One View Is Not Enough: Engagement, Recognition and Presentation of Forensic Facial Depictions

Journal:	*Journal of Forensic Practice*
Manuscript ID	JFP-12-2025-0113.R1
Manuscript Type:	Research Paper
Keywords:	Forensic visualisation, Forensic Facial Depiction, Forensic Art, Face Recognition, Static versus Dynamic Presentation, Generalised Linear Mixed Models (GLMM)

MANUSCRIPT DETAILS

TITLE: One View Is Not Enough: Engagement, Recognition and Presentation of Forensic Facial Depictions

ABSTRACT:

To assess whether dynamic rotation enhances the recognition of forensic facial depiction compared to traditional multi-view static presentation, and to examine how viewpoint availability and repeat exposure influence identification outcomes.

Two psychological experiments compared the recognition of facial depictions presented either as (i) multi-view static triptychs/polyptychs or (ii) 16-second dynamic rotations (animated GIFs). Recognition was measured using correct and incorrect name responses from participants and a combined accuracy index. Generalised Linear Mixed Models were used to evaluate the effect of presentation format while accounting for variation across participants and stimuli.

In Experiment 1 (static triptych vs. dynamic), dynamic presentation did not significantly improve correct naming overall. However, when faces were viewed twice, dynamic presentation produced significantly higher correct naming and overall accuracy relative to static presentation. In Experiment 2 (static five-view polyptych vs. dynamic), static multi-view presentation now slightly increased correct naming and significantly improved accuracy compared with dynamic rotation. Repeated exposure provided modest identification benefits but was also associated with some increase in mistaken names. Overall, additional static viewpoints enhanced recognition, while dynamic motion was thought to mainly increase engagement and to provide recognition gains only under repeated viewing.

CUST_RESEARCH_LIMITATIONS/IMPLICATIONS__(LIMIT_100_WORDS) :No data available.

For the presentation of forensic facial depiction to stimulate recognition, results suggest that increasing viewpoint availability is more effective than motion for promoting accurate recognition, though dynamic presentation may be useful when repeated exposure is expected.

CUST_SOCIAL_IMPLICATIONS_(LIMIT_100_WORDS) :No data available.

This study offers the first direct empirical comparison of static multi-view and dynamic rotational facial depictions in a forensic context, demonstrating distinct recognition outcomes and informing evidence-based display practice.

One View Is Not Enough: Engagement, Recognition and Presentation of Forensic Facial Depictions

Abstract

Purpose

To assess whether dynamic rotation enhances the recognition of forensic facial depiction compared to traditional multi-view static (polyptych) presentation, and to examine how viewpoint availability and repeat exposure influence identification outcomes.

Design/methodology/approach

Two psychological experiments compared the recognition of facial depictions presented either as (i) multi-view static polyptychs—either (three-view) triptychs or (five-view) pentaptychs—or (ii) 16-second dynamic rotations (animated GIFs). Recognition was measured using correct and incorrect naming responses from participants and a combined naming-accuracy index. Generalised Linear Mixed Models were used to evaluate the effect of presentation format while accounting for variation across participants and stimuli.

Findings

Experiment 1 involved facial depictions presented as static triptych (three views: front view plus two side profiles) and dynamic-view rotation. Dynamic presentation did not significantly improve correct naming overall. However, when faces were viewed twice, dynamic presentation produced significantly higher correct naming and overall accuracy relative to static presentation. In Experiment 2, two three-quarter profiles were added to the static condition to give a static (five-view) pentaptych. These pentaptychs now slightly *increased* correct naming and significantly *improved* accuracy compared with dynamic rotation. Repeated exposure provided modest identification benefits but was also associated with some increase in mistaken names. Overall, static three-quarter viewpoints enhanced recognition, while dynamic motion was thought to mainly increase engagement and to provide recognition gains under repeated viewing when there is a limited number of fixed views available.

Originality

This study offers the first direct empirical comparison of static multi-view and dynamic rotational facial depictions in a forensic context, demonstrating distinct recognition outcomes and informing evidence-based display practice.

Practical Implications

For the presentation of forensic facial depiction to stimulate recognition, results suggest that increasing the number of viewpoints is more effective than motion for promoting accurate recognition, although dynamic presentation may be useful when repeating exposure of limited (front-and-side view) perspectives.

Keywords: *Forensic Facial Depiction; Forensic Art; Face Recognition; Static versus Dynamic Presentation; Generalised Linear Mixed Models (GLMM)*

Introduction

Forensic facial depiction from human remains sits at the intersection of face perception science, forensic art and medico-legal investigation. Its primary purpose is to elicit familiar face recognition of an unidentified decedent by those who knew the individual in life (Taylor, 2001; Wilkinson *et al*., 2024). The process involves analysing skeletal and soft tissue remains to construct a plausible, public-facing representation, with the success of the depiction ultimately measured not only by anatomical accuracy but by its ability to prompt recognition.

Considerable progress has been made in improving the morphological accuracy of facial reconstruction. While facial reconstruction is the prediction of face morphology from skeletal and soft-tissue data, facial depiction refers to the visual presentation of that form with applied textural detail (Smith, 2020). Advances in the field include greater anatomical understanding (Taylor, 2001; Wilkinson, 2004, 2010), use of population-specific tissue depth datasets (Lee *et al*., 2012; Stephan *et al*., 2014), increased training for practitioners (Won-Joon *et al*., 2016) and development of sophisticated digital modelling techniques (Miranda *et al*., 2018; Roughly & Wilkinson, 2019; Short *et al*., 2014). More generally, predictive methods for key features such as the nose, mouth and eye fissures have become increasingly standardised (Ullrich & Stephan, 2011; Wilkinson *et al*., 2006, 2024), and morphometric studies demonstrate that reconstructed facial shapes can approximate the majority of the true surface within a tolerance of approximately 2 mm (Lee *et al.*, 2012; Miranda *et al.*, 2018; Wilkinson *et al.,* 2006). Nevertheless, reliance on tissue depth data, whether from cadavers or living subjects, using radiographs, CBCT, CT, MRI or ultrasound carries limitations. These issues include post-mortem artefacts, restricted sample diversity, postural changes and the inherently static nature of the measurements (Claes *et al.,* 2012; Stephan *et al.*, 2014). Moreover, metrical accuracy does not automatically guarantee recognition, particularly if deviations affect identity-critical features (Lee *et al.*, 2015; Smith and Wilkinson, 2024; Wilkinson, 2015).

In contrast to these well-established efforts to improve morphological accuracy, the depiction presentation remains comparatively under-theorised and inconsistently practised. Practitioner decisions regarding viewpoint, finishing effects, lighting, colouration, contrast and textural inclusion directly influence whether a face is recognisable, yet choices such as these are often guided by tradition, client expectations or heuristic experience rather than empirical evidence (Davy-Jow, 2013; Smith, 2020). Despite being central to recognition potential, this stage has attracted little systematic research.

The science of face perception provides relevant insights in this area of forensic investigation. Exposure to multiple viewpoints allows observers to access identity-diagnostic shape information more effectively than a single, fixed view (Bruce, 1994; Burton *et al.*, 2016). However, not all viewpoints contribute equally to recognition. Performance tends to be poorer for profile views than for more intermediate perspectives, such as three-quarter (¾) views, which provide a richer combination of structural and depth information (Hill *et al.,* 1997; Liu & Chaudhuri, 2002). While the benefit of additional viewpoints may be relatively modest for familiar faces, it becomes more important when stimuli are less precise or contain error, as is often the case in forensic facial depictions (Burton et al., 2016; Jenkins & Burton, 2011). In such contexts, intermediate viewpoints may help support recognition by conveying more of the overall facial structure. Similarly, variation across lighting and image contexts help build robust mental representations of a face, enhancing recognition (Bruce *et al.,* 1998; Hill & Bruce, 1996; Johnston & Edmonds, 2009). Dynamic formats, such as animated facial rotations, may increase attention and engagement (Knight & Johnston, 1997; O'Toole *et al.*, 2002), but evidence that they improve recognition accuracy beyond multi-view static arrays remains limited (Lander *et al.*, 1999; Lander & Bruce, 2003). Further, the distinction between increased viewer engagement and improved recognition outcomes is, therefore, a critical but under-explored issue in forensic practice (Bailenson *et al.*, 2002; Lander *et al*., 2001). The present study attempts to address this gap in the literature by directly comparing recognition performance across multi-view static arrays and dynamic presentation.

Design

Two mixed-factorial experiments were conducted, each with independent participant groups to assess the effect of presentation format of facial reconstructions (forensic depictions) on face recognition. Recruitment was coordinated through Liverpool John Moores University (LJMU) gatekeepers, with additional publicity via physical poster advertisements around campus and online circulation (*Twitter/X*) to ensure a broad demographic sample and to enhance generalisability.

In Experiment 1, participants completed recognition of both static (three-view) triptychs (three side-by-side images: frontal, left profile, right profile) and dynamic rotation (16-second animated GIF). In Experiment 2, a different group of participants followed the same procedure, but static stimuli also included two potentially important three-quarter viewpoints to give a pentaptych (five side-by-side images: frontal, left and right profiles, and two three-quarter views). In both experiments, participants were presented with facial reconstructions of six identities for polyptychs (triptychs / pentaptychs), and six (different) identities as animations (see Stimuli). A within-subject AB-AB design was conducted that contrasted the effectiveness of static (A) and dynamic (B) presentation formats.

Participants were asked to provide the correct name or a description that could identify the person (i.e., a unique semantic identifier). Each response was also accompanied by a confidence rating, a procedure that attempted to discourage participants from withholding names due to uncertainty about accuracy and provide an index of internal validity, to distinguish genuine recognition from guesswork. The block order of presentation was randomised (triptychs / pentaptych then animation, or the reverse) with equal sampling, with stimulus set randomised and counterbalanced across participants. Participants were given a different random order for the first presentation of these stimuli (with this random order maintained for the second presentation).

With recognition performance for celebrity identities dependent on pre-existing exposure, a ground-truth familiarity check was used to ensure that participant responses to the presented reconstructed faces sufficiently reflected genuinely familiar faces. The approach follows standard forensic practice: the primary interest lies in recognition performance assessed by Generalised Linear Mixed Models (GLMM) that are conditional (as in real life) based on prior familiarity with the relevant identities (i.e., for faces that have been seen previously in daily life), and where excluding unfamiliar identities improves interpretability (e.g., Davidson et al, 2025; Erickson et al., 2022; Frowd et al., 2025; Portch et al., 2025). Thus, as a final stage in the experimental process, participants viewed a reference photograph of each identity and were asked (using the same instruction as above for the reconstructions) to name them.

Two planned adjustments were conducted based on these additional (ground-truth) data. First, to allow good resolution of responses per person, each participant was required to correctly identify at least six (i.e., 50%) of the identities in this familiarity check for their data to be considered for analysis. As a result of this *a priori* rule, two additional participants were recruited to give the samples described in the follow section. Second, only responses to reconstructed stimuli (triptychs, pentaptychs and animated faces) with which participants were familiar (from the ground-truth familiarity check) were retained for inferential analyses.

Participants

The number of participants and stimulus items were based on considerable research involving recognition (naming) of error prone facial stimuli (e.g., Davidson et al., 2025; Erickson et al., 2022; Joyce & Frowd, 2025; Portch et al., 2025). This body of research demonstrates that a minimum of 10 stimulus items and 10 participants per condition are necessary to be able to detect a forensically-important medium effect [*Exp(B)* ≈ 2.5] with good power using the planned GLMM method of analysis, estimates that are supported by computer simulation (Davidson et al., 2025). We exceeded these minimal requirements in each experiment; the resulting design was indeed appropriate (e.g., see *Exp(B)* values Tables III, IV and X)—in fact, it was able to detect a small effect (e.g., see Table XII).

In each experiment, we recruited 48 unique participants whose data was used for GLMM analysis. The mean age of participants in Experiment 1 was 31.3 years (range: 19–64; 27 females, 21 males), and in Experiment 2 was 37.0 years (range:

20–70; 23 females, 25 males). Participants received a £5 Amazon voucher for their participation.

Stimuli

Stimuli comprised 12 well-known famous faces, half of each apparent sex, all of White European population affinity, and estimated to be between 25 and 45 years of age at the time of casting. The stimulus set was generated through a four-stage process: (1) 3D scanning of plaster casts ('life-masks'), (2) refinement of the digitised meshes, (3) staging head assets to produce high-fidelity 3D renderings and (4) creation of animated GIFs and static images for presentation.

The life-masks were from a shared and purchased resource between LJMU Face Lab and the Centre of Anatomy and Human Identification (CAHID) at the University of Dundee. Scanning was carried out with an Artec Space Spider scanner, and the resulting meshes were processed in Artec Studio 15 Professional (v.15.1.2.60, 64-bit) on a dedicated workstation (Intel® Core™ i7-7700 CPU @ 3.60 GHz, 32 GB RAM, NVIDIA GeForce GTX 1060, Windows 10 Enterprise 64-bit). Assets were exported as .obj files and refined using 3DSystems Geomagic® Freeform® (v.2021.1.25) with a 3D Systems Touch Haptic Arm (HID GEO-THA-35499), alongside Pixologic ZBrush (v.2021.7.1), running on a high-spec workstation (Intel® Core™ i9-10900X CPU @ 3.70 GHz, 128 GB RAM, Windows 10 Enterprise 64-bit). Final staging, rigging and rendering were conducted in Autodesk Maya (v.2023) with outputs composed into dynamic rotations in Adobe CC Premiere Pro (2021).

Refinement of the meshes was required because the masks varied in completeness and condition, reflecting differences in casting practices, environments and methodologies. For example, some meshes extended only to the zygomatics (cheeks), while others captured the full face, hairline and ears. Although all were cast in a neutral expression, subtle variations were evident; for example, there were differences in mouth occlusion, and distortions from supination and skin pulling during plastering. Signs of wear such as chipped nasal tips and cracking were also present. This variability does represent a limitation, but it does not greatly detract from the study: rather, it mirrors the kinds of inevitable inclusion of error present when depicting the faces of unknown decedents for recognition purposes.

These 12-item stimuli were divided into two equally sized sets. One set comprised half of the identities, randomly selected, for the polyptychs (triptychs or pentaptychs), with the remaining half used for the animations; the other set comprising the reverse identities (i.e., the identities used for animations were now used for triptychs / pentaptychs, and vice versa). Participants were presented with one of these two sets, randomly assigned, with equal sampling for each experiment.

Procedure

Participants were tested individually, and the task was self-paced. Each trial began with the presentation of a face stimulus, which remained onscreen until the participant responded. Participants were asked to identify the individual depicted, either by giving the correct name or a description that could be used to identify the person. Participants were asked to give a name if one came to mind, thereby limiting effects that might otherwise be caused by individual differences in response criterion. Participants were also asked for a confidence rating for each name given (1 = "extremely unconfident" to 7 = "extremely confident"). Participant responses were recorded by the researcher.

All 12 identities were presented over both conditions (six identities for triptychs / pentaptychs and the other six for animations), with participants randomly assigned with equal sampling to one of two stimuli sets, with block order of presentation (triptychs / pentaptychs shown first or second) randomised and counterbalanced across participants. Presentation of static reconstructions varied by experiment, with triptychs used in Experiment 1 and pentaptychs in Experiment 2. Each participant received a different random order of identities for the first presentation, with this order maintained for the second presentation. The task was completed in approximately 15 minutes, including debriefing as to the aims of the experiment.

Data Coding and Scoring

Responses were categorised into three mutually exclusive outcomes. A Correct Name Rate (CNR) was modelled for responses that were either the correct identity or a unique semantic description (e.g., by saying "Harry Potter" to the stimuli for Daniel Radcliffe). An Incorrect Name Rate (ICNR) captured cases where an incorrect name (i.e., a wrong identity) was given. A No Name Rate (NNR) represented cases where participants reported that no name came to mind.

Correct and incorrect naming responses were coded separately, as these outcomes represent qualitatively different types of error, each with consequences in forensic settings. For the former, a value of 1 was assigned for a correct response, and for the latter this value was given for an incorrect (mistaken) response (i.e., the name or description for another person). To provide an integrated measure of performance, an accuracy index was also calculated. This index was coded as +1 for an accurate (correct) name and –1 for inaccurate (mistaken) name, and was defined as the number of correct minus mistaken names relative to total (correct and mistaken) names. This metric allows differentiation between genuine recognition and systematic misidentification.

Statistical Analysis

Analyses were conducted using IBM SPSS Statistics for Windows (Version 29.0.2.0) using Generalized Linear Mixed Effects Models (GLMM). This regression-based approach is ideally suited to analysing individual data responses and provides a unified model that incorporates both fixed and random effects (Bolker *et al*., 2009)—see Davidson *et al.* (2025) and Erickson *et al.* (2022) for examples. The random effects were specified using a *maximal* structure, meaning that the model initially included all sources of random variation justified by the design (Barr et al., 2013). The random effects included random intercepts, to account for individual differences between participants and items, and random slopes, to account for the variation associated with the repeated presentation of items.

As mentioned above, three dependent variables were modelled: (i) correct naming, (ii) incorrect naming and (iii) naming accuracy. Correct and incorrect naming were analysed using binomial logistic regression, while accuracy was modelled using a multinomial structure.

Model convergence criteria were set to SPSS defaults (absolute difference of 1×10^{-8}, 100 maximum iterations). Both model-based and robust estimation were tested; as they produced identical results, the model-based approach was used as this method is relatively more common among statistical packages, facilitating replication of results (Erikson *et al.,* 2022). For selection of independent variables (IVs), up to three confirmatory candidate models were compared in the following order: (i) full factorial, containing individual predictors and their interaction, (ii) a model that included all main effects and (iii) a single predictor model. IVs and interaction terms were retained in the model where their p-value were equal to or less than the standard default for maintaining variables in a regression model ($\alpha < .10$; Field, 2024).

Effect sizes were reported as exponentiated Beta (*B*) regression coefficients [*Exp(B)*]. These effects were interpreted using sensible, approximate thresholds defined as small (1.5), medium (2.5) and large (4.5) effects (e.g., Davidson et al., 2025; Portch et al., 2025; Sporer & Martschuk, 2014). Also, to facilitate ease of comparison, effect sizes were presented with values greater than unity (Osbourne, 2017): negative values of *B* were made positive (hence the absolute value of *B*, |*B*|, was taken in these cases) and the associated confidence intervals were adjusted accordingly.

Results

The results are reported separately for each experiment, with descriptive statistics presented first, followed by inferential analyses of correct naming (CNR), incorrect naming (ICNR) and accuracy.

Experiment 1

Experiment 1 examined whether recognition performance differed between dynamic presentations (animated 16-second rotations) and static triptychs (frontal, left and right profiles), across two exposures. Figure 1 provides an example stimulus, and Table I reports descriptive means for correct, incorrect and accuracy naming by presentation modes and exposure orders.

Figure 1- Example stimulus. Angelina Jolie in display condition 1, static presentation of frontal view, two profile views (left) and display condition 2, screenshot of the dynamic gif presentation (right). The full animated stimulus is available in Supplementary Material 1.

Table I - Experiment 1 Correct, Incorrect and Accuracy Naming Means.

Naming Type	Order	Mode		Mean
		Static	Dynamic	
Correct	First	53.8 (135 / 251)	55.7 [b] (141 / 253)	54.8 [a] (276 / 504)
	Second	55.0 [c] (138 / 251)	71.1 [b, c] (180 / 253)	63.1 [a] (318 / 504)
	Mean	54.4 (273 / 502)	63.4 (321 / 506)	58.9 (594 / 1008)
Incorrect	First	34.5 (40 / 116)	39.3 (44 / 112)	36.8 [d] (84 / 228)
	Second	47.8 (54 / 113)	43.8 (32 / 73)	46.2 [d] (86 / 186)
	Mean	36.8 (84 / 228)	41.1 (76 / 185)	41.1 (170 / 414)
Accuracy	First	37.9 (95 / 251)	38.3 (97/ 253)	38.1 (192 / 504)
	Second	33.5 [e] (84 / 251)	58.5 [e] (148 / 253)	46.0 (232 / 504)
	Mean	35.7 (179 / 502)	48.4 (245 / 506)	42.1 (424 / 1008)

Note. For Correct Naming, figures are percentage correct; numbers in parentheses are total number of correct responses against total number of responses. For Incorrect Naming, figures are percentage incorrect; numbers in parentheses are total number of incorrect responses divided by total number of mistaken and no-naming responses. For Accuracy, figures are percentage; numbers in parentheses are number of correct minus mistaken responses divided by total number of responses. [a] $p = .006$, [b] $p < .001$, [c] $p < .001$, [d] $p = .018$, [e] $p < .001$.

The results suggest that correct naming improves somewhat from static to dynamic modalities (*MD* = 9.0%), and that there is some benefit (*MD* = 8.3%) for a repeated presentation. Incorrect naming also increased slightly for dynamic than static modalities (*MD* = 4.3%), and there was again a similar-size increase (*MD* = 9.4%) for a repeated presentation. In terms of accuracy (correct naming - incorrect naming), all means were positive, meaning that correct responses were overall more pervasive. The degree of accuracy was somewhat greater (*MD* = 12.7%) for dynamic than static modality, indicating an overall benefit for the former. There was again an overall positive benefit (*MD* = 7.9%) by repeating the presentation.

Correct Naming

A full factorial GLMM for correct naming was conducted between mode (coding as 1 = display condition 1, 2 = display condition 2 and presentation (coded as 1 = first presentation, 2 = second presentation). The interaction between mode and presentation [$F(1, 1004) = 7.50$, $p = .006$, $Exp(|B|) = 2.74$] was less than the planned value for alpha (.05) and so this term was retained, producing the final model (Table II).

Table II - Experiment 1 Correct Naming Final Model.

Fixed Effects	*F*	*DF1*	*DF2*	*p*
Presentation	7.52	1	1004	.006
Mode	11.39	1	1004	< .001
Presentation × Mode	7.50	1	1004	.006

A simple main effects analyses for the interaction (Table III) revealed that, although dynamic and static modes did not differ for the first presentation, there was benefit for dynamic over static for the second presentation. Also, while there was no difference between first and second presentation in the static condition, there was improvement in the dynamic condition. These results indicate that repetition improved naming performance in the dynamic condition but not in the static condition.

Table III - Experiment 1 Correct Naming Coefficients for Presentation × Mode

Fixed Effects	*B*	*SE(B)*	*p*	*Exp(B)*	*95% CI (– / +)*
Dynamic vs Static (First presentation)	0.10	0.21	.63	1.11	0.73 – 1.67
Dynamic vs Static (Second presentation)	0.93	0.25	< *.001*	2.53	1.45 – 4.07
Second vs First (Static mode)	0.08	0.23	.73	1.08	0.69 – 1.70
Second vs First (Dynamic mode)	0.91	0.26	< *.001*	2.48	1.43 – 4.02

Note. Overall Correct Classification: 76.9%. The model was specified with the lowest category of categorical predictors as reference (Static, First), and predictors and target were sorted in a descending order. (For the final model:) Information criteria are based on the -2 log likelihood (*AICC* = 4732.13, *BIC* = 4741.94); random effects were random intercepts for participants [σ = 0.73, *SE(σ)* = 0.22] and items [σ = 1.48, *SE(σ)* = 0.68], and random slopes for Presentation for items [σ = 0.05, *SE(σ)* = 0.08]. The intercept was [*B* = 0.13, *SE(B)* = 0.41].

Incorrect Naming

A full factorial GLMM for incorrect naming revealed that the interaction between mode and presentation was greater than alpha (*p* = .368, *Exp(|B|)* = 2.04) and was removed. In the resulting model, mode was also greater than alpha (*p* = .562, *Exp(|B|)* = 1.36) and was removed as well. In the final model, presentation was retained (Table IV).

Table IV - Experiment 1 Incorrect Naming Final Model.

Fixed Effects	*F*	*DF1*	*DF2*	*p*
Presentation	5.64	1	412	.018

Table V - Experiment 1 Coefficients for Presentation for Incorrect Naming.

Fixed Effects	*B*	*SE(B)*	*t(412)*	*p*	*Exp(B)*	*95% CI(-/+)*
Second - First	0.51	0.22	2.37	.018	1.67	1.09 – 2.55

Note. Overall Correct Classification: 73.4%. The model was specified with the lowest category of categorical predictors as reference (First), and predictors and target were sorted in a descending order. Information criteria are based on the -2 log likelihood (*AICC* = 1826.50, *BIC* = 1834.52); random intercepts for participants [σ = 0.82, *SE(σ)* = 0.32] and items [σ = 0.13, *SE(σ)* = 0.12]. Intercept [*B* = -0.57, *SE(B)* = 0.23].

This analysis (Table V) shows that incorrect names occurred more often for the second compared to the first presentation of the face stimuli. Also, as mode was not retained in the model, modality of presentation (static or dynamic) had no impact on ICNR.

Naming Accuracy

A full factorial GLMM for naming accuracy was conducted with mode (coded as 1 = display condition 1, 2 = display condition 2) and presentation (coded as 1 = first presentation, 2 = second presentation) entered as fixed factors. The interaction between mode and presentation was less than alpha ($F(1, 1003) = 9.13$, $p = .003$, $Exp(B) = 2.06$), and was therefore retained. The overall results of this (final) model are shown in Table VI.

Table VI - Experiment 1 Naming Accuracy Final Model

Fixed Effects	*F*	*DF1*	*DF2*	*p*
Mode	10.05	1	1003	.002
Presentation	4.55	1	1003	.033
Presentation × Mode	9.13	1	1003	.003

Table VII - Naming Accuracy – Coefficients for Comparisons for Presentation × Mode

Fixed Effects	*B*	*SE(B)*	*p*	*Exp(B)*	*95% CI (– / +)*
Dynamic vs Static (First presentation)	0.23	0.19	.90	1.02	0.71 – 1.47
Dynamic vs Static (Second presentation)	0.95	0.21	*< .001*	2.59	1.69 – 3.97
Second vs First (Static mode)	0.18	0.20	.60	1.20	0.75 – 1.61
Second vs First (Dynamic mode)	0.90	0.22	*< .001*	2.46	1.50 – 4.02

Note. Pairwise comparisons for the Presentation × Mode interaction in the Naming Accuracy GLMM (Experiment 1). Significant effects show that accuracy improved across presentations in the dynamic condition but not in the static condition. Overall Correct Classification: 65.2%. The model was specified with the lowest category of categorical predictors as reference (First, Dynamic), and predictors and target were sorted in a descending order. Information criteria are based on the -2 log likelihood ($AICC$ = 7832.67, BIC = 7842.48); random intercepts for participants [σ = 0.38, $SE(\sigma)$ = 0.13] and items [σ = 1.06, $SE(\sigma)$ = 0.49], and random slopes for Presentation for items [σ = 0.01, $SE(\sigma)$ = 0.05]. Intercept [B = -1.81, $SE(B)$ = 0.38].

A simple-main effects analysis reveals the same pattern of significant and non-significant differences as for correct naming: only the second (cf. first) presentation led to a benefit for dynamic over static mode, and that there was a benefit for dynamic (cf. static) for second presentations.

Experiment 2

Experiment 2 added two three-quarter views to the triptych, resulting in five viewpoints (frontal, left/right profiles, and left/right three-quarter views). This static condition was again contrasted with the dynamic presentation (16-second rotation). Figure 2 provides an example stimulus and Table VIII summarises the descriptive means for correct, incorrect, and accuracy naming across presentation modes and exposure orders.

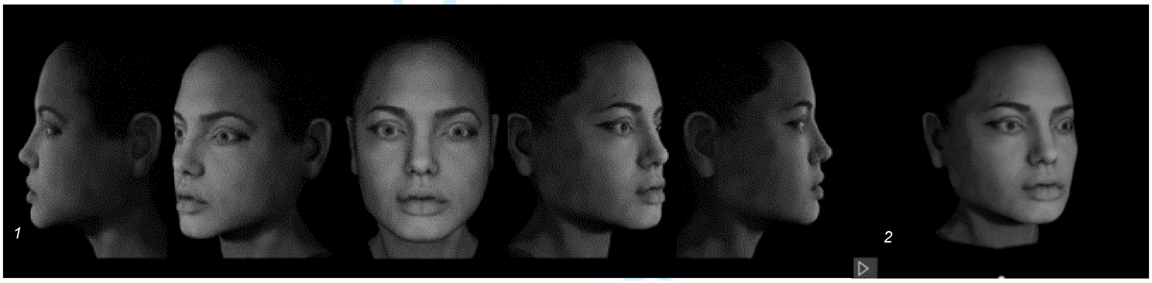

Figure 2 - Example stimulus. Angelina Jolie in display condition 1, static presentation of frontal view, two profile and two ¾ views (left) and display condition 2, screen shot of the dynamic gif presentation (right).

Table VIII – Experiment 2 Correct, Incorrect and Accuracy of Naming Means.

Naming Type	*Order*	*Mode*		Mean
Correct		Static	Dynamic	
	First	51.1 (114 / 223)	41.1 (85 / 207)	46.3[a] (199 / 430)
	Second	54.3 (121 / 223)	46.9 (97 / 27)	50.7[a] (218 / 430)
	Mean	52.7[b] (235 / 446)	44.0[b] (182 / 414)	48.5 (417 / 860)
Incorrect	First	11.2 (25 / 223)	13.5 (28 / 207)	12.3 (53 / 430)
	Second	10.8 (24 / 223)	15.5 (32 / 207)	13.0 (56 / 430)
	Mean	12.3 (53 / 430)	14.5 (60 / 414)	12.7 (109 / 860)
Accuracy	First	39.9 (89 / 223)	27.5 (57 / 207)	34.0 (146 / 430)
	Second	43.5 (97 / 223)	31.4 (65 / 207)	37.7 (162 / 430)
	Mean	41.7[c] (186 / 446)	29.5[c] (122 / 414)	35.8 (308 / 860)

Note. See Table I, *Note* for definition of figures. [a] p = .044, [b] p = .052, [c] p = .013.

In contrast to Experiment 1, the pattern of results was clearly different, particularly for the static condition. Here, correct naming showed a slight overall decline from static to dynamic modalities (*MD* = 8.7%), although a small positive benefit remained (*MD* = 4.4%) for a repeated presentation. Incorrect naming slightly increased from static to dynamic (*MD* = 2.2%), but there was almost no difference (*MD* = 0.7%) for a repeated presentation. In terms of accuracy (correct naming - incorrect naming), all means were positive. However, degree of accuracy was now somewhat greater for static than dynamic (*MD* = 12.2%), indicating that more correct names in general were produced in the *static* condition. There was a small overall positive difference (*MD* = 3.7%) by repeating presentation.

Correct Naming

A full factorial GLMM for correct naming was run between mode (coded as 1 = display condition 1, 2 = display condition 2) and presentation (coded as 1 = first presentation, 2 = second presentation).

The interaction between mode and presentation (p = .672) was greater than alpha and so was removed. The model was re-run with the main effects only. In this model, presentation was not significant (p = .165) and was removed. The final model retained mode, producing the final model (Tables IX and X). The effect of mode was significant, reflecting a small, between-subjects difference in correct naming between the two display conditions, with static presentation producing higher correct naming than dynamic presentation.

Table IX - Experiment 2 Correct Naming Final Model.

Fixed Effects	*F*	*DF1*	*DF2*	*p*
Mode	3.80	1	857	.018

Table X - Experiment 2 Coefficients for Mode for Correct Naming.

Fixed Effects	*B*	*SE(B)*	*t(857)*	*p*	*Exp(|B|)*	*95% CI(-/+)*
Static - Dynamic	-0.35	0.18	-1.95	.05	1.41	1.06 – 1.88

Note. Overall Correct Classification: 64.0%. The model was specified with the lowest category of categorical predictors as reference (Dynamic, Second), and predictors and target were sorted in a descending order. Information criteria are based on the -2 log likelihood (AICC = 3763.01, BIC = 3767.76); random intercepts for items [σ = 0.56, SE(σ) = 0.27]. Intercept [*B* = -0.05, *SE(B)* = 0.27].

Incorrect Naming

A full factorial GLMM for incorrect naming revealed that the interaction between mode and presentation was greater than alpha (p = .673, *Exp(|B|)* = 1.23) and so was removed. In the resulting model, mode was also greater than alpha (p = .180, *Exp(|B|)* = 1.38), as was presentation (p = .789, *Exp(|B|)* = 1.07), and so neither static nor dynamic display conditions influenced incorrect naming.

Naming Accuracy

A full factorial GLMM for naming accuracy showed that the interaction between mode and presentation was greater than alpha (p = .929, *Exp(|B|)* = 1.02), and so was removed. In the resulting model, presentation was also removed (p = .392, *Exp(|B|)* = 1.12), producing the final model (Tables XI and XII).

Table XI - Experiment 2 Naming Accuracy Final Model.

Fixed Effects	*F*	*DF1*	*DF2*	*p*
Mode	6.18	1	857	.013

Table XII - Experiment 2 Coefficients for Mode for Naming Accuracy.

Fixed Effects	B	SE(B)	t(857)	p	Exp(|B|)	95% CI (-/+)
Dynamic - Static	-0.33	0.13	2.49	.013	1.40	1.07 – 1.81

Note. Overall Correct Classification: 60.1%. The model was specified with the lowest category of categorical predictors as reference (Static), and predictors and target were sorted in a descending order. Information criteria are based on the -2 log likelihood (AICC = 6560.80, BIC = 6565.55); random intercepts for items [σ = 0.30, $SE(\sigma)$ = 0.15]. Intercept [B = 0.90, $SE(B)$ = 0.19].

This analysis reveals that the static (cf. dynamic) modality of presentation used in this experiment improved accuracy of naming with a small effect size.

Results Summary

Across both experiments, presentation format played a critical role in recognition performance, while effects of repetition depended on the experimental context. Dynamic rotation did not outperform the static triptych on first exposure, but when faces were encountered for a second time, dynamic presentation yielded a recognition advantage. However, this benefit was accompanied by an increased risk of misidentification, highlighting potential limitations for applied use.

Conversely, when the static condition provided richer (more varied) visual information (five viewpoints that included two three-quarter views of the face), this multi-view presentation outperformed dynamic rotation on overall accuracy and showed a marginal advantage in correct naming. Importantly, this enhanced performance did not increase the likelihood of misidentification (unlike Experiment 1). Taken together, the results suggest that while dynamic presentation can support recognition for repeated exposure, static multi-view presentations offer a more robust and reliable advantage under conditions of higher visual information. These complementary patterns underscore the importance of considering both stimulus richness and the opportunity for repeated viewing when evaluating face recognition performance.

Discussion

The findings highlight the nuanced relationship between the format of facial depiction and recognition. When faces were presented in three static views, dynamic presentation did not improve recognition on first exposure but conferred an advantage after repeated presentation. This suggests that motion enhances consolidation of identity rather than initial recognition. By contrast, when the static condition offered five distinct views, static images outperformed dynamic display on both overall accuracy and (marginally) on correct naming, without increasing the likelihood of misidentification. This advantage likely reflects the inclusion of intermediate (¾) viewpoints, which provide a richer combination of structural and depth information (e.g., to improve perception of the nose and brow ridge areas) than profile views alone, allowing observers to better integrate identity-diagnostic features across perspectives. This is particularly important for unfamiliar or error-prone stimuli, where additional structural information may help compensate for missing or ambiguous cues. Multiple static perspectives can therefore rival or exceed the informational richness of a continuous rotation.

This pattern aligns with theories of face recognition that emphasise the integration of structural information across different perspectives (Bruce & Young, 1986; O'Toole *et al*., 2002). Motion appears to provide an opportunity for deeper engagement, yet this does not always translate into improved recognition accuracy when sufficiently rich static information is available. A key implication is that dynamic display enhances engagement rather than recognition itself. Motion captures and sustains attention, encouraging viewers to invest more effort in processing a face or in offering a potential identity. However, recognisability ultimately depends on the quality of the cues and how effectively they align with stored memory representations. This distinction is increasingly relevant in the digital age where richly textured, animated faces are the norm and viewers are motivated to engage with them. The findings therefore raise an applied question: can the heightened engagement elicited by dynamic displays be harnessed to improve recognition when uncertainty is unavoidable?

Forensic facial depictions are not portraits, but rather approximations of decedent appearance, and with this, they inevitably contain error. Dynamic or interactive formats may encourage sustained engagement with a face, potentially increasing the likelihood of successful identification despite missing or erroneous data.

Current investigative systems and display platforms are built primarily around static images. Integrating dynamic or interactive formats would require not only technical adaptation, but also procedural, legal and ethical considerations. Enhancing engagement may create optimum conditions for recognition, but the ethics of representing faces in such vivid and extended ways must also be acknowledged. Issues of consent, preserving the dignity of the decedent, appropriateness, sensitivity to families and the risk of overstating the certainty of a depiction all weigh on the use of dynamic techniques.

At present, practitioners typically manage uncertainty in facial depictions by obscuring unknown details, for example through blurring, cropping or shading absent regions. This is most often applied to external features such as the hairline, hairstyle or ears, aspects of reconstruction that cannot be predicted with precision from skeletal remains. From a face perception perspective, this practice interacts with established findings that recognition for familiar faces is more tolerant and guided by internal features (eyes, nose, mouth), whereas recognition of unfamiliar faces relies more heavily on external features (Clutterbuck & Johnston, 2002; Ellis *et al.*, 1979). Cropping or blurring therefore reduces the risk of presenting inaccurate information, but the cost may be perceptual completeness and interference to recognition. Raymond (2022), for instance, demonstrated that occluding facial regions impairs recognition performance. However, techniques such as these have been found to be valuable to recognition by concealing information that is inaccurate in the face (e.g., Frowd, 2021; Frowd et al., 2014).

Part of the challenge for depiction of facial reconstructions arises because faces are inherently variable. A face changes across the lifespan in response to BMI, ageing, environmental influences such as light exposure, trauma, or cultural practices like body modification (Coleman & Grover, 2006; Meinhardt-Injac & Hildebrandt, 2016). Factors such as these alter facial texture in ways that do not seem to be recoverable from skeletal remains, yet such textural cues are important for recognition (e.g., Johnston & Edmonds, 2009; Liu *et al.*, 2005; Troje & Bülthoff, 1995). Establishing how much of this texture should be added to a reconstruction is therefore central to maximising recognition.

This challenge is directly relevant to Valentine's (1991) multidimensional face space (MDFS) framework, in which each identity is represented as a cluster of exemplars in a multidimensional perceptual space, defined by attributes such as feature proportions, lighting or viewpoint. More distinctive faces are located further from the central "average" face and are therefore easier to recognise (Valentine, 1991, 2001)[1]. Familiarity strengthens recognition because repeated and varied exposures enrich the subspace for that identity, enabling observers to match novel appearances against a more flexible and distributed representation (Blank & Yovel, 2011; Hole & Bourne, 2010). Forensic depictions, by contrast, often obscure precisely those variable external or textural cues that would help broaden such subspaces, reducing the likelihood of a successful match.

The development of more advanced dynamic tools raises the possibility of an alternative approach, and rather than concealing uncertainty, depictions could incorporate variance to represent it. A depiction might, for example, present alternative plausible textures, hairstyles or surface details in a morphing-type rotation, enabling viewers to engage with the variability inherent in the depiction. This strategy of presenting variance, rather than obscuring unknown information, may also make depictions more inclusive across cultural contexts where reliance on different facial features varies (Latif & Moulson, 2022; Megreya & Bindemann, 2009; Suhkre *et al.*, 2015). It may further help to mitigate the influence of practitioner bias during construction (Lee & Wilkinson, 2016). By adopting such a strategy, depictions could acknowledge uncertainty more transparently and offer a more constructive way of communicating their limitations.

This shift has the potential to make depictions more engaging while simultaneously clarifying the limits of their precision. Crucially, depictions are not intended to function as definitive portraits of identity but as prompts to recognition, and their provisional role must be explicitly communicated. Communicating this effectively is a challenge that extends beyond the laboratory or studio to investigative partners and the public. The persistent expectation, amplified by the "CSI effect," that science delivers certainty and precision stands in sharp contrast to the realities of forensic practice, where error and uncertainty are inevitable (Kaplan *et al.*, 2020; Schweitzer & Saks, 2007; Shelton, Kim & Barak, 2006). Moving beyond the reliance on a single static image offers one way to address this gap. Dynamic or variable presentations can make the limits of a forensic facial depiction more apparent, supporting clearer

[1] A similar account can be made for a contrasting, exemplar-based model that does not assume the presence of a central prototype (e.g., Lewis, 2004).

communication of uncertainty to both investigative partners and the public. Far from undermining forensic science, explicitly acknowledging these limits strengthens investigative practice by situating depictions realistically as aids to recognition rather than proof of identity. Dynamic formats may contribute here, as motion or morphing variation can enrich how observers locate a depiction within their multidimensional face space (e.g., Valentine, 1991), increasing flexibility in recognition while also foregrounding its provisional nature. However, broadening representational space also increases the risk of false positives, which may dilute investigative resources in contexts that are already tightly constrained—however, it should be noted that false positives provide value by usefully allowing persons-of-interest to be discounted (e.g., Davidson et al., 2025). This dual effect, offering both potential gains and new risks, marks a central challenge for forensic practice. It is important to note, however, that the shape information in the present stimuli constitutes ground truth; a craniofacial reconstruction will inevitably contain morphological inaccuracies, and it remains an open question whether further views of a face that is not totally accurate would yield equivalent results. Nonetheless, advances in dynamic presentation offer a route toward addressing such limitations.

One promising direction within this agenda is the incorporation of naturalistic motion and viewer agency (e.g., Lander *et al*., 1999). Dynamic features such as subtle expressions, controlled movements, or smooth rotations, now increasingly feasible with advanced rendering tools like *Epic Games' MetaHuman Creator*, could provide richer perceptual information. Interactive functions, such as toggling between motions, adjusting rotation speed, or manually exploring a face in 3D space, may encourage deeper engagement and help viewers focus on structural cues in which we can place greater confidence, rather than surface details that are more prone to error. In this way, interactive dynamic presentation offers a practical extension of the current findings: broadening the perceptual 'space' in which recognition can occur, while foregrounding the need for empirical testing to ensure that such innovations enhance, rather than dilute, investigative effectiveness. Substantial advances in reconstruction have produced increasingly precise models, capable of approximating true morphology within millimetre tolerances (e.g., Wilkinson *et al*., 2006). Yet, recognition cannot be secured by anatomical accuracy alone. The way a face is depicted, whether static or dynamic, occluded or variable, plays an equally critical role in whether it will be recognised. To move forward, the discipline must not only continue refining reconstruction methods but also systematically evaluate how presentation formats can balance recognition benefits against the risks of error.

A potential limitation of the present design is that familiarity was established retrospectively using a post-task ground-truth check. Although this procedure is standard in naming-based studies using celebrity identities, it cannot perfectly separate identities that were already known from identities that may have become easier to identify following exposure during the experiment. However, any such learning would have been distributed across conditions due to randomisation, and would likely reduce overall naming performance rather than inflate differences between presentation formats. For this reason, the check was treated as a conservative inclusion filter rather than as a graded measure of baseline familiarity. Future work could strengthen this aspect further by collecting a brief pre-task familiarity screen or by obtaining independent familiarity ratings prior to exposure, while balancing this against the risk that pre-screening itself can prime naming performance. Alternatively, participants can be recruited from a defined population where familiarity with the stimulus identities is established a priori, for example by using locally familiar targets such as lecturers on a programme and their students, where exposure history is predictable and can be verified directly.

Conclusion

This study demonstrates that recognition from forensic facial depictions depends critically on both the quantity of structural information provided and the format in which that information is displayed. Dynamic presentation did not improve recognition on first exposure when compared to three static views, but it did confer an advantage after repeated presentation, suggesting that motion supports engagement and consolidation rather than immediate recognition. By contrast, when static formats offered richer information through five distinct views, they outperformed dynamic displays in both overall accuracy and correct naming, without increasing misidentification. Together, these findings highlight two central insights: providing more than one view of a face (esp. those that include three-quarter views) enhances recognition potential, and dynamic presentation serves primarily to heighten engagement.

Theoretically, this distinction clarifies the role of motion in face perception. Motion appears to capture and sustain attention, encouraging greater cognitive investment in processing a face, yet recognisability ultimately depends on the informational richness of the depiction and how well its cues align with stored memory representations. Multiple static perspectives can therefore rival or exceed the benefits of motion when they deliver sufficiently comprehensive structural detail.

Practically, the findings suggest that standardising multi-view static dissemination remains the most reliable approach under current conditions. At the same time, dynamic or interactive formats may hold particular value when depictions are unavoidably incomplete or uncertain. While this study did not directly examine viewer agency, the engagement benefits observed for dynamic presentation suggest that giving observers greater control over how a face is explored, for example, rotating it manually or toggling between plausible variants, could prove a useful direction for future research. Such approaches would build on the attentional advantages identified here and may ultimately enrich recognition opportunities.

Finally, it is important to acknowledge the ethical and practical challenges of depicting faces in increasingly vivid and technologically advanced ways. Questions of consent, dignity, sensitivity to families and the risk of overstating certainty weigh heavily on the adoption of dynamic techniques. As digital capabilities expand, the discipline must balance these responsibilities with the opportunities offered by innovation. Through transparent communication of both the limits and the provisional nature of forensic facial depictions, their value can be strengthened, ensuring they are used responsibly and effectively while also leveraging their engaging qualities to support recognition.

Practical Implications

- Multi-view static presentation (esp. those that include a three-quarter view) should be prioritised for public dissemination when the goal is to maximise accurate recognition, as increasing viewpoint availability produced the most reliable benefits across experiments.
- Dynamic rotation may be useful in contexts where repeated exposure is expected, because recognition gains for dynamic presentation emerged most clearly at the second viewing.

- Repeated exposure can increase correct naming, but it can also, to a lesser extent, increase mistaken naming. For applied casework, this reinforces the need for clear reporting and triage processes to manage false leads.
- Where platform constraints limit the number of images that can be shared, practitioners should prioritise additional static viewpoints rather than substituting motion, particularly when only a single exposure is likely.
- Decisions about depiction presentation should be documented as part of casework reporting, since presentation format can influence both correct recognition and the rate of mistaken identifications.

References:

Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review, 12*(6), 1043–1047. https://doi.org/10.3758/BF03206441

Bailenson, J. N., Beall, A. C., Loomis, J. M., Blascovich, J., & Turk, M. (2002). Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, 13(4), 428–441. https://doi.org/10.1162/1054746027602043l8

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Blank, H., & Yovel, G. (2011). The structure of face space: How holistic processing is reflected in individual differences. *Journal of Experimental Psychology: Human Perception and Performance, 37*(3), 616–625. https://doi.org/10.1037/a0021431

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution, 24*(3), 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Bruce, V. (1994). Stability from variation: The case of face recognition. *Quarterly Journal of Experimental Psychology, 47A*(1), 5–28. https://doi.org/10.1080/14640749408401141

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*(3), 305–327. https://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Bruce, V., Hancock, P.J.B., Burton, A.M. (1998). Human Face Perception and Identification. In: Wechsler, H., Phillips, P.J., Bruce, V., Soulié, F.F., Huang, T.S. (eds) Face Recognition. NATO ASI Series, vol 163. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-72201-1_3

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science, 40*(1), 202–223. https://doi.org/10.1111/cogs.12231

Claes, P., Vandermeulen, D., & Suetens, P. (2012). Morphological variability of facial soft tissue depth: A metric analysis. *Forensic Science International, 219*(1–3), 75–80. https://doi.org/10.1016/j.forsciint.2011.11.017

Clutterbuck, R., & Johnston, R. A. (2002). Exploring levels of face familiarity by using an indirect face-matching measure. *Perception, 31*(8), 985–994. https://doi.org/10.1068/p3335

Coleman, S. R., & Grover, R. (2006). The anatomy of the aging face: Volume loss and changes in 3-dimensional topography. *Aesthetic Surgery Journal, 26*(1S), S4–S9. https://doi.org/10.1016/j.asj.2005.09.012

Davidson, C. I., Frowd, C. D., & Houlton, T. M. R. (2025). Witness artistic rendition and its impacts on visual memory for forensic facial composite creation. *Journal of Forensic Practice*. https://doi.org/10.1108/JFP-03-2025-0026

Davy-Jow, S. (2013). The visibility of forensic facial depiction: Contexts, methods, and public engagement. *Forensic Science Policy & Management, 4*(3–4), 85–94. https://doi.org/10.1080/19409044.2013.858367

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8*(4), 431–439. https://doi.org/10.1068/p080431

Erickson, W. B., Brown, C., Portch E., Lampinen, J. M., Marsh, J. M., Fodarella, C., Petkovic, A., Coultas, C., Newby, A., Data, L., Hancock, P. J. B., & Frowd, C. D. (2022). The impact of weapons and unusual objects on the construction of facial composites, *Psychology, Crime & Law*. DOI: 10.1080/1068316X.2022.2079643

Field, A. P. (2024). *Discovering statistics using IBM SPSS statistics* (6th edition.). Sage.

Frowd, C. D. (2021). Forensic Facial Composites. In M. Toglia, A. Smith, A., & J. M. Lampinen (Eds.) *Methods, Measures, and Theories in Forensic Facial-Recognition*. Taylor and Francis (pp. 34-64). Taylor and Francis.

Frowd, C. D., Jones, S., Fodarella, C., Skelton, F. C., Fields, S., Williams, A., Marsh, J., Thorley, R., Nelson, L., Greenwood, L., Date, L., Kearley, K., McIntyre, A., & Hancock, P. J. B. (2014). Configural and featural information in facial-composite images. *Science & Justice*, *54*, 215-227.

Frowd, C. D., Portch, E., Estudillo, A. J., Ford, C. J., Purcell, A., Pitchford, M., & Brown, C. (2025). The value of whole-face procedures for the construction and naming of identifiable likenesses for recall-based methods of facial-composite construction. *Applied Cognitive Psychology*, *39*(4). https://doi.org/10.1002/acp.70015

Garson, G. D. (2019). *Logistic regression: Regression analysis and generalized linear models.* Statistical Associates Publishers.

Hill, H. & Bruce, V. 1996. Effects of lighting on the perception of facial surfaces. Jour nal of experimental psychology. Human perception and performance, 22, 986-1004.

Hill, H., Schyns, P. G., & Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition*, 62(2), 201–222. https://doi.org/10.1016/S0010-0277(96)00785-8

Hole, G. J., & Bourne, V. J. (2010). *Face perception: A practical guide to social perception.* Psychology Press.

IBM Corp. (2020). *IBM SPSS Statistics for Windows, Version 29.0.2.0.* IBM Corp.

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition*, 121(3), 313–323.

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1571), 1671–1683. https://doi.org/10.1098/rstb.2010.0379

Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory, 17*(5), 577–596. https://doi.org/10.1080/09658210902976969

Joyce, S., & Frowd, C. D. (2025). Drawing the artist: the advantage of artistic ability for the construction of holistic facial-composite images. *Journal of Forensic Practice*, 1–20. https://doi.org/10.1108/JFP-07-2025-0073

Kaplan, J., Ling, S., & Cuellar, M. (2020). Public beliefs about the accuracy and importance of forensic evidence in the United States. *Science & Justice, 60*(3), 263–272. https://doi.org/10.1016/j.scijus.2020.01.001

Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition, 4*(3), 265–273. https://doi.org/10.1080/713756764

Lander, K., & Bruce, V. (2003). The role of motion in learning new faces. *Visual Cognition, 10*(8), 897–912. https://doi.org/10.1080/13506280344000149

Lander, K., Bruce, V., & Hill, H. (2001). Evaluating the effectiveness of movement in recognising unfamiliar faces. *Perception, 30*(8), 969–979. https://doi.org/10.1068/p3132

Lander, K., Christie, F., & Bruce, V. (1999). The role of movement in the recognition of famous faces. *Memory & Cognition, 27*(6), 974–985. https://doi.org/10.3758/BF03201228

Latif, A., & Moulson, M. C. (2022). The effect of face covering experience on holistic face processing in Egyptian adults. *Cognitive Research: Principles and Implications, 7*(1), 73. https://doi.org/10.1186/s41235-022-00434-4

Lee, W.-J., Wilkinson, C. M., & Hwang, H.-S. (2012). An accuracy assessment of forensic computerized facial reconstruction employing cone-beam computed tomography from live subjects. *Journal of Forensic Sciences, 57*(2), 318–327. https://doi.org/10.1111/j.1556-4029.2011.01969.x

Lee, W.-J., Wilkinson, C. M., & Hwang, H.-S. (2015). Correlation between facial reconstruction accuracy and recognition rates: Using 3D CT data from living subjects.

Journal of Forensic Sciences, 60(1), 180–187. https://doi.org/10.1111/1556-4029.12628

Lewis, M. B. (2004). Face-space-R: Towards a unified account of face recognition. *Visual Cognition, 11*(1), 29–69. https://doi.org/10.1080/13506280344000194

Liu, C. H., Collin, C. A., Burton, A. M., & Chaudhuri, A. (2005). Lighting direction affects recognition of untextured faces. *Vision Research, 45*(2), 243–250. https://doi.org/10.1016/j.visres.2004.07.052

Liu, C. H., & Chaudhuri, A. (2002). Reassessing the 3/4 view advantage in face recognition. *Cognition*, 83(1), 31–48.

Megreya, A. M., & Bindemann, M. (2009). Revisiting the processing of internal and external features of unfamiliar faces: The headscarf effect. *Perception, 38*(12), 1831–1848. https://doi.org/10.1068/p6380

Meinhardt-Injac, B., & Hildebrandt, A. (2016). Age-related changes in face processing: Development from childhood to old age. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 517–536). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199559053.013.0028

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*(1), 3–35. https://doi.org/10.1037/1076-8971.7.1.3

Miranda, G. E., Wilkinson, C. M., Roughley, M., & Montoya, F. (2018). Assessment of the accuracy of computer-generated facial reconstruction. *International Journal of Legal Medicine, 132*(3), 1061–1071. https://doi.org/10.1007/s00414-017-1699-8

Osborne, J. W. (2017). Regression & Linear Modeling: Best Practices and Modern Methods. Simple Linear Models With Categorical Dependent Variables: Binary Logistic Regression. Ch. 5. https://dx.doi.org/10.4135/9781071802724

O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences, 6*(6), 261–266. https://doi.org/10.1016/S1364-6613(02)01908-3

Portch, E., Brown, C., Fodarella, C., Jackson, E., Hancock, P. J. B., Tredoux, … Frowd, C. D. (2025). The impact of forensic delay: facilitating facial composite construction using an early-recall retrieval technique. *Ergonomics*, 1. https://doi.org/10.1080/00140139.2025.2519876

Raymond, J. E. (2022). Visual recognition and perceptual completeness: The role of occlusion in face processing. *Journal of Vision, 22*(4), 1–14. https://doi.org/10.1167/jov.22.4.1

Roughley, M., & Wilkinson, C. (2019). Computerized forensic facial reconstruction: A review of current methods and future directions. *Forensic Sciences Research, 4*(4), 315–326. https://doi.org/10.1080/20961790.2019.1662201

Schweitzer, N. J., & Saks, M. J. (2007). The CSI effect: Popular fiction about forensic science affects public expectations about real forensic science. *Jurimetrics, 47*(3), 357–364.

Shelton, D. E., Kim, Y. S., & Barak, G. (2006). A study of juror expectations and demands concerning scientific evidence: Does the “CSI Effect” exist? *Vanderbilt Journal of Entertainment & Technology Law, 9*(2), 331–368.

Short, J. M., Cattaneo, C., Gibelli, D., & Poppa, P. (2014). 3D computerized forensic craniofacial reconstruction: Evaluation of a new approach. *International Journal of Legal Medicine, 128*(5), 873–880. https://doi.org/10.1007/s00414-013-0940-y

Smith, E. (2020). Artistic interpretation in forensic facial reconstruction: Limits and possibilities. *Forensic Imaging, 21,* 200368. https://doi.org/10.1016/j.fri.2020.200368

Smith, K. (2020). *Laws of the face: Re-staging forensic art as a counter-forensic device* (Doctoral dissertation, Liverpool John Moores University). https://researchonline.ljmu.ac.uk/id/eprint/14304

Smith, K., & Wilkinson, C. (2024). The Doppelgänger effect? A comparative study of forensic facial depiction methods. *Forensic Science International, 356,* 111935. https://doi.org/10.1016/j.forsciint.2024.111935

Sporer, S. L., & Martschuk, N. (2014). The accuracy–confidence relation in eyewitness identification: Effects of repetition and retention interval. *Psychology, Crime & Law, 20*(6), 574–594. https://doi.org/10.1080/1068316X.2013.876500

Stephan, C. N., Huang, A. J., & Davidson, P. L. (2014). Assessing facial approximation accuracy: How to measure it, and why. *Forensic Science International, 237,* 147.e1–147.e13. https://doi.org/10.1016/j.forsciint.2014.01.004

Suhkre, R., Wattenmaker, W. D., & Johnston, R. A. (2015). The effect of partial occlusion on recognition of unfamiliar faces. *Attention, Perception, & Psychophysics, 77*(7), 2300–2312. https://doi.org/10.3758/s13414-015-0935-y

Taylor, K. T. (2001). *Forensic art and illustration.* CRC Press.

Ullrich, H., & Stephan, C. N. (2011). On the matching accuracy of facial approximations generated by computer modeling: Comparing computer-generated 3D reconstructions with photographs of living persons. *Journal of Forensic Sciences, 56*(3), 672–678. https://doi.org/10.1111/j.1556-4029.2010.01685.x

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology, 43A*(2), 161–204. https://doi.org/10.1080/14640749108400966

Valentine, T. (2001). Face-space models of recognition. In M. J. Wenger & J. T. Townsend (Eds.), *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges* (pp. 83–113). Lawrence Erlbaum Associates.

Wilkinson, C. (2004). *Forensic facial reconstruction.* Cambridge University Press.

Wilkinson, C., Rynn, C., Peters, H. (2006). A blind accuracy assessment of computer-modeled forensic facial reconstruction using computed tomography data from live subjects. *Forensic Science, Medicine, and Pathology, 2*, 179-187.

Wilkinson, C. (2010). Facial reconstruction – anatomical art or artistic anatomy? *Journal of Anatomy, 216*(2), 235–250. https://doi.org/10.1111/j.1469-7580.2009.01182.x

Wilkinson, C. (2015). Forensic facial reconstruction: Theory and practice of identification. In R. Siegel & P. Saukko (Eds.), *Encyclopedia of Forensic Sciences* (2nd ed., pp. 354–361). Academic Press.

Wilkinson, C., Liu, C. Y. J., Shrimpton, S., & Greenway, E. (2024). Craniofacial identification standards: A review of reliability, reproducibility, and implementation.

Forensic Science International, 365, 111935. https://doi.org/10.1016/j.forsciint.2024.111935

Wilkinson, C., Rynn, C., Peters, H., Taister, M., Kau, C. H., & Richmond, S. (2006). A blind accuracy assessment of computer-modeled forensic facial reconstruction using computed tomography data from live subjects. *Forensic Science, Medicine, and Pathology, 2*(3), 179–187. https://doi.org/10.1007/s12024-006-0012-9

Won-Joon, C., & Wilkinson, C. (2016). The effect of practitioner background on forensic facial reconstruction accuracy. *Journal of Forensic Sciences, 61*(2), 401–407. https://doi.org/10.1111/1556-4029.12922

Won-Joon, L., Chung, J., & Wilkinson, C. (2016). Accuracy of three-dimensional computerized forensic craniofacial reconstruction. *Forensic Science International, 262,* 182–189. https://doi.org/10.1016/j.forsciint.2016.03.026

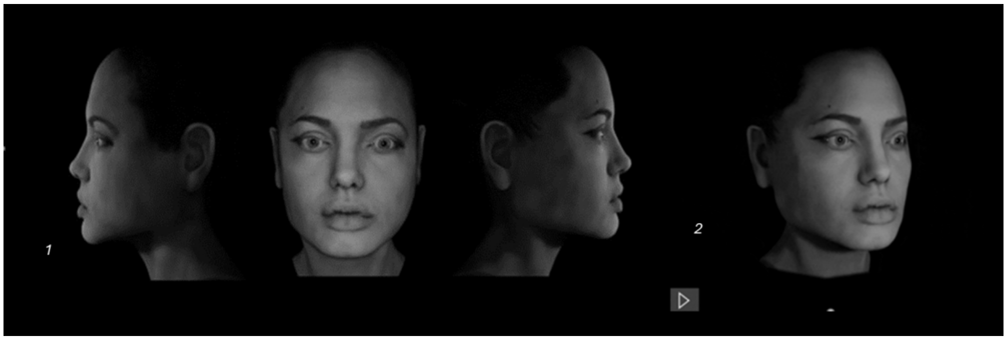

Figure 1- Example stimulus. Angelina Jolie in display condition 1, static presentation of frontal view, two profile views (left) and display condition 2, screen shot of the dynamic gif presentation (right). Source: Authors own work.

344x115mm (130 x 130 DPI)

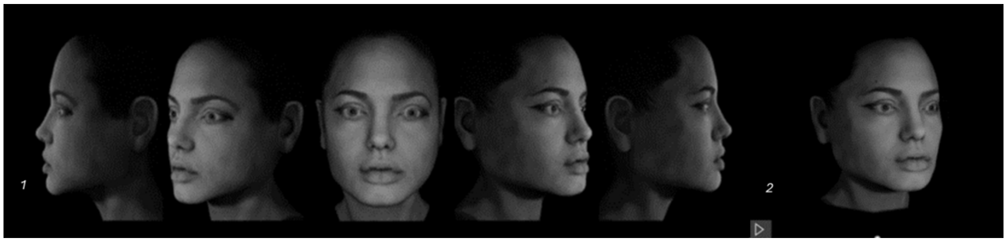

Figure 2 - Example stimulus. Angelina Jolie in display condition 1, static presentation of frontal view, two profile and two ¾ views (left) and display condition 2, screen shot of the dynamic gif presentation (right). Source: Authors own work.

392x91mm (130 x 130 DPI)

118x118mm (72 x 72 DPI)

Table I - Experiment 1 Correct, Incorrect and Accuracy Naming Means.

Naming Type	*Order*	*Mode*		*Mean*
		Static	Dynamic	
Correct	First	53.8 (135 / 251)	55.7 [b] (141 / 253)	54.8 [a] (276 / 504)
	Second	55.0 [c] (138 / 251)	71.1 [b, c] (180 / 253)	63.1 [a] (318 / 504)
	Mean	54.4 (273 / 502)	63.4 (321 / 506)	58.9 (594 / 1008)
Incorrect	First	34.5 (40 / 116)	39.3 (44 / 112)	36.8 [d] (84 / 228)
	Second	47.8 (54 / 113)	43.8 (32 / 73)	46.2 [d] (86 / 186)
	Mean	36.8 (84 / 228)	41.1 (76 / 185)	41.1 (170 / 414)
Accuracy	First	37.9 (95 / 251)	38.3 (97/ 253)	38.1 (192 / 504)
	Second	33.5 [e] (84 / 251)	58.5 [e] (148 / 253)	46.0 (232 / 504)
	Mean	35.7 (179 / 502)	48.4 (245 / 506)	42.1 (424 / 1008)

Note. For Correct Naming, figures are percentage correct; numbers in parentheses are total number of correct responses against total number of responses. For Incorrect Naming, figures are percentage incorrect; numbers in parentheses are total number of incorrect responses divided by total number of mistaken and no-naming responses. For Accuracy, figures are percentage; numbers in parentheses are number of correct minus mistaken responses divided by total number of responses. [a] $p = .006$, [b] $p < .001$, [c] $p < .001$, [d] $p = .018$, [e] $p < .001$.

Table II - Experiment 1 Correct Naming Final Model.

Fixed Effects	*F*	*DF1*	*DF2*	*p*
Presentation	7.52	1	1004	.006
Mode	11.39	1	1004	< .001
Presentation × Mode	7.50	1	1004	.006

Table III - Experiment 1 Correct Naming Coefficients for Presentation × Mode

Fixed Effects	*B*	*SE(B)*	*p*	*Exp(B)*	*95% CI (– / +)*
Dynamic vs Static (First presentation)	0.10	0.21	.63	1.11	0.73 – 1.67
Dynamic vs Static (Second presentation)	0.93	0.25	< *.001*	2.53	1.45 – 4.07
Second vs First (Static mode)	0.08	0.23	.73	1.08	0.69 – 1.70
Second vs First (Dynamic mode)	0.91	0.26	< *.001*	2.48	1.43 – 4.02

Note. Overall Correct Classification: 76.9%. The model was specified with the lowest category of categorical predictors as reference (Static, First), and predictors and target were sorted in a descending order. (For the final model:) Information criteria are based on the -2 log likelihood (*AICC* = 4732.13, *BIC* = 4741.94); random effects were random intercepts for participants [σ = 0.73, *SE(σ)* = 0.22] and items [σ = 1.48, *SE(σ)* = 0.68], and random slopes for Presentation for items [σ = 0.05, *SE(σ)* = 0.08]. The intercept was [*B* = 0.13, *SE(B)* = 0.41].

Table IV - Experiment 1 Incorrect Naming Final Model.

Fixed Effects	*F*	*DF1*	*DF2*	*p*
Presentation	5.64	1	412	.018

Table V - Experiment 1 Coefficients for Presentation for Incorrect Naming.

Fixed Effects	*B*	*SE(B)*	*t(412)*	*p*	*Exp(B)*	*95% CI(-/+)*
Second - First	0.51	0.22	2.37	.018	1.67	1.09 – 2.55

Note. Overall Correct Classification: 73.4%. The model was specified with the lowest category of categorical predictors as reference (First), and predictors and target were sorted in a descending order. Information criteria are based on the -2 log likelihood (*AICC* = 1826.50, *BIC* = 1834.52); random intercepts for participants [σ = 0.82, *SE(σ)* = 0.32] and items [σ = 0.13, *SE(σ)* = 0.12]. Intercept [*B* = -0.57, *SE(B)* = 0.23].

Table VI - Experiment 1 Naming Accuracy Final Model

Fixed Effects	*F*	*DF1*	*DF2*	*p*

Mode	10.05	1	1003	.002
Presentation	4.55	1	1003	.033
Presentation × Mode	9.13	1	1003	.003

Table VII - Naming Accuracy – Coefficients for Comparisons for Presentation × Mode

Fixed Effects	*B*	*SE(B)*	*p*	*Exp(B)*	*95% CI (– / +)*
Dynamic vs Static (First presentation)	0.23	0.19	.90	1.02	0.71 – 1.47
Dynamic vs Static (Second presentation)	0.95	0.21	< .001	2.59	1.69 – 3.97
Second vs First (Static mode)	0.18	0.20	.60	1.20	0.75 – 1.61
Second vs First (Dynamic mode)	0.90	0.22	< .001	2.46	1.50 – 4.02

Note. Pairwise comparisons for the Presentation × Mode interaction in the Naming Accuracy GLMM (Experiment 1). Significant effects show that accuracy improved across presentations in the dynamic condition but not in the static condition. Overall Correct Classification: 65.2%. The model was specified with the lowest category of categorical predictors as reference (First, Dynamic), and predictors and target were sorted in a descending order. Information criteria are based on the -2 log likelihood (*AICC* = 7832.67, *BIC* = 7842.48); random intercepts for participants [σ = 0.38, *SE(σ)* = 0.13] and items [σ = 1.06, *SE(σ)* = 0.49], and random slopes for Presentation for items [σ = 0.01, *SE(σ)* = 0.05]. Intercept [*B* = -1.81, *SE(B)* = 0.38].

Table VIII – Experiment 2 Correct, Incorrect and Accuracy of Naming Means.

Naming Type	*Order*	*Mode*		Mean
		Static	Dynamic	
Correct	First	51.1 (114 / 223)	41.1 (85 / 207)	46.3[a] (199 / 430)
	Second	54.3 (121 / 223)	46.9 (97 / 27)	50.7[a] (218 / 430)
	Mean	52.7[b] (235 / 446)	44.0[b] (182 / 414)	48.5 (417 / 860)
Incorrect	First	11.2 (25 / 223)	13.5 (28 / 207)	12.3 (53 / 430)
	Second	10.8 (24 / 223)	15.5 (32 / 207)	13.0 (56 / 430)
	Mean	12.3 (53 / 430)	14.5 (60 / 414)	12.7 (109 / 860)
Accuracy	First	39.9 (89 / 223)	27.5 (57 / 207)	34.0 (146 / 430)
	Second	43.5 (97 / 223)	31.4 (65 / 207)	37.7 (162 / 430)
	Mean	41.7[c] (186 / 446)	29.5[c] (122 / 414)	35.8 (308 / 860)

Note. See Table I, *Note* for definition of figures. [a] p = .044, [b] p = .052, [c] p = .013.

Table IX - Experiment 2 Correct Naming Final Model.

Fixed Effects	*F*	*DF1*	*DF2*	*p*
Mode	3.80	1	857	.018

Table X - Experiment 2 Coefficients for Mode for Correct Naming.

Fixed Effects	*B*	*SE(B)*	*t(857)*	*p*	*Exp(|B|)*	*95% CI(-/+)*
Static - Dynamic	-0.35	0.18	-1.95	.05	1.41	1.06 – 1.88

Note. Overall Correct Classification: 64.0%. The model was specified with the lowest category of categorical predictors as reference (Dynamic, Second), and predictors and target were sorted in a descending order. Information criteria are based on the -2 log likelihood (AICC = 3763.01, BIC = 3767.76); random intercepts for items [σ = 0.56, SE(σ) = 0.27]. Intercept [*B* = -0.05, *SE(B)* = 0.27].

Table XI - Experiment 2 Naming Accuracy Final Model.

Fixed Effects	*F*	*DF1*	*DF2*	*p*
Mode	6.18	1	857	.013

Table XII - Experiment 2 Coefficients for Mode for Naming Accuracy.

Fixed Effects	B	SE(B)	t(857)	p	Exp(|B|)	95% CI (-/+)
Dynamic - Static	-0.33	0.13	2.49	.013	1.40	1.07 – 1.81

Note. Overall Correct Classification: 60.1%. The model was specified with the lowest category of categorical predictors as reference (Static), and predictors and target were sorted in a descending order. Information criteria are based on the -2 log likelihood (AICC = 6560.80, BIC = 6565.55); random intercepts for items [σ = 0.30, $SE(\sigma)$ = 0.15]. Intercept [B = 0.90, $SE(B)$ = 0.19].