UK facial average and the facial approximation were standardised to the interpupillary plane (a horizontal line connecting the centre of the pupils) in Adobe CS4 Photoshop. All images were then visually assessed for head pose, face shape and facial expression. Head pose was estimated by the position of the ears, which can appear raised in a downwards head pitch and obscured in a head turn [26, 28], face shape according to whether it appeared short/wide, tall/thin or an indeterminate combination, and facial expression by the shape of the lips and

cheeks, which appear raised during smiling [48, 49]. This resulted in the following distributions for the database images (n=64):

15 upwards, 21 downwards, 28 indeterminate head pitch;
 9 left turned, 11 right turned, 33 indeterminate head turn;
 16 tall/thin, 27 short/wide, 21 indeterminate facial shape;
 47 serious or neutral, 17 slightly smiling facial expression.

A significant inverse correlation was found between head pitch and turn (Pearson's $r = -0.26$, $p = 0.02$) when all of the frequencies were analysed in the statistics software *PAST* (v.3.11, 2016 [50]), indicating that the right turned faces tend towards a downward pose and vice versa. Visual assessment of faces by one individual is, however, known to be inaccurate with regards to estimating head pose, and requires a minimum of 10 assessors to statistically correlate with the accuracy of a geometric morphometric analysis [26, 28].

A total of 40 landmark coordinates capturing face and facial feature shape (Table 1 and Figure 2) were recorded using the 2D digitising software *tpsDig2* (v.2.17, 2013 [51]). Of these, 30 landmarks were selected for their relative robusticity and homology, with an emphasis on those landmarks used in the facial approximation to estimate the facial features and/or resulting from the estimation. An additional 10 semi-landmarks (2 x 5) were used to capture the shape of the outer face and jaw, and converted to landmarks within *tpsUtil* (v.1.56, 2013 [51]). The resulting *tpsDig2* file (image sets of the facial landmark x,y coordinates) was checked for outliers within the geometric morphometric software package *MorphoJ* (v.1.06d, 2013, [52]), and those due to landmark errors (placement and/or numerical sequence) were corrected. To test whether the shape variation within the images is viable for undertaking multivariate statistical analyses, the corrected *tpsDig2* file was entered into *tpsSmall* (v.1.33, 2016 [51]), with the result that the shape variation is appropriately narrow (regression slopes 0.99, correlation 1.0), and a digital facial average of the database

photographs (n=64, Figure 3a) was produced using *tpsSuper64* (v.2.02, 2015 [51]) (see [53] for further information regarding the Rohlf *tps* series of geometric morphometric software).

The *tpsDig2* file was uploaded into the statistical shape analysis tool, *morphologika2* (v.2.5, 2007 [54]). Because all of the photographs were taken from a range of unknown and unreported lens-subject distances, all 69 images (including the UK facial average and the facial approximation) were first Procrustes registered prior to the analyses within *morphologika2*, *MorphoJ*, *tpsSuper64* and *PAST*. Procrustes registration removes the effect of size, some of which is due to variation in image resolution, and involves each set of landmark coordinates being scaled and rotated for comparable fit.

To identify the main patterns of shape variance following Procrustes registration, Principle Components (PC) analyses were undertaken within *morphologika2*, which has the facility to display the shape variance captured by the suite of landmarks as dynamic and static wireframes. When extracted from the maximum positive and negative values for individuals on each PC axis (i.e. within the variance of the group), the wireframes illustrate the dominant morphological aspects being captured by each PC (see Figures 3 and 5). Multivariate regression analyses were undertaken within *morphologika2* to identify patterns of specific variance across a group of selected PCs, and to identify which of the PCs are significant partial regression coefficients. In addition, Thin Plate Spline (TPS) deformation grids illustrate the deformation pattern between a selected reference point and selected target point following a multivariate regression (see Figure 4). To both verify the results and supplement the statistical output arising from the analyses undertaken in *morphologika2*, additional analyses were undertaken with the statistical software package *PAST* (Principle Components Analyses (relative warps) and similarity/distance matrices).

The following independent variables were used to identify the overall composition and distribution of the database images (n=64) regarding depicted head pose orientation, face shape and facial expression, the source of the 64 database photographs, and the extent to which this database deviates from the UK facial average.

1. *Head pitch*: Visual assessment: upwards = 1, indeterminate = 2, downwards = 2;
2. *Head turn*: Visual assessment: left = 1, indeterminate = 2, right = 3;
3. *Face shape*: Visual assessment: tall-narrow = 1, indeterminate = 2, short-wide = 3;
4. *Facial expression*: Visual assessment: neutral/serious = 1, slight smile = 2;
5. *Source*: PAL and Utrecht Face Databases = 1; internet images = 2; Pasco Sheriff Arrest Database = 3;
6. *UK facial average*: UK facial average = 1; database images = 2.

The analyses of head pose orientation, face shape and facial expression (1-4 above) were repeated including the UK facial average (n=65) to record the location of this image along each of the PC axes (i.e. the PC score).

The morphological evaluation of the facial approximation involved multivariate regressions using the following independent variables, and apart from a general assessment of head pose, did not include the UK facial average:

7. The three photographs of Ms Pearce-Stevenson = 1; the database images = 2 (n=67);
8. The facial approximation = 1, the database images = 2 (n=65);
9. The three photographs of Ms Pearce-Stevenson and the facial approximation = 1; the database images = 2 (n=68).

Results

The results of the database composition analyses concerned with variance in head pitch, turn, face shape, facial expression, database source, and how the database compares to the UK facial average (multivariate regressions 1-6), are summarised in Table 2. The morphological analyses of the photographs of Ms Pearce-Stevenson and the facial approximation (multivariate regressions 7-9) are summarised in Table 3. As can be seen, all of the statistically significant multivariate regressions result in partial regression coefficients with a standard error (S) of $< 25\%$ (0.025).

Database composition analyses

The first 20 PCs capture 98% of the overall variance, and these were used for the database composition analyses. As can be seen in Figure 3, including the facial average rotates the orientation of the database data in Cartesian space, but not the overall configuration of the points. The multivariate regressions involving the database, and those involving the database plus the UK facial average, were statistically significant, and both produced very similar results. Therefore only the results of the analyses including the UK facial average (i.e. $n=65$) are shown in Table 2.

Each of the independent variables, with the exception of the regression by the UK facial average, resulted in significant correlation coefficients: head pitch PC 1; head turn PC 2, which is also present to a slight degree in PC 1 (and agrees with the what was indicated by the visual assessment data); facial shape PC 3 and, to a lesser extent, PC 7 (PC 7 discriminates between triangular versus rounded jaws, which is also captured by PC 6) and PC 1 (there are slightly more short/wide faces with a right head turn); facial expression PC 4, and, to a lesser extent, PC 5 (which is concerned with mouth width, and does not include the smiling related expansion of nose width [48]). However, and as mentioned previously, these

independent variables are derived from inherently inaccurate visual assessments, and on the basis of past research findings [26, 28] can be assumed to under-represent the statistical significance of the related shape variance.

The Procrustes registered landmark coordinates for the 64 database images were entered into a PCA (relative warps) in *PAST* (bootstrap 1000), to see which PCs are statistically significant without a multivariate regression. The results are that PC 1-3 are clearly significant, and PC 4 is within the 95% confidence interval (broken stick). Furthermore, and as with the analyses undertaken in *morphologika2*, the *PAST* deformation grids show the variance is due to head pitch and a slight degree of turn (PC 1), head turn (PC 2), tall-thin/short-wide facial shapes (PC 3), and smiling/not smiling facial expressions (PC 4). In both *morphologika2* and *PAST*, PC 1-4 accounts for 83.5% of the overall variance, and the percentage variance for each PC is close to identical in agreement (< 0.03% difference).

The *morphologika2* multivariate regression by image source involved only the database images and results in PC 1 and PC 2 being statistically significant. PC 1 and PC 2 account for 66% of the cumulative variance, which, as discussed above, is predominantly due to the influence of head pitch and head turn, and combinations of these poses. The scatter plot of PC 1 (x axis) and PC 2 (y axis) is shown in Figure 3. As can be seen, the PAL and Utrecht Database photographs (rectangles) display the most upwards and neutrally pitched faces, and the most left and neutrally turned faces. The most right turned and most downwardly pitched photographs are the internet images (circles) that were selected for these characteristics, but moderate displays of head turn and pitch are also present within the photographs sourced from the Pasco Sheriff Arrest Database (triangles).

When included in the analysis, the UK facial average is located within the PAL and Utrecht Database cluster within PC 1/PC 2 (-0.39, -0.036), which is indicated by an arrow in Figure 3. This location, and the comparison of this database's facial average with the UK facial average (Figure 3, insert b) shows that the database is, on average, more turned and pitched. Within the remaining PCs the facial average is located close to the database mean of 0,0 (see Table 2), and the multivariate analysis involving regressing the UK facial average against the database images is not significant (< 1% of the total sample variance) and has no significant partial correlation coefficients.

What the results of the UK facial average multivariate regressions indicate is that the inclusion of the internet sourced photographs have resulted in a photographic database that is in greater agreement with the head pose variation displayed in the three photographs of Ms Pearce-Stevenson, but does not deviate from the UK facial average in facial shape, facial expression or other facial characteristics captured by the geometric morphometric analyses (the average database face produced in *tpsSuper64* is shown in Figure 3a). The database composition therefore enables a more accurate geometric morphometric evaluation of the photographs of Ms Pearce-Stevenson and the facial approximation. This is because while PC 1 and PC 2 capture the statistically significant variance due to head pitch and turn, these aspects are still present to some extent within the variance captured by the remaining PCs, and earlier trials indicated that head pose continued to confound analyses involving photographs predominantly sourced from the more neutrally posed PAL and Utrecht Facial Databases, and also when only one image of Ms Pearce-Stevenson was analysed.

Morphological analysis of the facial approximation

When the three photographs of Ms Pearce-Stevenson, the facial approximation and the database images are analysed within *morphologika2*, the variance due to head pitch and turn are significant for both PC 1 (37% overall variance) and PC2 (31% overall variance). PC 1 is capturing a downwards left turned to neutral and upwards right turned head pose (head pitch: 14% variance, $S = 0.006$; head turn: 13% variance, $S = 0.006$). PC 2 is the inverse of this pattern, downwards right turned to upwards left turned (head pitch: 8% variance, $S = 0.006$; head turn: 6% variance, $S = 0.006$). There is a difference of 26% between the cumulative variance of PC1-2 (68%) and the variance arising from the multivariate regressions based on visual assessment (41%). That the visual assessment results are 0.60 of the geometric morphometric variance, however, is in agreement with the inaccuracy of one person undertaking a visual assessment of head pose from photographs. When 10 individual assessors are compared there is a tendency to under-estimate head pose, and inter-assessor agreement has a low Pearson's correlation (pitch $r = 0.641$, turn $r = 0.747$) [26, 28].

When the UK facial average is entered into the analysis, it again groups with the PAL and Utrecht database images. The frontal/downwards image of Ms Pearce-Stevenson (Figure 1d) clusters within the downwardly orientated database photographs and is close to the UK facial average in head turn. The right turned/downwards image of Ms Pearce-Stevenson (Figure 1e) is right turned, and shares a similar downwardly pitched head pose to the more frontally orientated photograph. The younger image of Ms Pearce-Stevenson (Figure 1c, e) has less of a right head turn and clusters within the range of the slightly upwardly pitched images in the database. The facial approximation is located close to the UK average in both head orientations, but is more upwardly pitched. Therefore, relative to the UK facial average, the head pose variation within the three photographs of Ms Pearce-Stevenson is neutrally turned,

moderately right turned, strongly right turned, moderately downwards, and slightly upwards. All subsequent analyses exclude the UK average.

Separate multivariate regressions were undertaken with the photographs of Ms Pearce-Stevenson and the facial approximation with and without PC 1 (head turn) and PC 2 (head pitch). The partial regression coefficients are the same (PC1-20, PC3-20), but a multivariate regression involving PC 3-20 enables morphological comparisons with the impact of head pose variation significantly reduced (PC1-2 cumulative variance = 68%), though not entirely removed.

The regression involving the three photographs of Ms Pearce-Stevenson and the database photographs (Table 3, multivariate regression 7, Figure 4a) indicates Ms Pearce-Stevenson had, compared to the database, a short/wide face, wide spaced eyes, slightly wide nose and short nose-mouth distance, narrow and thin lips, and a longer than average distance between the mouth and chin (stomion-menton). When the facial approximation is similarly compared to the database (Table 3, multivariate regression 8, Figure 4b) this results in the facial approximation also differing from the database in having a short/wide face, a slightly wide nose and slightly long mouth-chin distance (stomion-menton). The facial approximation, however, has a longer nose, much shorter than average nose-mouth distance, a mouth that is only slightly narrower, and lips that are close to the database average in height.

When the two multivariate regression results are directly compared (Figure 4c), the facial approximation deviations from the database mean can be clearly seen to differ from how the photographs of Ms Pearce-Stevenson deviate. The facial approximation has a shorter and wider face shape, larger and narrower spaced eyes, a longer nose, shorter nose-mouth

distance, fuller and wider mouth, and a shorter distance between the mouth and base of the chin. (Note that the wireframe results represented in Figure 4 are mathematical visualisations of morphological deviations from the database average, and do not represent the landmark coordinates of the photographs of Ms Pearce-Stevenson, the facial approximation, or the database average.)

The final multivariate regression (Table 3, multivariate regression 9) results in the facial approximation clustering with the photographs of Ms Pearce-Stevenson, and this clustering is significant when PC 1-20 ($p < 0.001$) and PC 3-20 ($p = 0.002$) are included in the regression. The primary significant partial regression coefficients are PC3 (face shape), and PC8 (eye spacing, nose width, mouth width, lip heights, chin height). PC 10 and PC 17 are also significant partial regression coefficients, but capture $< 1\%$ of the overall variance. PC10 separates the facial approximation out from the cluster by iris size and nose-mouth distance, while PC17 contains subtle aspects of facial asymmetry that could be accounted for by landmark error and/or variance due to photographic distortion.

When the analysis is undertaken in *PAST* the PCA (relative warp) scatter plot for PC 3 and PC 8 is identical to the PC 3/PC 8 scatter plot results in *morphologika2*, and the difference in variance for PC 3-20 is $< 0.01\%$. The *PAST* scatter plot for PC 3/PC 8 is illustrated in Figure 5 because of the added functionality of a 95 % confidence ellipsis and a Minimal Spanning Tree within an eigenvalue scaled scatter plot. As can be seen, the facial approximation clusters with the younger right turned, more upwardly posed image of Ms Pearce-Stevenson (Figure 1c, f), which in turn is co-joined with the photograph where she is depicted in a frontal/downwards orientation (Figure Figure 1d). All three images are fairly distantly connected to the more markedly right turned photograph of Ms Pearce-Stevenson (Figure 1e),

but in this analysis the frontal/downwards photograph falls outside of the 95% confidence ellipsis.

A similarity-distance matrix undertaken in *PAST* and involving the PC 3-20 scores for the 68 images (an extract of the full matrix is shown in Table 4) results in the facial approximation being most similar to the frontal/downwards photograph of Ms Pearce-Stevenson, and least similar to the image in which she is depicted right turned/downwards. The matrix also indicates that the frontal/downwards image is most similar to the younger right turned/upwards image of Ms Pearce-Stevenson, and that the right turned/downwards image is considerably distant from both. In both the PCA (relative warp) minimal spanning tree and in the similarity-distance matrix, the intervening image is the same database photograph. That is, between the facial approximation and the younger image of Ms Pearce-Stevenson (PCA) and between the facial approximation and the frontal/downwards image (similarity-distance matrix) is Shape 56, which is the PAL Database photograph shown in Figure 2.

The similarity-distance matrix indicates the right turned/downwards image (Figure 1e) is morphologically very different to the other two photographs of Ms Pearce-Stevenson (Figure 1c, f and 1d). This is because aspects of head pose variance is still present even when PC1 and PC2 are removed from the analysis. However, excluding the right turned/downwards photograph from a multivariate regression comparing Ms Pearce-Stevenson to the 64 database images produces very similar results. As can be seen in Table 3 (multivariate regression 7a), there is a slight increase in overall variance, face shape (PC 3) reaches statistical significance, and there is an additional significant correlation coefficient (PC 11) capturing < 1% of the variance. Furthermore, and as can be seen in Figure 4c, the wireframe resulting from this analysis is only slightly dissimilar to that resulting from the three

photographs of Ms Pearce-Stevenson, and highlights the same discrepancies in the facial approximation's prediction of Ms Pearce-Stevenson's facial morphology.

Discussion

Both in the forensic report that accompanied the facial approximation and in the publication that describes the methods [1] facial indices were calculated from the results. These indicated that:

relative to the average, this young woman had a wide face, wide jaw, and wide-spaced eyes, a long nose, short upper and lower lip (nose to mouth, and mouth to start of chin), and the suggestion of a high chin. (e35)

The geometric morphometric evaluation of the facial approximation, relative to the average of the 64 database images (Table 3, multivariate regression 8 and Figure 4b), is in general agreement with these anthropometric indices. When the facial approximation is included in an analysis involving the three photographs of Ms Pearce-Stevenson, the results are that the facial approximation is statistically similar to, and clusters closely with, Ms Pearce-Stevenson in facial morphology (Table 3, multivariate regression 9 and Figure 5). These findings are supported by the similarity-distance matrix (Table 4), though this matrix also shows that the right turned/downwards image of Ms Pearce-Stevenson is noticeably dissimilar to the facial approximation, and dissimilar to the frontal/downwardly orientated image and the younger photograph of Ms Pearce-Stevenson. This indicates that head pose variance in photographs requires more than one antemortem image, and ideally a number of images expressing a range of head poses, to be viable for a geometric morphometric evaluation of a facial approximation.

Although the facial approximation clusters with the images of Ms Pearce-Stevenson and can be considered to have succeeded in generally predicting her facial morphology, the results also indicate that, relative to the average of the 64 database images (Table 3, multivariate regression 7 and Figure 4a), Ms Pearce-Stevenson's facial configuration and feature dimensions were somewhat different to what was predicted. Figure 4c is a direct comparison of how the facial approximation and the photographs of Ms Pearce-Stevenson deviate from the database average, and as can be seen, the facial approximation has poorly predicted facial height and width, eye height and spacing, nasal length, nose-mouth distance, lip fullness and mouth width.

Some of these errors are due to inherent limitations in the methods. All of the verified methods that were applied in 2011 are derived from statistical averages, and individuals are not statistical averages. So a general pattern of inaccuracy in the results is to be expected. However, not all of the methods that were applied in 2011 were verified methods. As discussed in the paper that describes the facial approximation [1], some were untested recommendations that have since been proven invalid, and not all aspects of the face have a verified method to apply. Furthermore, the geometric morphometric analyses also show that an assumption that there needed to be inter-feature agreement between the different methods is a false assumption.

As can be seen in Figure 4c, the geometric morphometric results show that the facial approximation predicted a too short/wide facial shape. This is likely due to errors in how the facial soft tissue depths (fSTD) were applied, rather than the measures themselves. As pointed out by Stephan and Simpson [55] and illustrated by George [56], very few fSTDs are located perpendicular to the skull. A recent evaluation of fSTDs taken from CT data [57]

clarifies the misinterpretation of 'perpendicular' across the facial reconstruction literature, and which is due to a misreading of the tangent method described in Aulsebrook et al. [58]. This facial approximation should have followed the tangent method described by Aulsebrook et al. [58], which incorporates the specific curvature of the bone. The foreshortening and widening of the face may also be due to photographic distortion, which causes the lower face to recede and the upper face to expand when the skull is photographed at distances of less than 2.5 m [38, 39]. It also the case that errors in assigning the landmarks onto a 2D image [59] may have contributed to the resulting facial shape, and in particular at the mid-ramus and gonion. Had the fSTDs been applied directly onto the skull and subsequently photographed, as is recommended [60], then this would have reduced, but not entirely removed, the problem of photographic distortion and facilitated a more accurate landmark placement. However, this approach also results in contaminating the evidence, which is not desirable in forensic facial approximations. CT scans or 3D scans are a far more accurate basis on which to estimate facial appearance, and have an additional advantage of more easily facilitating a methodological revision of the original results [61].

The over-estimation of eye height was due to the application of an invalid surgical recommendation [62] that exaggerated the height of the iris. With regards to eye spacing, the facial approximation also slightly underestimated eyeball width [63], and did not fully comply with research findings that the eyeball is displaced from the orbital centre by, on average, 1.4 mm superiorly and 2.3 mm laterally [64, 65]. This was in part due to the cited studies involving a small number of cadaveric subjects of advanced age, but a large validation study involving 375 living adults [63] has since verified the findings. The decision to only displace the eyeballs by 1 mm superiorly and laterally, however, was in the main part because a greater displacement did not agree with the verified methods used to estimate

mouth width. Estimation of mouth width included the infraorbital foramen [66, 67], inter-canine width [68] and a combination of inter-canine width and the location of the pupils [69], and resulted in a narrow mouth width that did not align with the medial border of the iris when the eyeball was displaced by 2.3 mm. So, on the false assumption that there should be inter-feature agreement, lateral eyeball displacement was reduced by 1.4 mm and the mouth slightly widened. As a consequence the facial approximation has underestimated Ms Pearce-Stevenson's eye spacing and overestimated her mouth width.

The estimation of the nasal dimensions followed the algorithms of Rynn et. al [70], who validated their findings in reference to five subjects (2 males, 3 females). These recommendations have since been evaluated by Mala [71] regarding lateral projection referring to 86 lateral cephalograms. Mala found that for women of European population affinity (n=34, 19-39 years), the algorithms have an error value of ~ 2.2 mm (6.7% of the dimensions), though tend to perform better regarding estimating nasal height. It is not clear, therefore, why the algorithms resulted in a nasal length that is slightly longer than Ms Pearce-Stevenson's, and longer than the average dimensions of the image database, but this may be an inherent limitation of this method.

Lip heights were predicted particularly poorly, as was the vertical orientation of the oral fissure. When taken together with the slightly too long nose, these estimations resulted in the facial approximation predicting an unusually short nose-mouth distance. The geometric morphometric analysis suggests that, relative to the database images, Ms Pearce-Stevenson had narrow lip heights, and a nose-mouth distance that is only slightly shorter than the database average. The predicted lip heights were known to be highly approximate, for in the absence of the central incisors a published average [72] was applied. This also meant that the

recommendation for the vertical orientation of the oral fissure [73, 74] could not be used, and instead followed the description of oral fissure location in a well respected anatomical text [75]. This respected, but unverified, location, resulted in the mouth sitting slightly too high within the face. Mala and Veleminska [76] recently tested facial reconstruction recommendations for the estimation of lip heights and the vertical orientation of the oral fissure (51 males, 35 females; European population affinity). Their study found the oral fissure was best predicted at $\frac{3}{4}$ the height of the upper central incisors [73], which had an error rate of 1.3 mm. However, given the central incisors are frequently lost post-mortem, it would seem that a new method is required to estimate lip heights and vertical placement.

Not every feature of the facial approximation has been covered by the geometric morphometric evaluation, and this includes the estimation of the eyebrow location and the hair style. It is fairly clear from all of the photographs of Ms Pearce-Stevenson that the eyebrows are located too low on the face, and therefore the unverified facial reconstruction recommendation that links eyebrow location to the superior orbital rim [77] likely requires revision. Hair style was accomplished in collaboration with a professional hairdresser, and this was reasonably accurate with regards to the shape of the hair (straight) and style of cut (long side fringe). Where the hair style predicts less well is in mass (dense) and texture (coarse). These inaccuracies are probably due to the hairdresser predicting mass and texture from photographs of hair that had been exposed to the elements for a considerable period of time, coupled with a manipulation of the hair style to largely obscure the highly speculative estimation of ear shape and height. It would seem advisable, therefore, that a hairdresser be able to view any remains of head hair *in situ*, which in this instance was not possible due to this aspect being completed more than 3000 km from the remains.

Although the geometric morphometric analyses show the facial approximation is statistically similar to Ms Pearce-Stevenson in facial morphology, it is highly likely that the prediction of a too short/wide face, the reduction in eye spacing, the exaggerated lip fullness and in particular the under-estimation of the nose-mouth distance would not have facilitated recognition. As stated previously, familiar face recognition is known to involve configural processing, which includes how the features are spatially organised within the face [78]. A recent study of the impact of aspect ratio distortion [15] notes that previous studies have found familiar face recognition is not disrupted when the face is stretched vertically [79], and subsequent studies have found that linear manipulations of the aspect ratio of the face slows identification, but does not prevent it. This explains why familiar face recognition is tolerant to the spatial distortions that occur with head pose variation. These studies, however, involve manipulating photographs of a familiar person, and so are linear distortions of a face and facial features that accurately represent the individual prior to being manipulated (though there will be aspects of photographic distortion). Furthermore, these studies of configural processing distort the whole face, and not individual features.

A different series of face perception studies have manipulated individual features to see the extent to which facial memory of familiar faces is sensitive to internal feature displacement. The findings are that face memory is particularly sensitive to the vertical orientation of the mouth and to a reduction, but not expansion, of eye spacing [80-83]. The facial approximation of Ms Pearce-Stevenson contains no feature that is accurately predicted, and while there is some uniformity in the over-estimation of size (the face is too wide, the eyes are too large, the nose is too long, the lips are too full), this pattern is reversed when it comes

to eye-spacing, nose-mouth distance and facial height. As mentioned, familiar face recognition is generally agreed to be neither ethically nor easily accomplished with a forensic facial approximation. However, it seems reasonable to assume that while the facial approximation may have been 'recognised' through an automated face recognition system (that had more than one image of Ms Pearce-Stevenson in its reference image database), familiar face recognition would have been disrupted by the non-uniform displacement of the facial feature configuration, and in particular the under-estimation of eye spacing and nose-mouth distance.

Conclusion

Geometric morphometrics has been applied to evaluate the accuracy of a facial approximation of a young woman discovered in Australia's Belanglo State Forrest in 2010. Three antemortem images of the young woman were used in this evaluation to allow for the influence of head pose variation, and a database of 64 images displaying a comparable range of head pose variations formed the basis of the analyses. Although geometric morphometrics cannot entirely remove the effect of head pose on the face, and all of the images in the analyses are confounded to an unknowable degree by photographic distortion, the results are indicative as to how well the facial approximation predicted facial appearance. Overall the facial approximation is significantly similar to the young woman it predicted, but within this similarity there are noticeable differences that would have likely disrupted familiar face recognition. Some of these are due to the limitations of the methods applied, some are due to a misapplication of these methods, most of the unverified recommendations are simply wrong, and there was a false assumption that the verified methods that were applied required inter-feature agreement. Although metric assessments of actual forensic facial

approximations are rare, this study shows that geometric morphometrics can be used to evaluate a facial approximation for shape accuracy when the analyses include more than one image of the identified individual and the comparative database is compiled of sex, age, population and head pose matched images.

Acknowledgements

The author is grateful for the support of the New South Wales Police, and in particular Detective Sergeant Tim Attwood, NSW Police Force Homicide Squad (Australia). The author also appreciates the comments provided by the anonymous Reviewers of this research, who were instrumental in strengthening the methods and results. However, all errors in the manuscript, and in the facial approximation, are the sole responsibility of the author.

References

- [1] Hayes S. Facial approximation of 'Angel': Case specific methodological review. *Forensic Sci Int.* 2014;237:e30-e41.
- [2] Cunningham R. Information pours in after murder victims identified. Channel 7 News. <https://au.news.yahoo.com/video/watch/29883079/police-inundated-with-information-following-identification-of-karlie-stevenson-and-kandalyce/#page122> October 2015.
- [3] Fleming K. Amazing likeness as 'Angel' identified. *The West Australian*. Australia: Seven West Media; 25 October 2015.
- [4] Stephan CN. Beyond the sphere of the English facial approximation literature: Ramifications of German papers on Western method concepts. *J Forensic Sci.* 2006;51:736-9.
- [5] Prag J, Neave R. *Making Faces: Using Forensic and Archaeological Evidence*. London: British Museum Press; 1997.
- [6] Clement JG, Ranson DL. Craniofacial identification. In: Clement JG, Ranson DL, editors. *Craniofacial identification in forensic medicine*. London: Arnold; 1998. p. 3-8.
- [7] Clement JG, Thomas CDL, Taylor RG, Craig PJG, Ismaniati N, Sheridan CS, et al. The dilemma of the forensic artist. *First International Conference on Forensic Human Identification*. London 1999.
- [8] Megreya AM, Burton AM. Unfamiliar faces are not faces: evidence from a matching task. *Mem Cognit.* 2006;34:865.
- [9] Sinha P, Balas BJ, Ostrovsky Y, Russell R. Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About. *Proceedings of the IEEE.* 2006;94:1948.
- [10] Freiwald W, Yovel G, Duchaine B. *Face Processing Systems: From Neurons to Real World Social Perception*. *Annu Rev Neurosci.* 2016;39.
- [11] Mike Burton A. Why has research in face recognition progressed so slowly? The importance of variability. *The Quarterly Journal of Experimental Psychology.* 2013;66:1467-85.
- [12] Davies G, Ellis HD, Shepherd J. Cue saliency in faces as assessed by the 'Photofit' technique. *Perception.* 1977;6:263-9.
- [13] Carbon C-C. *Face Processing: early processing in the recognition of faces [Doctoral Dissertation]*. Berlin: Freie Universitat Berlin; 2002.
- [14] Burton AM, Wilson S, Cowan M, Bruce V. Face recognition from poor quality video: evidence from security surveillance. *Psychol Sci.* 1999.
- [15] Sandford A, Burton AM. Tolerance for distorted faces: Challenges to a configural processing account of familiar face recognition. *Cognition.* 2014;132:262-8.
- [16] Davies-Thompson J, Newling K, Andrews TJ. Image-invariant responses in face-selective regions do not explain the perceptual advantage for familiar face recognition. *Cereb Cortex.* 2012:bhs024.
- [17] Barragan-Jason G, Cauchoux M, Barbeau EJ. The neural speed of familiar face recognition. *Neuropsychologia.* 2015;75:390-401.
- [18] Bruce V, Henderson Z, Newman C, Burton AM. Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied.* 2001;7:207-18.
- [19] Fernandes CMS, da Costa Serra M, da Silva JVL, Yoshito Noritomi P, de Sena Pereira FDA, Melani RFH. Tests of one Brazilian facial reconstruction method using three soft tissue depth sets and familiar assessors. *Forensic Sci Int.* 2012;214:211.e1-e7.
- [20] Solla HE, Iscan MY, McCabe B. Identification of the skeletal remains of a child in Uruguay. *The Forensic Examiner.* 2004;13:32-6.
- [21] Stephan CN, Cicolini J. The reproducibility of facial approximation accuracy results generated from photo-spread tests. *Forensic Sci Int.* 2010;201:133-7.
- [22] Stephan CN, Arthur RS. Assessing facial approximation accuracy: how do resemblance ratings of disparate faces compare to recognition tests? *Forensic Sci Int.* 2006;159, Supplement:S159-S63.
- [23] Stephan CN, Cicolini J. Measuring the accuracy of facial approximations: a comparative study of resemblance rating and face array methods. *J Forensic Sci.* 2008;53:58-64.
- [24] Farkas LG. Photogrammetry of the face. In: Farkas LG, editor. *Anthropometry of the Head and Face*. 2nd Edition ed. New York: Raven; 1994. p. 79-88.

- [25] Hayes S, Tullberg C. Police witness identification images: a geometric morphometric analysis. *J Forensic Sci.* 2012;57:1487-94.
- [26] Hayes S, Milne N. What's wrong with this picture? An experiment in quantifying accuracy in 2D portrait drawing. *Visual Communication.* 2011;10:149-74.
- [27] Hayes S, Sutikna T, Morwood M. Faces of *Homo floresiensis* (LB1). *Journal of Archaeological Science.* 2013;40:4400-10.
- [28] Hayes S. Seeing and Measuring the 2D Face: PhD Thesis. University of Western Australia; 2010.
- [29] ABC News. Murder victim Karlie Pearce-Stevenson's identity stolen, mobile phone used after death (27 October, 5.07pm). <http://www.abc.net.au/news/2015-10-27/karlie-pearce-stevenson-identity-stolen-phone-used-after-death/68886382015>.
- [30] Benny-Morrison A. Karlie Pearce-Stevenson's daughter's birth certificate found at house allegedly linked to accused killer Daniel James Holdom. *The Sydney Morning Herald* <http://www.smh.com.au/nsw/girl-in-suitcase-birth-certificate-found-at-house-allegedly-linked-to-accused-killer-20151111-gkwnbn.html>; Fairfax Media; 2015.
- [31] News Limited. The Australian crimes in 2015 that shocked us. *newscomau*. <http://www.news.com.au/national/crime/the-australian-crimes-in-2015-that-shocked-us/news-story/305841a6ba56f8801c585ae219314c2a>; News Limited; 2015.
- [32] Hayes S. Facial Identification Report: Missing Person. Coronial Investigation Unit, Western Australia Police. 2014.
- [33] Hayes S. Facial Identification Report: Adult Women. State Crime Command, Victoria Police. 2014.
- [34] Hayes S. Facial Identification Report: Subadults. State Crime Command, Victoria Police. 2014.
- [35] Hayes S. Facial Identification Report: Kinship Analysis. Australian Federal Police. 2015.
- [36] Gavan J, Washburn S, Lewis P. Photography: an anthropometric tool. *Am J Phys Anthropol.* 1952;10:331-54.
- [37] Geoghegan B. The determination of body measurements, surface area and body volume by photography. *Am J Phys Anthropol.* 1953;11:97-120.
- [38] Eliášová H, Krsek P. Superimposition and projective transformation of 3D object. *Forensic Sci Int.* 2007;167:146-53.
- [39] Stephan CN. Perspective distortion in craniofacial superimposition: Logarithmic decay curves mapped mathematically and by practical experiment. *Forensic Sci Int.* 2015;257:520.e1-e8.
- [40] Mitteroecker P, Windhager S, Müller GB, Schaefer K. The Morphometrics of "Masculinity" in Human Faces. *PLoS ONE.* 2015;10:e0118374.
- [41] Meindl K, Windhager S, Wallner B, Schaefer K. Second-to-fourth digit ratio and facial shape in boys: the lower the digit ratio, the more robust the face. *Proceedings of the Royal Society of London B: Biological Sciences.* 2012;rsb20112351.
- [42] Anies OS, Torres MAJ, Manting MM, Demayo CG. Landmark-Based Geometric Morphometrics in Describing Facial Shape of the Sama-Banguingui Tribe from the Philippines. *Journal of Medical and Bioengineering Vol.* 2013;2.
- [43] Perrett DI, May K, Yoshikawa S. Facial shape and judgements of female attractiveness. *Nature.* 1994;368:239-42.
- [44] Rhodes G, Sumich A, Byatt G. Are average facial configurations attractive only because of their symmetry? *Psychol Sci.* 1999;10:52-8.
- [45] Minear M, Park DC. A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers.* 2004;36:630-3.
- [46] AAP. Search halted for missing Finnish exchange student. *Sydney Morning Herald.* Sydney: Fairfax Media; 2011.
- [47] Internet Images F. https://500px.com/photo/10600575/evil-by-bartek-g%C4%85sierek?ctx_page=1&from=user&user_id=272925; [http://vi.sualize.us/power close up black and white face sephia portrait freckles picture kXgV.html](http://vi.sualize.us/power%20close%20up%20black%20and%20white%20face%20sephia%20portrait%20freckles%20picture%20kXgV.html); https://florianadamblog.files.wordpress.com/2014/12/img_6283.jpg; <http://pelfind.co/photo/56149p1293851/colorful-unusual-women-portraits-by-emma-uber>; [http://clairegill.photography/wp-content/uploads/2014/12/27-1-post/141022Rosie002\(pp_w599_h900\).jpg](http://clairegill.photography/wp-content/uploads/2014/12/27-1-post/141022Rosie002(pp_w599_h900).jpg); <http://favim.com/orig/201104/29/Favim.com-28029.jpg>;

- <https://s-media-cache-ak0.pinimg.com/564x/4e/e1/28/4ee128e23fac27e54cf006ba33f39e13.jpg>; <http://cache.desktopnexus.com/thumbseg/1191/1191574-bigthumbnail.jpg>; www.imagebase.net/People/girlface; <http://www.acnenadvice.com/images/acne-miracle-pro.jpg>; <http://i2.cdn.turner.com/cnnnext/dam/assets/160203182545-bitch-face-large-169.jpg>; <http://hairzstyle.com/wp-content/uploads/2015/08/2-hairstyle-tips-for-round-faces-11.jpg>; http://cache3.asset-cache.net/x/476951987.jpg?v=2&c=IWSAsset&k=2&d=Qu34XwHPS5CienYdXd-e5azNNMJuxOh6oQrceLOtApGGZEi_gJvVKV4P_43sE86-0; <http://cache4.asset-cache.net/gc/83455374-young-woman-with-a-necklace-close-up-gettyimages.jpg?v=1&c=IWSAsset&k=2&d=7dfFxpOik8JNxs1IilpqINdH16h5qXLRTyEPH7CPod0DrHxYgX08SdebEORB6X8O>; https://encrypted-tbn3.gstatic.com/images?q=tbn:ANd9GcR6C21kbERKZavth6aTcc8WkJHU6uHk4Rdh4xwx4DwnQklv8_qh5A; <http://img4.bdbphotos.com/images/230x300/v/w/vw3cyt23hwxbtc2y.jpg?skj2io4l>; <http://imgc.allpostersimages.com/images/P-473-488-90/76/7666/X5NG300Z/posters/1930s-portrait-of-brunette-woman-wearing-printed-blouse-half-smile-mona-lisa-like-beautiful-eyes.jpg>. 2016.
- [48] McEntire C. Three-dimensional soft tissue changes upon smiling: Virginia Commonwealth University; 2013.
- [49] Okamoto H, Haraguchi S, Takada K. Laterality of asymmetry in movements of the corners of the mouth during voluntary smile. *The Angle Orthodontist*. 2010;80:223-9.
- [50] Hammer Ø, Harper D, Ryan P. PAST-Palaeontological statistics. www.uv.es/~pardonmv/pe/2001_1/past/pastprog/past.pdf, accessed em. 2001;25:2009.
- [51] Rohlf FJ. tpsDig. 2.12 ed: Department of Ecology and Evolution, State University of New York at Stony Brook; 2008.
- [52] Klingenberg CP. MorphoJ: an integrated software package for geometric morphometrics. *Molecular ecology resources*. 2011;11:353-7.
- [53] Rohlf FJ. The tps series of software. *Hystrix, the Italian Journal of Mammalogy*. 2015;26:9-12.
- [54] O'Higgins P, Jones N. Tools for statistical shape analysis. Hull York Medical School; 2006.
- [55] Stephan CN, Simpson EK. Facial soft tissue depths in craniofacial identification (part 1): an analytical review of the published adult data. *J Forensic Sci*. 2008;53:1257-72.
- [56] George RM. Anatomical and artistic guidelines for forensic facial reconstruction. In: Iscan MY, Helmer RP, editors. *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*. New York: Wiley-Liss; 1993. p. 215-27.
- [57] Caple JM, Stephan CN, Gregory LS, MacGregor DM. Effect of Head Position on Facial Soft Tissue Depth Measurements Obtained Using Computed Tomography. *J Forensic Sci*. 2015.
- [58] Aulsebrook W, Becker P, Yagar M. Facial soft-tissue thicknesses in the adult male Zulu. *Forensic Sci Int*. 1996;79:83-102.
- [59] Caple J, Stephan CN. A standardized nomenclature for craniofacial and facial anthropometry. *Int J Legal Med*. 2015;1-17.
- [60] Taylor RG, Angel C. Facial reconstruction and approximation. In: Clement JG, Ranson DL, editors. *Craniofacial Identification in Forensic Medicine*. London: Arnold; 1998. p. 177-86.
- [61] Hayes S. Faces in the Museum: Revising the methods of facial reconstructions. *Museum Management and Curatorship*. 2015;1-28.
- [62] Larrabee W, Makielski K. *Surgical Anatomy of the Face*. New York: Raven; 1993.
- [63] Guyomarc'h P, Dutailly B, Couture C, Coqueugniot H. Anatomical placement of the human eyeball in the orbit - validation using CT Scans of living adults and prediction for facial approximation. *J Forensic Sci*. 2012;1271-5.
- [64] Stephan C, Davidson PL. The placement of the human eyeball and canthi in craniofacial identification. *J Forensic Sci*. 2008;53:612-9.
- [65] Stephan CN, Huang AJR, Davidson PL. Further evidence on the anatomical placement of the human eyeball for facial approximation and craniofacial superimposition. *J Forensic Sci*. 2009;54:267-9.

- [66] Song W-C, Kim S-H, Paik D-J, Han S-H, Hu K-S, Kim H-J, et al. Location of the infraorbital and mental foramen with reference to the soft-tissue landmarks. *Plast Reconstr Surg*. 2007;120:1343-7.
- [67] Stephan C, Murphy S. Mouth width prediction in craniofacial identification: cadaver tests of four recent methods, including two techniques for edentulous skulls. *J Forensic Odontostomatol*. 2008;27:2-7.
- [68] Stephan C, Henneberg M. Predicting mouth width from inter-canine width-a 75% rule. *J Forensic Sci*. 2003;48:725-7.
- [69] Stephan CN. Facial approximation: an evaluation of mouth-width determination. *Am J Phys Anthropol*. 2003;121:48-57.
- [70] Rynn C, Wilkinson M, Peters H. Prediction of nasal morphology from the skull. *Forensic Science, Medicine and Pathology*. 2010;6:20-34.
- [71] Mala PZ. Pronasale Position: An Appraisal of Two Recently Proposed Methods for Predicting Nasal Projection in Facial Reconstruction. *J Forensic Sci*. 2013;58:957-63.
- [72] Wilkinson C, Motwani M, Chaing E. The relationship between the soft tissues and the skeletal detail of the mouth. *J Forensic Sci*. 2003;48:728-32.
- [73] George RM. The lateral craniographic method of facial reconstruction. *J Forensic Sci*. 1987;32:1305-30.
- [74] Taylor KT. *Forensic Art and Illustration*. Boca Raton: CRC Press; 2001.
- [75] Standing S. *Gray's Anatomy: The Anatomical Basis of Clinical Practice*. 40th ed: Elsevier; 2008.
- [76] Mala PZ, Veleminska J. Vertical lip position and thickness in facial reconstruction: a validation of commonly used methods for predicting the position and size of lips. *J Forensic Sci*. 2016:n/a-n/a.
- [77] Fedosyutkin B, Nainys J. The relationship of skull morphology to facial features. In: Iscan MY, Helmer RP, editors. *Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification*. New York: Wiley-Liss; 1993. p. 119-213.
- [78] Maurer D, LeGrand R, Mondloch CJ. The many faces of configural processing. *Trends Neurosci*. 2002;6:255.
- [79] Hole GJ, George P, Eaves K, Razek A. Effects of geometric distortions on face-recognition performance. *Perception*. 2002;31:1221.
- [80] Haig ND. The effect of feature displacement on face recognition. *Perception*. 1984;13:505.
- [81] Hosie JA, Ellis HD, Haig ND. The effect of feature displacement on the perception of well-known faces. *Perception*. 1988;17:461.
- [82] Brédart S, Devue C. The accuracy of memory for faces of personally known individuals. *Perception*. 2006;35:101-6.
- [83] Ge L, Luo J, Nishimura M, Lee K. The lasting impression of Chairman Mao: hyperfidelity of familiar-face memory. *Perception*. 2003;32:601-14.

Figure captions

Figure 1. The facial approximation of the Belanglo 'Angel' and Ms Pearce-Stevenson

The facial approximation (a-b), one of the photographs of Ms Pearce-Stevenson that has appeared in the Australian media (c), wireframes representing the three photographs used in this evaluation: frontal/downwardly posed (d), right turned/downwardly posed (e), right turned/upwardly posed (f). The photograph of Ms Pearce-Stevenson is reproduced with permission (refer main text).

Figure 2. Facial landmark coordinates (n=40)

This suite of 40 landmarks were applied to the facial approximation, the UK average and each of the 64 database photographs for the geometric morphometric analyses. The white horizontal lines are parallel with the interpupillary plane (a horizontal line connecting landmarks 1 and 2). The photograph is one of the PAL Database images, coded within this database as Shape 56 (access to the Database includes permission to reproduce the images [45]). See Table 1 for the landmark definitions.

Figure 3. PC1 and PC2, database facial average, and the UK facial average

Head pose variation within the database images: blue rectangles PAL (light blue) and Utrecht (dark blue) Face Database images, black triangles Pasco Sheriff Arrest Database images, pink circles internet images. The UK facial average is circled and indicated with an arrow. The sepia data points represent the same analysis without the UK facial average. Insert figure 3a is the facial average of the database images, insert b is an overlay of the wireframe of the database average (blue) and the UK facial average (black).

Figure 4. Morphological differences between the facial approximation and the three photographs of Ms Pearce-Stevenson

A: The multivariate regression results of the three photographs of Ms Pearce-Stevenson and the database images. The left wireframe is how Ms Pearce-Stevenson differs from the database, the right wireframe overlay is a comparison of Ms Pearce-Stevenson and the database. The centre images are the TPS deformation grids arising from the multivariate regression. Left centre is how Ms Pearce-Stevenson differs from the database, right centre is how the database differs from Ms Pearce-Stevenson.

B: The multivariate regression of the facial approximation and the database images. The left wireframe is how the facial approximation differs from the database, the right wireframe overlay is a comparison of the facial approximation and the database. The centre images are the TPS deformation grids arising from the multivariate regression. Left centre is how the

facial approximation differs from the database, right centre is how the database differs from the facial approximation.

C: The morphological differences between the facial approximation and Ms Pearce-Stevenson. The centre image is how the results of A and B compare (centre wireframe overlay and marked within a rectangle; the green wireframe is the facial approximation, the black wireframe is Ms Pearce-Stevenson). The left wireframe overlay shows how a regression involving only two images of Ms Pearce-Stevenson (red wireframe) compares to the comparisons involving all three images (black wireframe). The right wireframe overlay shows how a regression involving only two images of Ms Pearce-Stevenson (red wireframe) compares to the facial approximation.

Note that these wireframes are mathematical constructs illustrating the variance following the multivariate regression analyses, and do not represent the actual coordinate data of the photographs of Ms Pearce-Stevenson, the facial approximation or the database average.

Figure 5. The morphological similarity of the facial approximation with Ms Pearce-Stevenson (PC3 and PC 8)

The facial approximation clusters closely with two photographs of Ms Pearce-Stevenson on PC3 and PC8. The minimal spanning tree indicates the closest affinities within the scatter plot and the 95% confidence ellipsis. The wireframes illustrate the variance being captured by each PC, and are taken from the maximum and minimum values of the data points from each axis. The photographs of Ms Pearce-Stevenson are red diamonds (circled), the facial approximation is an open square. The cluster (shown encapsulated in a rectangle) involves the two photographs of Ms Pearce-Stevenson (Figure 1d, f) and the facial approximation. The blue rectangles are the PAL and Utrecht Face Database images, the black triangles are the Pasco Sheriff Arrest Database images, and the pink circles are the internet images.

Table 1: Landmark Definitions

<i>Right</i>	<i>Left</i>	<i>Mid</i>	<i>Landmark</i>	<i>Landmark Definition</i>
1	2		<i>pupil</i>	Centre of the pupils. Where not visible, approximate from the centre of the iris
3	6		<i>lateral iris</i>	Lateral border of the iris on the IPL*
4	5		<i>medial iris</i>	Medial border of the iris on the IPL*
7	13		<i>exocanthion</i>	The outer eye corner
8	12		<i>superior lid</i>	The highest point of the upper eyelid at the lash root line immediately above the pupil
9	11		<i>endocanthion</i>	The inner eye corner
10	14		<i>inferior lid</i>	The lowest point of the lower eyelid immediately below the pupil and at the root of the lashes
15	19		<i>alare height</i>	The highest point of the nose wing on the edge of the nose (often more medially located than the wing width)
16	18		<i>alare wing width</i>	The widest points of the nose wing
		17	<i>subnasale</i>	Where the septum meets the philtrum at the base of the nose. When not visible, approximate as slightly lower than the alare base
20	24		<i>cheilion</i>	The most lateral point of the oral fissure, before it turns downwards at the mouth corners. Often this is an intersection point with a lip edge
21	23		<i>crista philtri</i>	Where the vermillion border of the upper lip meets the philtrum
		22	<i>superior upper lip</i>	The highest midpoint of the upper lip
		27	<i>inferior upper lip</i>	The lowest midpoint of the upper lip
		26	<i>superior lower lip</i>	The highest midpoint of the lower lip. When the mouth is closed this will be the same as the inferior upper lip
		25	<i>inferior lower lip</i>	The lowest midpoint of the lower lip
40	28		<i>upper face</i>	The widest point of the face at the lower eye lid following a horizontal plane parallel to the IPL*
	46		<i>menton</i>	The lowest point of the chin often on the midline
36-39	29-33		<i>outer face</i>	10 semi-landmarks (2 x 5) between the upper face and menton

* IPL – Interpupillary Line: A horizontal line connecting the midpoint of the pupils

Table 2. Morphological analysis of the database composition (*morphologika2*)

Independent variable includes the number of PCs and % of cumulative variance (cum. var.). Partial regression coefficients with the largest effect (PC variance x adjusted R^2) are shown bolded. S = standard error. See Methods for details of independent variable coding. The scatter plot for PC1 and PC2 is shown in Figure 3.

	Independent variable	n	Multivariate regression results	Partial regression coefficients					
				PC	PC var.	R^2 (adj.)	S	<i>p</i> -value	UK $X\ddagger$
1	Head pitch PC1-20 98 % cum. var.	65	24.5 % var. $p < 0.0001$	PC 1*	38.1 %	0.574	0.005	< 0.0001	-0.039
2	Head turn PC 1-20 98 % cum. var.	65	18.5 % var. $p < 0.0001$	PC 2* PC 1*	28.3 % 38.1 %	0.350 0.180	0.006 0.007	< 0.0001 < 0.001	-0.036
3	Face shape PC 1-20 98 % cum. var.	65	9.0 % var. $p < 0.0001$	PC 3* PC 1* PC 7	11.8 % 38.1% 1.9 %	0.402 0.047 0.066	0.003 0.007 0.001	< 0.0001 = 0.05 < 0.02	-0.002 +0.005 +0.005
4	Facial expression PC 1-20 98 % cum. var.	65	3.6 % var. $p = 0.0001$	PC 4* PC 5	5.3 % 2.7 %	0.259 0.056	0.004 0.003	< 0.0001 = 0.03	-0.004 -0.003
5	Image source PC 1-20 98 % cum. var.	64	19.1 % var. $p < 0.0001$	PC 1* PC 2*	38.1 % 28.1 %	0.282 0.230	0.006 0.005	< 0.0001 < 0.0001	<i>n/a</i>
6	UK average PC 1-20 98 % cum. var.	65	1 % var. <i>n. significant</i>	-	-	-	-	-	<i>see 1-4 above</i>

*PCs marked with an asterisk are significant within a PCA (relative warps) in PAST (bootstrap 1000, above broken stick and within the 95% confidence interval).

Table 3. Morphological analysis of Ms Pearce-Stevenson and the facial approximation (*morphologica2*)

Independent variable includes the number of PCs and % of cumulative variance (cum. var.). Partial regression coefficients (PC) with the largest effect (PC var. x adjusted R^2) are shown bolded, results that borderline according to the standard error ($S < 0.025$) are shown in brackets. See Methods for details of independent variable coding. The wireframes and TPS deformation grids resulting from multivariate regressions 7 and 8 (PC 3-20) are shown in Figure 4, the scatter plot of PC 3 and 8 (multivariate regression 9) is shown in Figure 5.

	Independent variable	n	Multivariate regression results	Partial regression coefficients				
				PC	PC var.	R^2 (adj.)	S	p-value
7	Ms Pearce-Stevenson (x 3) PC 1-20 98 % cum. var. PC 3-20 30 % cum. var.	67	PC 1-20 4.7 % var. $p = 0.0001$ PC 3-20 2.1 % var. $p = 0.004$	PC 2 (PC3) PC 8 PC 10	30.1 % 11.5 % 1.3 % 0.8 %	0.075 (0.018) 0.172 0.108	0.023 (0.01) 0.004 0.004	= 0.01 (= 0.1) < 0.001 = 0.004
7a	Ms Pearce-Stevenson (x 2) PC 3-20 31 % cum. var.	66	PC 3-20 3.0 % var. $p = 0.001$	PC3 PC 8 PC 10 PC 11	11.9 % 1.3 % 0.9 % 0.08 %	0.043 0.144 0.128 0.068	0.018 0.005 0.005 0.004	= 0.05 < 0.001 = 0.002 = 0.02
8	Facial approximation PC 1-20 98 % cum. Var. PC 3-20 32 % cum. var.	65	PC 1-20 2.7% var. <i>n. significant</i> PC 3-20 2.3 % var. <i>n. significant</i>	(PC 3)	(12.3 %)	(0.036)	(0.025)	(= 0.07)
9	Ms Pearce-Stevenson (x 3) & Facial approx. PC 1-20 98 % cum. var. PC 3-20 30 % cum. var.	68	PC 1-20 3.4 % var. $p < 0.001$ PC 3-20 3.8 % var. $p = 0.002$	PC 3 PC 8 PC 10 PC 17	11.2 % 1.3 % 0.8 % 0.3 %	0.066 0.171 0.052 0.069	0.01 0.004 0.003 0.002	= 0.02 < 0.001 = 0.03 = 0.02

Table 4. Similarity-distance matrix for the 3 photographs of Ms Pearce-Stevenson, the facial approximation and the PAL Database image (database no. 56, and see Figure 3). The rank of the matrix score is in brackets (1-67, most similar to most distant), the matrix score which is the most similar is shown in bold.

	Ms Pearce-Stevenson			Facial Approx.
	Front/down	Right/down	Right/up	
PAL Database	0.037 (1)	0.053 (49)	0.039 (2)	0.024 (1)
Front/down	-	0.060 (60)	0.044 (7)	0.038 (2)
Right/down	0.060 (19)	-	0.046 (12)	0.057 (56)
Younger Right/up	0.044 (4)	0.046 (29)	-	0.043 (5)
Facial Approx.	0.038 (2)	0.057 (56)	0.043 (5)	-
Range	0.037-0.109	0.030-0.081	0.036-0.089	0.024-0.105

a

b

c

d

e

f

a

b

A

B

C