Carl N. Stephan,[1] B.H.Sc. and Maciej Henneberg,[1] D.Sc.

Building Faces from Dry Skulls: Are They Recognized Above Chance Rates?*

REFERENCE: Stephan CN and Henneberg M. Building faces from dry skulls: are they recognized above chance rates? J Forensic Sci 2001;46(3):432–440.

ABSTRACT: Methods of facial approximation have successfully aided the identification of deceased individuals. Successes may be due to either accurate facial approximation techniques or chance. This study aims to determine if any of 16 facial approximations, built using standard techniques, are sufficiently accurate to produce correct identifications of target individuals above chance. Four skulls were approximated using four commonly used methods of facial approximation. The resulting 16 facial approximations were judged by 37 assessors of varying ages. Assessors attempted to identify the target individual of each facial approximation from a face pool of ten photographed faces. Only one facial approximation resulted in true positive identification rates above chance at statistically significant levels. It is concluded that it is rare for facial approximations to be sufficiently accurate to allow identification of a target individual above chance. Since 403 incorrect identifications were made out of 592 identification scenarios, facial approximation should be considered to be a highly inaccurate and unreliable forensic technique. These results suggest that facial approximations are not very useful in excluding individuals to whom skeletal remains may not belong. Evidence from this experiment supports suggestions by others that facial approximation should be used in forensic science when all other methods of identification have failed and only to provide tentative identification.

KEYWORDS: forensic science, facial reconstruction, facial reproduction, facial approximation, human identification, skeletal remains, recognition, accuracy, face pools

The term "facial approximation" refers to the process of building a face over a skull to create an image of what a person looked like during life. This technique has often been used in forensic science when attempting to identify human remains (1,2) and has also been used to recreate the facial appearance of hominids (3) and modern humans (4). Facial approximation has been referred to as facial reconstitution (1), facial restoration (5), facial reproduction (6), facial reconstruction (7), and forensic sculpture (8). Facial approximation techniques are varied, encompassing drawing (9,10), sculpting (4,8,11), and computer generated methods (12,13). Facial approximation techniques also differ in their approach. Some techniques require building of the facial muscles, e.g., Gerasimov's technique (3), and are often referred to as Russian methods (4). Others techniques, e.g., Gatliff's technique (8), use average measurements to determine facial soft tissue depths, and are often referred to as American techniques (4). Others use a combination of the Russian and American techniques, e.g., Neave's technique (4). The purpose of forensic facial approximation is to promote recognition of the person to whom the skull belonged (the target individual). An accurate forensic facial approximation should, therefore, be easily recognized as the person to whom the skull belonged.

Published rates of successful identifications from facial approximations are generally high, suggesting that methods of facial approximation are accurate. Gerasimov (3) claims 100% success, Bender (14) 85%, and Gatliff (15) 70%. Facial approximation has also been reported to be a useful forensic tool because it can be used to exclude suspected individuals (16,17).

The accuracy of facial approximation techniques has been previously tested by either comparing a facial approximation to the target individual to determine the facial approximation's *similarity* (1,4,18,19), or by comparing a facial approximation to a face pool to determine the ability for a target individual to be *recognized* from a facial approximation (16,17).

From the direct comparison (i.e., resemblance ratings) of a facial approximation to a target individual, it appears that facial approximations are accurate. Krogman (19) reports a facial approximation that was "recognizable as that of the subject chosen." Suzuki (1) reports that "the resemblance between the two [a target individual and a facial approximation] was quite striking." Helmer et al. (18) conclude that "in general it can be said that at least a slight and often even a close resemblance was achieved" from the facial approximations to target individuals. Prag and Neave (4) also report a "reconstructed face [that] bore an uncanny resemblance to the photograph [of the target individual]."

By asking assessors to identify a target individual from a range of faces (i.e., a face pool) after examining a facial approximation, identification rates for each face can be calculated. Snow et al. (17) found identification rates to be significantly above chance for a male and a female facial approximation made using a sculpting, American technique. Assessors attempted to identify the antemortem photograph of the target individual from a face pool of seven photographed faces by forced choice method, i.e., assessors were forced to identify a face even if they thought the facial approximation did not correspond to any face. The male facial approximation was examined by 200 assessors with a true positive identification rate of 68% (17). The female facial approximation was examined by 102 assessors with a true positive identification rate of 26% (as reported in Table 2 of the Snow et al.'s (17) results, text and Table 1 state 104 assessors). Although incorrect identifications were also made they were not abundant. Two nontarget faces were often identified from the female approximation (17) and

[1] Department of Anatomical Sciences, University of Adelaide, Adelaide 5005, Australia.

* Portions of this work have been previously presented at the 14th International Symposium on the Forensic Sciences.

Received 8 Oct. 1999; and in revised form 28 March 2000, 20 June 2000; accepted 22 June 2000.

appear to be above chance at statistically significant rates ($p \leq 0.07$). The results of this experiment indicate that the 3D sculpting American technique of facial approximation is fairly accurate.

An average identification rate of 26% has been recorded by van Rensburg (16) who tested fifteen facial approximations. Individual identification rates for each facial approximation were not reported, however, the average rate was 19% above chance rates (16). The facial approximations were made using a three dimensional (3D) sculpting combination method that requires building the facial muscles and soft tissues whilst taking into account average soft tissue depths. The facial approximations were compared to death masks. Comparison of facial approximations to death masks in this study may not be close to reality since forensic facial approximation is aimed at building faces that are representative of living individuals whose facial appearance and soft tissue structure differs from that of a deceased person's face.

Suk (20), Montagu (21), and Brues (22) have expressed doubt that accurate facial approximations can be easily produced because many facial soft tissues leave little clue to their physical structure on the skull. Consequently, facial approximation is expected to involve subjective interpretation of skull/face relationships and is expected to be inaccurate and unreliable.

Since a successful forensic facial approximation depends on the facial approximation being *recognizable* as the target individual, face pool comparison appears to be a more reliable method of assessing a facial approximation's accuracy than resemblance ratings. Resemblance ratings appear not to be optimal since they measure the *similarity* between the facial approximation and the target individual and not the ability for the target individual to be recognized from a group of faces. Resemblance ratings also ignore the facial approximation's resemblance to nontarget individuals, which may be greater than the target individual's. The identification rates of target individuals, from the few face pool comparison studies conducted, despite being rather low, have tended to be above chance. This appears to indicate that facial approximation techniques, despite being inaccurate, may actually work.

The present study aims to objectively determine the accuracy of four commonly used methods of forensic facial approximation: a 3D sculpting American method; a two dimensional (2D) drawing American method; a 2D computer "FACE" assisted American method; and a 3D sculpting combination method. In the case of 3D approximations being identified above chance levels, an additional investigation was made to determine the effect of facial pose on the identification rate.

Materials and Methods

Skulls

Four dry skulls were used for facial approximation. Each had a corresponding recent antemortem photograph that was concealed from the investigator (CNS) involved in the anthropological identification and facial approximation process. Only information interpreted from the dry skulls was used in the facial approximation process. For this reason, facial approximations were not complemented with make up, or with head or facial hair, as this could in no way be determined from the skull. This method ensured that *facial approximation methods alone* were tested for their accuracy. Sex, age, and population of origin were estimated by visual examination of the skull, following standard anthropological methods (Table 1). Each skull was given a fictitious name for ease of reference (Table 1).

Skull Casting and Preparation

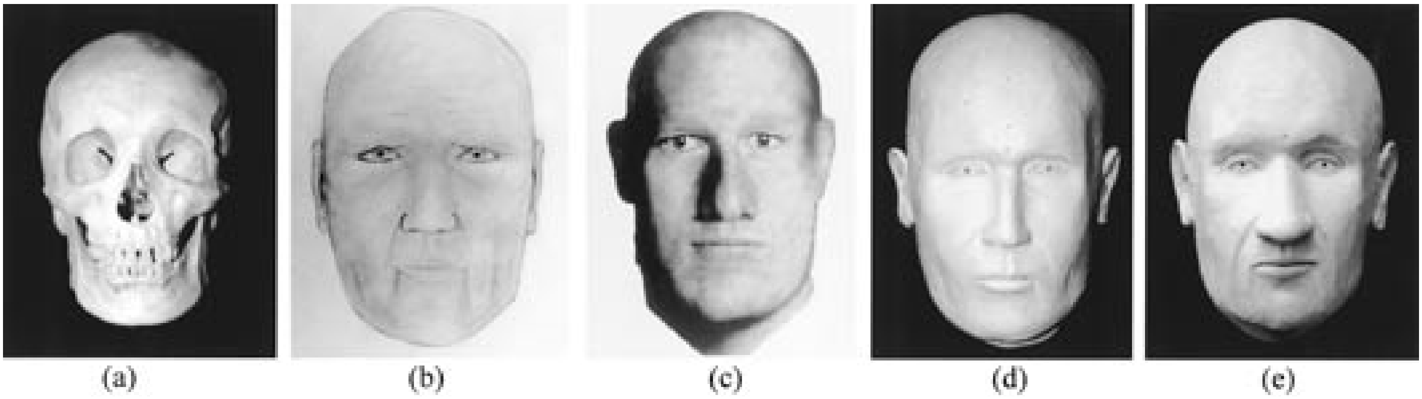
Each cranium and mandible was cast separately using a split mould technique. The cranium was cast in plaster of Paris and the mandible was cast in cold cure resin. The process was repeated, producing two casts of each skull. All casts were cross-checked against the original skulls using GPM® sliding and spreading calipers to ensure accurate dimensions. The mandible was rearticulated with the cranium in its appropriate position using dental wax, which also simulated the interarticular disc of the tempero-mandibular joint. Teeth which had been lost postmortem were replaced by casting the opposite side tooth and gluing it in place. This method was used, despite its crudeness, because it was quick and facial approximations with visible front teeth (e.g., produced by a class II malocclusion) would appear more realistic if teeth were "replaced" in contrast to filling the gap with a wax arch. Skull casts were mounted in the Frankfurt horizontal plane and small bore holes were made at 34 reference points (23). Wooden dowels, 2.5 mm in diameter, were then glued into the bore holes. Dowels were marked at appropriate average soft tissue depths according to Helmer (23) using a metal ruler to the nearest 0.5 mm and trimmed to appropriate depth. Helmer's soft tissue depths were used since they are calculated from ultrasound measurements on living people and are, therefore, likely to be a more accurate representation of a living person's face in contrast to measurements taken on cadavers. One cast of each skull (with soft tissue markers in place) was photographed in the full face position, using a digital camera. These images were used for facial approximations made by the 2D computer "FACE" assisted American method and the 2D drawing American method.

Methods of Facial Approximation

Four facial approximations were made of each skull. All skulls were positioned in the Frankfurt horizontal plane during facial approximation. As an example, Fig. 1 shows completed approximations for each method, on one skull.

TABLE 1—*Details of skulls used for facial approximation.*

		Sex		Age (years)		Ancestry	
Skull	Name	Estimated	Actual	Estimated	Actual	Estimated	Actual
1	"Sam,"	male	male	55 ± 5	50–60	White	White
2	"Fred,"	male	male	45 ± 5	40	White	White
3	"Kate,"	female	female	28 ± 5	20–30	White	White
4	"Jane"	female	female	32 ± 5	30–40	White	White



NOTE: the actual target individual is not shown to protect confidentiality.

FIG. 1—*An example of a skull approximated with four methods of facial approximation.* (a) *"Sam's" skull;* (b) *2D drawing American facial approximation of "Sam's" skull;* (c) *2D computer "FACE" assisted American facial approximation of "Sam's" skull;* (d) *3D sculpting American facial approximation of "Sam's" skull;* (e) *3D sculpting combination facial approximation of "Sam's" skull.*

3D Sculpting American Method—

Step 1 Eyes—Eyeballs were anatomically shaped balls of plasticine, 26 mm in diameter for males and 25 mm for females (Neave, 1998 Personal Communication). The pupil was represented by depressing the plasticine with the writing end of a pencil. This eliminated the need for choosing a prosthetic eyeball with a colored iris, the color of which can not be determined from the skull. The eyeballs were positioned so that the eyeball and pupil were centrally located within the orbit (8,24). The eyeballs projected from the orbits so the apex of the cornea touched a centrally located tangent from the superior to the inferior orbital margins (8).

Step 2 Soft tissue—The soft tissue of the face was added by joining all the average soft tissue depths with strips of plasticine, approximately 5 mm in width, at the appropriate depth. This formed a triangular framework over the skull similar to the method of Stewart (11). This framework was blocked in with plasticine leaving the area around the nose and mouth.

Step 3 Mouth—The area of the mouth was blocked in to the appropriate soft tissue depth, with lines in the plasticine representing the mouth slit (8). The junction between the canine and first premolar, on both sides, determined the positions of the corners of the mouth (4,8). The stomion was located a third of the way up the central maxillary incisors for females, and a quarter of the way up for males (9). Each lip was equal in height to the corresponding enamel of the upper and lower anterior incisors (8).

Step 4 Nose—The shape of the nasal bridge was determined by the shape of the nasal bones according to George (9). The base of the nose was positioned a few millimeters below the anterior nasal spine (25). The angulation of the nasal septum was determined by the direction of the nasal spine (4,25). The projection of the nose from subnasale to pronasale was determined by a distance equal to 3× the length of the nasal spine, as measured from the lower margin of the nasal aperture to the tip of the spine (8,24). The width of the nose (measured across the wings) was determined by dividing the maximum width of the nasal aperture by 0.6 (4,24,25).

Step 6 Eyelids—The endocanthion and ectocanthion were determined in accordance with the apparent intentions of Krogman (26). The original citation appears to have confused *lateral* and *medial* and if followed would appear to place the corners of the eye slit outside the orbits. The endocanthion of each eye was positioned lateral to the lacrimal duct, 4 mm below the dacryon (27) and was approximately 3 mm *lateral* (reported as medial) to the medial border of the orbit (26). The ectocanthion of each eye was positioned at the level of Witnall's tubercle when present (28) and approximately 5 mm *medial* (reported as lateral) to the lateral border of the orbit (26). When Witnall's tubercle was not present the level of the ectocanthion was calculated according to Stewart (28). The lower lid approximated a straight line and was positioned so that its upper border was equal to the inferior border of the iris (8). The upper lid followed a more exaggerated curve and was positioned so that the inferior border covered some of the upper portion of the iris (8).

Step 7 Auricles—The auricles were the last facial feature to be added to the facial approximation. The general length of the auricles was estimated by the length of the nose (8,24,25). The ear canal was placed at a level equal to that of the external auditory meatus (24). The level of the superior border of the external auditory meatus was equal to the superior edge of the tragus (8). The position of the ear was determined by the artistic canon that it is equal to the relative position of the nose. The ear anatomy was "average" since no detailed anatomy of the ears can presently be determined from the skull.

Step 8 Final Harmonization—Once the face was complete, it was harmonized for a final time checking that all soft tissue depths were adhered to, and that the proportions and positions of the facial features worked harmoniously together. The facial approximation was then photographed in the Frankfurt horizontal plane in three views: full face, right three quarter (30° rotation from full face) and right profile.

2D Drawing American Method

The digitized pictures of the skull casts (with soft tissue depth measures in place) in full face view were printed onto 210 × 297 mm glossy paper using an Epson® 740 Stylus printer. Plain, 210 × 297 mm paper was superimposed over the printed picture and viewed on a light box. This enabled the skull to be viewed while the facial approximation was drawn onto the top piece of paper with B pencils. The outline of the face was drawn by joining the soft tissue depths. The eyes, nose, mouth, and ears were then approximated as in the 3D American technique, but by drawing in two dimensions. The projection of the face, particularly the nose, had to be represented two dimensionally and therefore required artistic interpretation when appropriately toning the face.

2D Computer "FACE" Assisted American Method

This method is similar to Ubelaker and O'Donnell's computer assisted technique (29) and the Face Imaging Reconstructive Morphography method (30), but differs because the "FACE" assisted method uses a photofit (facial feature photograph) identification system called FACE®, rather than an identikit system of artistic drawings. The digitized full face pictures of each of the skull casts (with soft tissue depths in place) were opened in Adobe® photoshop 4.0. A 1994 demonstration model of the FACE® system, at the University of Melbourne, was then used to select photographs of facial features for each skull. The face was divided into five basic features: eyes, nose, mouth, jaw, and hair. The FACE® system used had a limited selection of facial feature images available (approximately 20 or more for each facial feature for both sexes). Facial features were selected based on their appearance and shape which was in accordance with skull morphology. The face was approximated by morphing the features to the appropriate soft tissue depths in FACE®. By superimposing the facial features over the skulls in Adobe® Photoshop 4.0 and varying the opacity of picture layers, the facial features could be regularly checked for correct proportions against the skull and soft tissue depths. The most receding hair feature was selected so that minimal alterations in PhotoShop were needed to remove the hair completely. Once the matching and morphing of facial features was complete the edges of each facial feature image were blended with the next and the tone of each facial feature equalized by adjusting the brightness and contrast in the FACE® program. Completed facial approximations were printed on 210 × 297 mm photo quality glossy paper using an Epson® 740 Stylus printer.

3D Sculpting Combination Method

This method was the same as the 3D sculpting American method except for aspects outlined below.

Soft Tissue (Replaces Step 2 of the 3D Sculpting American Method)—The muscles of the face were built on the skull, in plasticine, according to the method of Prag and Neave (4). The position, size, and shape of the temporalis and masseter muscles was determined by the ruggedness, size, and shape of the origins and insertions of these muscles on the skull. The general size and robustness of the skull was also taken into account. The remaining facial muscles (buccinator, orbicularis oris, levator anguli oris, levator labii superioris, zygomaticus major and minor, depressor labii inferioris and depressor anguli oris and risorius) were built as an idealized model since their origins and insertions can not be precisely located. Support was given to the muscles by blocking under them with plasticine. This prevented the muscles from collapsing when covered with sheets of plasticine. An idealized model of the parotid gland was also added to the reconstruction. Sheets of plasticine, approximately 5 mm thick, were laid over the muscles. This layer of plasticine conformed to the contours of the underlying muscles and simulated the cutaneous and superficial subcutaneous tissue. Soft tissue depths were used as a guide, but were not strictly adhered to if the morphology of the skull suggested different depths.

Nose (Replaces Nose Projection Guideline in Step 4 of the 3D Sculpting American Method)—The nose projection was determined by projecting a tangent (simulated by wooden dowels), following the direction of the distal third of the nasal bones, downwards until it crossed another tangent projected anteriorly following the direction of the anterior nasal spine (3,4,25).

Identification Process

Photographs of the facial approximations were presented in a random order to 37 assessors who attempted to identify the target individual from a face pool. Assessors had a background in the medical sciences. Faces in the face pools were of similar age and same sex as the target individual. All facial approximations were presented in full face view. Three dimensional facial approximations were also presented to assessors in separate, single views of right three quarter (30° rotation from full face), right profile, and a combined view (all three together) in an attempt to determine if facial pose affected recognition.

Each face pool consisted of ten photographs of human faces, standardized for size. This resulted in some photographs differing in resolution. The use of antemortem photographs of target individuals limited the choice of these photographs. As a result, one target individual was wearing a hat ("Fred") and another target individual was wearing sunglasses ("Sam"). In these instances the corresponding face pool was photographed with similar attire, e.g., all had a hat or sunglasses. Nontarget individuals were randomly selected volunteers from the public. All photographs were developed and printed on black and white, standard (Ilford® IS3.1M) photographic paper (127 × 100 mm). Not all face pools included the target individual. Face pools that did not include the target individual had one randomly selected nontarget face, from the face pool, repeated in a slightly altered position and were developed on a slightly higher contrasting (Ilford® IS4.1M) photographic paper (127 × 100 mm). This kept the number of faces in the face pool consistent without introducing a new nontarget individual's face.

Three dimensional facial approximations and most nontarget individuals were photographed (without a flash) in a fluorescent-lit room. This simulated an average, indoor, amateur "snap shot," which the majority of the antemortem target individual photographs were expected to be. One face pool was used to compare facial approximations made on the same skull. Since assessors were unaware of the number of individuals approximated they were forced to assess each identification scenario as an independent case.

The 16 facial approximations with corresponding face pools were presented to assessors, one at a time, in a random order, in a single sitting. The assessors were asked to identify a face from the face pool as being the person who was approximated. Assessors had the option of not being able to make an identification. Identification responses were grouped under three classifications: (1) true positive identifications (identifications of facial approximations as their respective target individual); (2) false positive identifications (identifications of facial approximations as a face other than that of the target individual); or (3) no identification (the assessor could not recognize any face from the facial approximation). From this experimental design, subjects had a 50% chance of deciding if the target face was actually in the face pool and a 10% chance of correctly guessing the correct face when the target individual was present in the face pool. Therefore, in identification scenarios that included the face pool, the face of the target individual (and every other nontarget face in the face pool) was expected to have an

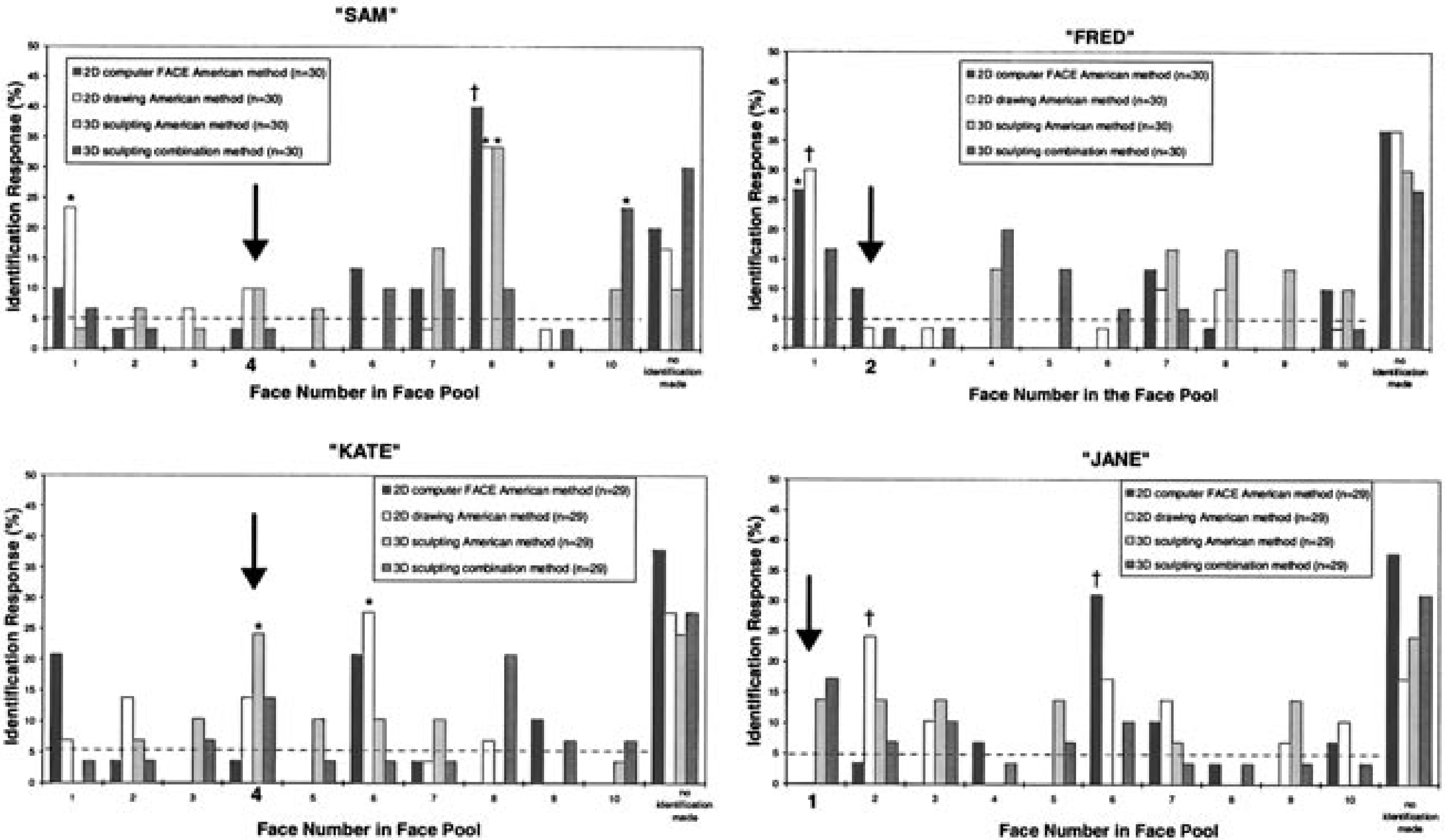


FIG. 2—*Identification responses for faces in the face pools when the target face is present. * indicates a significant difference above chance rate* ($p < 0.05$)*; † indicates a significant difference above chance rate* ($p < 0.05$) *and above the identification rate of the target individual* ($p < 0.05$)*; ↓ indicates the target face; ----- indicates 5% expected chance identification response for each face in the face pool.*

identification rate of 5% due to chance (50% × 10% = 5%). When the target individual's face was missing from the face pool subjects had a 50% chance of correctly deciding that no identification could be made or a 5% chance of identifying a nontarget face.

The JMP® (3.0.2) statistical package was used to test all results for statistical significance using Fisher's exact test.

Results

Of the 592 identification scenarios conduced, 472 included the target face in the face pool and 120 did not. Of the 472 identification scenarios including the target face, 38 true positive identifications were made, 316 false positive identifications were made, and 118 instances of no identification were reported. Of the 120 identification scenarios when the target face was not present in the face pool, 87 false positive identifications were made and there were 33 instances where no identifications were correctly made. Overall, 38 true positive identifications and 403 false positive identifications were made with 151 instances of no identification.

Identification responses for each skull, when target faces are present in the face pools, are shown in Fig. 2. Actual identification rates for each face in the face pools were compared to those expected by chance for statistical significance ($p < 0.05$). The 3D American facial approximation of "Kate" was the only facial approximation to give true positive identification rates above those expected by chance (Fig. 2).

TABLE 2—*Summary of the number of faces identified for each facial approximation when the target face was present in the face pool.*

	Sam	Fred	Kate	Jane
2D computer FACE assisted American method	6 (5)	5 (4)	6 (5)	6 (6)
2D drawing American method	7 (6)	7 (6)	6 (5)	6 (6)
3D sculpting American method	8 (7)	5 (5)	7 (6)	6 (5)
3D sculpting combination method	8 (7)	8 (7)	10 (9)	10 (9)

() indicates the number of nontarget faces identified.

When the target faces were present in the face pools, an average of five out of nine nontarget faces were identified for the 2D computer "FACE" assisted method (Table 2). This increased to six out of nine nontarget faces for both the 2D drawing American method and the 3D sculpting American method (Table 2). Eight out of nine nontarget faces were identified for the 3D sculpting combination technique (Table 2). All target faces were identified by at least one assessor except in three cases: 2D computer "FACE" assisted method of "Jane"; 2D drawing American method of "Jane"; and the 3D sculpturing American method of "Fred" (Table 2). Overall, ten nontarget faces were identified at statistically significant rates above chance (Fig. 2). Four of these nontarget faces were identified at a statistically significant rate ($p < 0.05$) that was also above that of their target individual (Fig. 2).

Comparison of the facial views to each other for the 3D American approximation of "Kate" failed to reveal any statistically significant differences (Fig. 3). This may be because the facial views did not differ in their ability to promote recognition, or (more likely) that the small sample sizes ($n = 8$ for the 3/4 facial view and $n = 7$ for each other facial view) were not adequate to detect statistical significance. While the power for detecting significance is low, there may be evidence that the three quarter view is best (Fig. 3). This aspect of the study will not be further discussed.

Discussion

According to the methods of Gail and Gart (31), it can be reliably concluded at a 0.05 confidence level for making a type I error (rejecting the null hypothesis when it is in fact true) and a probability of 0.20 (power of 0.80) for making a type II error (accepting the null hypothesis when it is in fact false), that the techniques of facial approximation studied do not give true positive identification rates at or above 35%, for the skulls approximated. We acknowledge that the power of this experiment was not high enough to reliably determine if identification rates of target individuals much lower than 35% were statistically significant in comparison to chance rates. In order to reach higher power much larger samples are needed. However an identification rate of 8%, for example, although statistically significant may also cause numerous nontarget

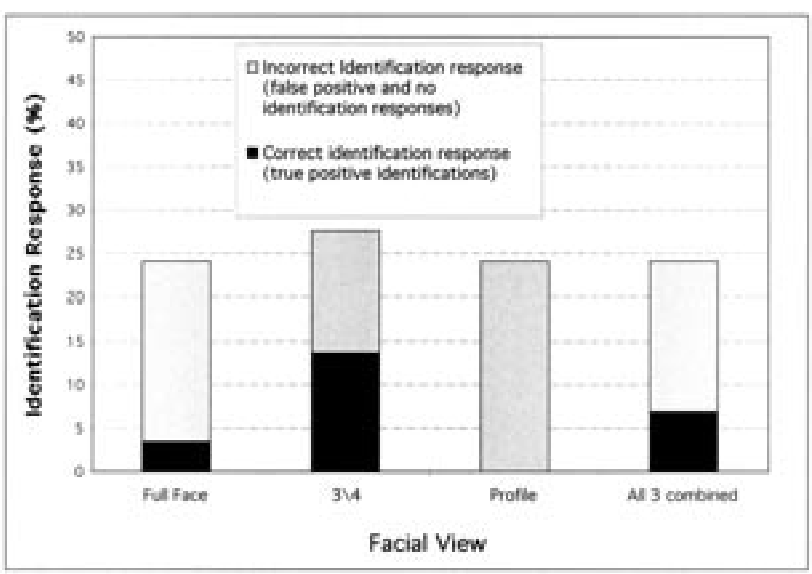


FIG. 3—*Identification responses for each facial view of facial approximations identified at a statistically significant rate above chance when the target face is present in the face pool. * indicates significance at 95% confidence level.*

individuals to be identified above chance rates. This may be problematic as it may not be possible or practical to investigate all of the leads generated.

Since only 1 out of 16 facial approximations were identified above chance at statistically significant levels, facial approximation appears to be an unreliable and inaccurate identification aid. This appears to contradict the results of published forensic cases and claimed success rates by people building facial approximations. Published forensic reports appear to be unrepresentative of the actual success of facial approximations. There have been 16 published successful cases (1,2,4,5,10,15,30,32–34), only one published article of limited to no success (35) and five published cases where no identification has been made (4,33,35). The situation may be that the few successful cases are given much attention, while many failures go unreported.

The many successful forensic reports and the high claimed success rates by people creating facial approximations, may be due to the inclusion of cases where additional information promoted identification, independent of the facial approximation. Information such as hair, jewelry, and clothes found at the crime scene may be responsible for successful identifications (35). If such information has aided in or been responsible for successful identification then the success should not be attributed to facial approximation alone. Realistic success rates of facial approximation techniques may be suggested by Ubelaker in a comment that he knows of only three cases, in his experience, where facial approximation alone resulted in positive identification (36). Successful target identification in other studies may be explained by chance or presence of additional information or better subjective interpretation of the skull and/or artistic modeling.

Although interpretation of the skull and approximation of its features may become somewhat standardized with the use of scientific guidelines, "approximationists" should have good artistic sense. Anyone can model a face, but to create a face that looks alive requires additional talent and manual dexterity. Although science has not yet determined if a finely modeled, life-like face is more easily recognized than a poorly modeled face, the former is preferred since recognition is expected to be promoted when a facial approximation appears more like the target individual. Figure 1 was provided in this study to illustrate the quality of some final approximations based on the methods used and artistic skills of CNS. It is difficult, however, from these pictures, to judge true artistic talent since facial approximation methods contribute to overall visage quality.

The usefulness of facial approximation may be questioned since the method is not highly effective but is time consuming and costly. However, since it potentially only takes one person to believe they recognize a facial approximation and report it to law enforcement agencies, for a tentative identification to be made, facial approximation may be seen as a useful identification technique. Thirteen of the sixteen facial approximations (81%) in this experiment could be expected to be successfully identified in a forensic environment since they were identified by at least one individual (Fig. 2).

The large number of false positive identifications (403 out of 592 identification scenarios) and the identification of many nontarget individuals (almost 70% of all nontarget individuals, Table 2) confirms the opinions of others (17,33,37), that facial approximations should not be used to positively identify an individual. Facial approximation should only be used when more precise methods of identification have failed and only to make, what Caldwell (26) calls, a "tentative" identification.

We disagree with Rhine (6) and Ubelaker (36) that the term facial reproduction is appropriate, since it implies a perfectly accurate replication of the face (9). We consider the most appropriate name to be "facial approximation" as suggested by George (9) since it indicates an inexact technique and does not imply that the method of building a face from a skull is accurate. We also agree with Rhine (6) that many of the other names used to describe facial approximation are inappropriate, e.g., facial reconstruction, facial restoration, and facial reconstitution, since they do not adequately describe the actual procedure.

Since the same techniques of facial approximation produced different identification results for different skulls (Fig. 2) it is difficult to determine which method is best. It may be that different methods are more effective for certain skulls. The present results indicate that only the 3D sculpting American technique produced true positive identifications at a statistically significant rate above chance and may, therefore, be more accurate in contrast to other methods studied. However, the 3D sculpting combination method was the only method, of those tested, to generate identifications of all target individuals and may, therefore, be considered superior despite its inaccuracies, e.g., target faces were not recognized above chance rates and many nontarget individuals (an average of eight out of nine) were identified (Table 2).

Facial approximations are largely unscientific as much of the method relies upon subjective interpretation of the skull itself. Also many guidelines used for determining facial features of facial approximations are artistic and their accuracy is unknown. Often a decision between two or more guidelines is needed when approximating facial features (Table 3) and without knowing the accuracy of each guideline, they cannot be intelligently selected to maximize the accuracy of facial approximation techniques. Consequently, the use of such guidelines introduces unknown quantities of error to fa-

TABLE 3—*Various generalizations used for facial feature determination.*

Feature	Generalization Used for Facial Approximation
Mouth width	1. Equal to the distance between the junction of the maxillary canine and the first premolar on each side (8,24,27).
	2. Equal to the distance between two perpendiculars dropped, one from the center of the pupil of each eye (24,26).
	3. Corresponds to the distance between the mandibular second molars (25).
Mouth closure line	1. Anywhere in the region of the upper central incisors (3,27).
	2. Lower third of central maxillary incisors for females or lower quarter for males (9).
	3. Equal to the line formed by the teeth when the mouth is closed (25).
Nose width	1. Equals the width of the nasal aperture +10 mm for whites and +16 mm for blacks (8).
	2. Equals the width of the nasal aperture +10 mm for whites and +15 mm for blacks (44).
	3. In Caucasoids the nasal aperture is approximately 3/5 of the total nose width (4,24,25).
	4. Is between the mid points of the canines or their alveoli (25).
Nose projection	1. Is 3× the length of the nasal spine (8,24).
	2. Projection of two lines, one at a tangent to the last third of the nasal bone the other as a continuation of the main direction of the nasal spine. The point of intersection gives the position of the tip of the nose and the nose projection (3,4,25).

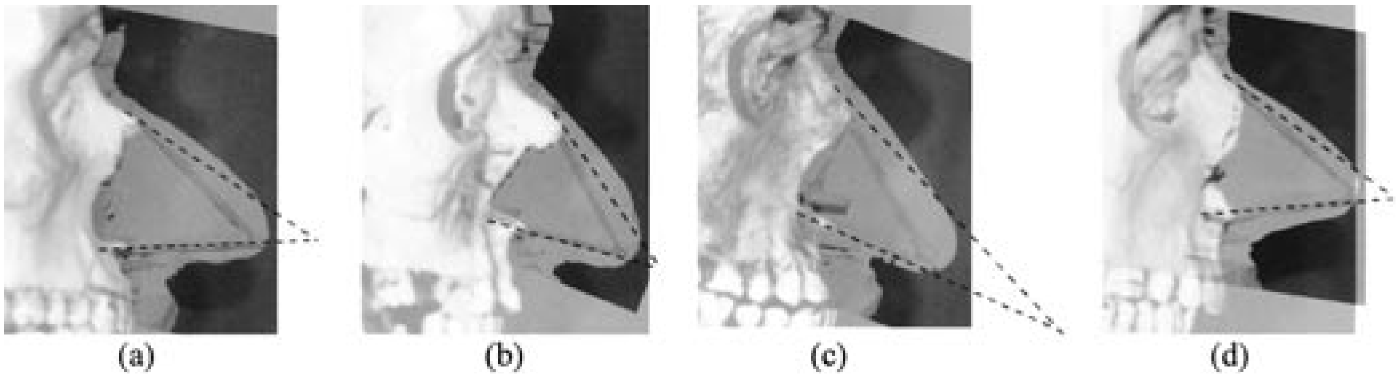


FIG. 4—*Superimposition of the 3D combination method facial approximations over their respective target individual's skull showing prediction of the nose tip. Note the unrealistic size of the nose and conservative use of the method.* (a) *"Sam";* (b) *"Fred";* (c) *"Kate";* (d) *"Jane"; ---- Shows projection of the tangents.*

cial approximations. One commonly used guideline that has been scientifically studied is that the height of the ear equals the height of the nose (8,24,25). This guideline is unreliable because in a sample of 103 young adults, it has been found that 95% had an ear bigger than their nose (38). Determining the position of the nose tip in the 3D combination technique also appears to be inaccurate as it consistently produced unrealistically large noses, even when the technique is conservatively followed (Fig. 4). While comprehensive knowledge of facial anatomy and extensive experience in facial approximation may reduce the amount of subjective interpretation possible, facial approximation must become based on scientific guidelines if subjectivity is to be limited and the amount of error introduced into a facial approximation is to be determined. Macho (39,40) has published some scientific guidelines for determining the nose as has Hoffman and colleagues (41), however, further scientific guidelines are needed for all facial features.

Face pool comparison studies are disadvantaged because most have used assessors who are unfamiliar with target individuals. Unfamiliar scenarios are not representative of a real forensic scenario, where people who know the victim usually recognize the facial approximation. Since it appears that familiar faces are easier to identify than unfamiliar faces (42,43), it may be argued that the unfamiliar testing scenario may reduce the true positive identification rates of facial approximations compared to those in a forensic scenario. While it would be more realistic to study identifications of facial approximations in a familiar scenario, it is difficult to do. People highly familiar with the victims are often relatives, who may become further traumatized if they participated in scientific study.

We did not approximate the hair of individuals in this study since it can not be determined from the skull alone. Helmer and colleagues (18) also report that the addition of a particular hair type and style to a facial approximation is subjective and that the addition of correct hair is due to chance (18). It is, therefore, surprising that the 24 facial approximations of Helmer et al. (18) have hair types and styles that are remarkably similar to their respective target individual. Helmer et al. (18) report that their facial approximations, without hair, bore limited resemblance to their respective target individual. Therefore, the result of Helmer et al. (18) experiment that independently made facial approximations are similar to the target individual, is probably an artifact of the hair they added to the facial approximations and consequently their conclusion that facial approximations are reproducible is probably incorrect. Information concerning hair type may be available if hair is found at the crime scene, however, the similarity of hair type and hair styles between many (but not all) published facial approximations and their target individuals may suggest that many forensic artists/scientists retouch their approximations to display the correct hair style after an identification has been made. This seems to be done to make the facial approximation appear more accurate.

It may be argued that in this experiment, the use of two target individuals, one wearing sunglasses and the other a hat, could have reduced the true positive identification rates. This is probably correct, however, the use of target individuals wearing such attire may also be representative of a forensic situation where a person making the identification may have seen the target individual in public, wearing a hat or sunglasses. It appears that hats had a limited effect on identification responses but sunglasses may increase identification of nontarget individuals since, overall, there were more statistically significant identifications of nontarget faces in "Sam's" face pool (Fig. 2). The effect of attire on target individual identification is not clearly apparent. Both "Sam" and "Fred" were identified more frequently for the 2D computer FACE American method and the 2D drawing American method in comparison to naked target faces. However, the 3D sculpting methods resulted in less frequent identification of "Sam" and "Fred" than the naked target faces.

Despite 3D facial approximation being a lengthy procedure (approximately 20 h), we agree with Ubelaker (36) that 3D facial approximations are probably superior to other methods since they can be made with prosthetic eyes and dressed up in real clothes making them appear extremely lifelike. So far, computer generated facial approximation methods have produced facial images that appear to be unacceptable for identification due to their lifelessness and limited detail. With further improvements, computer generated facial approximation may become the method of choice because it has the potential to limit subjective interpretation of the skull and reduces the time of the procedure.

Conclusions

The results of this study indicate that standard techniques of facial approximation rarely produce faces that are correctly identified at a statistically significant rate when compared to chance. Results also show that many incorrect responses are made (403 false positive identifications out of a total of 592 identification scenarios). Of the four methods studied, only the 3D sculpting American method of facial approximation gave identification rates of a target individual above chance at a statistically significant rate. From these results it is concluded that facial approximation is an inaccurate and unreliable technique. It is also concluded that facial approximation generates many tentative identifications, but is not useful to exclude suspected individuals to whom the skeletal remains may not belong.

Acknowledgments

We wish to thank Dr. Jane Taylor and Dr. Kenneth Brown at the Forensic Odontology Unit at the University of Adelaide for their invaluable technical assistance. Much thanks to Richard Neave for his untiring reply to emails and assistance with details of the facial approximation technique. Special thanks to Prof. John Clement, David Thomas, and Ronn Taylor for allowing us to use their computer facilities at the University of Melbourne. We are also indebted to Chris Leigh, Department of Anatomical Sciences, University of Adelaide, whose knowledge of and help with skull casting was deeply appreciated. We are most grateful to James Rogers and staff from the University of Adelaide's Dental School for their additional help.

References

1. Suzuki T. Reconstitution of a skull. Int Crim Police Rev 1973;264: 76–80.
2. Phillips V, Rosendorff S, Scholtz H. Identification of a suicide victim by facial reconstruction. J Forensic Sci 1996;14:34–8.
3. Gerasimov M. The face finder (Translated by Alan Houghton Brodrick). London: Hutchinson & Co, 1971.
4. Prag J, Neave R. Making faces. Using forensic and archaeological evidence. London: British Museum Press, 1997.
5. Farrar F. From skull to visage: a forensic technique for facial restoration. IPCH 1977;44:78–80.
6. Rhine J. Coming to terms with facial reproduction. J Forensic Sci 1990;35:960–3.
7. Tyrell J, Evison M, Chamberlain A, Green M. Forensic three-dimensional facial reconstruction: historical review and contemporary developments. J Forensic Sci 1997;42:653–61.
8. Gatliff B. Facial sculpture on the skull for identification. Am J Forensic Med Pathol 1984;5:327–32.
9. George R. The lateral craniographic method of facial reconstruction. J Forensic Sci, 1987;32:1305–30.
10. Angel J, Cherry D. Personality reconstruction from unidentified remains. FBI Law Enf Bul 1977;46:12–5.
11. Stewart T. Identification by the skeletal structures. In: Camps F, editor. Gradwohl's Legal Medicine 2nd Ed. Bristol, John Wright & Sons, 1968;123–54.
12. Vanezis P, Blowes R, Linney A, Tan A, Richards R, Neave R. Application of 3-D computer graphics for facial reconstruction and comparison with sculpting techniques. Foren Sci I 1989;42:69–84.
13. Quatrehomme G, Cotin S, Subsol G, Delingette H, Garidel G, Grevin G, Fidrich M, Bailet P, Ollier A. A fully three-dimensional method for facial reconstruction based on deformable models. J Forensic Sci 1997;42:649–52.
14. Bender F. In: Rubin S. Murder he sculpted. Read Dig 1998;May:62–7.
15. Gatliff B, Snow C. From skull to visage. J Biocommun 1979;6:27–30.
16. Van Rensburg J. Accuracy of recognition of 3-dimensional plastic reconstruction of faces from skulls (Abstract). Anatomical Society of Southern Africa 1993; 23rd Annual Congress: 20.
17. Snow C, Gatliff B, McWilliams K. Reconstruction of facial features from the skull: an evaluation of its usefulness in forensic anthropology. Am J Phys Anthropol 1970;33:221–8.
18. Helmer R, Rohricht S, Petersen D, Mohr F. Assessment of the reliability of facial reconstruction. In: İşcan M, Helmer R, editors. Forensic analysis of the skull: craniofacial analysis, reconstruction, and identification. New York: Wiley-Liss, 1993;229–46.
19. Krogman W. The reconstruction of the living head from the skull. FBI Law Enf Bul 1946;17:7–12.
20. Suk V. Fallacies of anthropological identification and reconstructions: a critique based on anatomical dissections. Publications of the Faculty of Science, University of Masaryk. 1935;207:1–18.
21. Montagu M. A study of man embracing error. Tech Rev 1947;49:345–7.
22. Brues A. Identification of skeletal remains. J Crim Law Criminol Police Sci 1985;48:551–66.
23. Helmer R. Schadelidentifizierung durch elekronicshe bildmischung: zugl e beitr zur konstitutionsbiometrie U Dickermessung d Gesichtsweichteile. Heidelberg: Kriminalistik-Verlag, 1984.
24. Krogman W, İşcan M. The human skeleton in forensic medicine. 2nd ed. Springfield, Charles C Thomas, 1986.
25. Fedosyutkin B, Nainys J. The relationship of skull morphology to facial features. In: İşcan M, Helmer R, editors. Forensic Analysis of the Skull. New York: Wiley-Liss, 1993;199–213.
26. Caldwell P. New questions (and some answers) on the facial reproduction techniques. In: Reichs K, editor. Forensic Osteology. Springfield, Charles C Thomas, 1986.
27. Angel J. Restoration of head and face for identification. Paper presented at the 30th Annual Meeting of the American Academy of Forensic Sciences. 20–25 February. St. Louis, MO, 1978.
28. Stewart T. The points of attachment of the palpebral ligaments: their use in facial reconstructions on the skull. J Forensic Sci 1983;28: 858–63.
29. Ubelaker D, O'Donnell G. Computer-assisted facial reproduction. J Forensic Sci 1992;37:155–62.
30. Perper J, Patterson G, Backner J. Face imaging reconstructive morphography: a new method for physiognomic reconstruction. Am J Forensic Med Pathol 1988;9:126–38.
31. Gail M, Gart J. The determination of sample sizes for use with the exact conditional test in 2×2 comparative trials. Biom 1973;29:441–8.
32. Rathbun, T. Personal identification: facial reproductions. In: Rathbun T, Buikstra J, editors. Human identification: case studies in forensic anthropology. Springfield: Charles C Thomas, 1984;347–62.
33. Ubelaker D. Facial reproduction. In: Wedit C, editor. Forensic Science, New York: Mathew Bender and Co. 1993;28E.01–28E.100.
34. Stoney M, Koelmeyer T. Facial reconstruction: a case report and review of development of techniques. Med Sci Law 1999;39:49–60.
35. Haglund W, Reay D. Use of facial approximation techniques in identification of Green River serial murder victims. Am J Forensic Med Pathol 1991;12:132–42.
36. Ubelaker D. Facial reconstruction. J Forensic Sci 1993;37:1442–4.
37. Taylor R, Angel C. Facial reconstruction and approximation. In: Clement J and Ranson D, editors. Craniofacial identification in forensic medicine, Great Britain, Arnold, 1998;177–85.
38. Farkas L, Munro I, Kolar J. The validity of neoclassical facial proportion canons. In: Farkas L, Munro I, editors. Anthropometric facial proportions in medicine. Illinois, Charles C Thomas, 1987;59.
39. Macho G. An apprasial of plastic reconstruction of the external nose. J Forensic Sci 1986;31:1391–1403.
40. Macho G. Descriptive morphological features of the nose—an assessment of their importance for plastic reconstruction. J Forensic Sci 1989;34:902–11.
41. Hoffman B, McConathy D, Coward M, Saddler L. Relationship between the piriform aperture and interalar nasal width in adult males. J Forensic Sci 1991;36:1152–61.
42. Bruce V. Changing faces: visual and nonvisual coding processes in face recognition. Brit J Psychol 1982;73:105–16.
43. Burton A, Wilson S, Cowan M, Bruce V. Face recognition in poor-quality video: evidence from security surveillance. Psychological Science 1999;10:243–8.
44. Schultz A. Relation of the external nose to the bony nose and nasal cartilages in whites and negroes. Am J Phys Anthropol 1918;1:329–38.

Additional information and reprint requests:
Carl N. Stephan
Department of Anatomical Sciences
University of Adelaide
Adelaide 5005, Australia
Email: carl.stephan@student.adelaide.edu.au