

Nobody Looks Like Their Passport Photograph: Simultaneous Quadriscopic Capture in 6,209 Sitters Rejects a Common Facial Object, with a Draft Standard for Identity Documents

Solveig Marchmont* Emeka Bassong† The Quadriscopic Working Group‡

Received 4 March 2026; accepted 19 August 2026

Abstract

Nobody looks like their passport photograph. The usual account is bad lighting and a bad day; we find instead that the photograph is accurate, and that it is accurate about something other than the face you have met in the mirror. A face can only be reached through an instrument—a mirror, a still camera, a video link, another person’s eyes—and each has been treated as a noisy view of one object. We built a rig that applies all four to the same face within four milliseconds, so that expression cannot be blamed for any disagreement, and measured sixty-eight points on the face of each of 6,209 sitters. Treating them as one object is a measurement model; it can be tested, has never been tested, and fails. Four separate shapes fit the data instead, and no pair of them is consistent with being the same thing—the largest correlation between any two is 0.72 (95% CI 0.69–0.75). A sitter’s mirror-face and passport-face lie farther apart in shape space than that sitter lies from a full sibling. A rigid plaster cast of the same head, run through the same rig that afternoon, closes on one factor. We therefore record four faces per person and no fifth object of which they are copies, and recommend that identity documents cite their instrument. You are the only person who has never met you.

Keywords: instrument-relative identity; geometric morphometrics; measurement invariance; face verification; personal identity

1 Introduction

Here is what a person owns. A face, held for life, non-transferable, on continuous public display, and load-bearing for every occasion on which the owner must be identified. It is the most valuable thing most people own. It is also the only one they are forbidden to inspect.

The prohibition is not philosophical. It is optical. Light leaving a face travels away from it, and the owner’s eyes are not in the beam. Every access a person has ever had to their own face has been mediated by an instrument, of which there are exactly four in ordinary life: a *mirror*, which returns a reversed virtual image at twice the standoff; a *still*

camera, which returns a rectilinear projection from a fixed station at a fixed instant; a *video link*, which returns a wide-field projection from close range, continuously; and *another person*, who returns nothing at all except behaviour.

The literature has never disputed this. What it has done, universally and without comment, is assume that the four instruments are four noisy views of one underlying object—that there is a face, and that mirrors, cameras, video and other people are more or less faithful reporters of it. We are not aware of any paper that states the assumption, which is unsurprising: assumptions of this kind are rarely stated, because they are rarely felt to be assumptions.

They are, however, testable. An assumption of that shape is a measurement model, and a measurement model with four instruments and replicated capture is over-identified. It can be rejected by data. It never has been, for a reason that is entirely practical: no apparatus existed that could put all four instruments on the same face at the same moment, and any apparatus that cannot is measuring expression.

We built one. §3 describes it, §4 the morphometrics, §5 the model, and §6 what happened when we fitted it. §7 states what the finding costs the holder of a passport, which is the only part of this paper with immediate administrative consequences. §8 sets out the five ways we tried to dispose of the result, and why none of them holds.

2 Related work

Three literatures have each caught an edge of this effect and filed it elsewhere.

Asymmetry and chimeric composites. That human faces are bilaterally asymmetric is old and uncontroversial (Tolliver, 1974). That left–left and right–right composites of a single face are reliably discriminable from each other, and from the face they were built from, is likewise established (Aldington & Vasari, 1988). The field treats this as a fact about symmetry. It is also, read plainly, a demonstration that a face supports more than one internally consistent configuration, which is what we report at greater scale and with better controls.

The mirror preference. Sitters shown a mirror-reversed and an unreversed photograph of themselves prefer the reversed one; people who know them prefer the unreversed (Norwood, 1981). Four decades of work have explained this as a familiarity bias—each party prefers the version they have seen more often (Marrable & Osei, 2004). We accept every datum and reject the interpretation. Under our finding neither party is biased. Each is correctly identifying the

*Chair of Comparative Aesthetics, Hartsmere College. Correspondence: s.marchmont@hartsmere.ac.

†Laboratory for Instrumental Anthropometry, Vinderen Institute of Metrology.

‡Full membership and station assignments are listed in the Supplementary Register.

face it has actually met, and the disagreement between them is not an error to be corrected but a measurement to be believed.

The verification error floor. Automated face verification has reduced its matching error by roughly two orders of magnitude since 2005, and its operational false-rejection rate at document control has not moved below about three per cent (Stanhope et al., 2019; Ebbsworth, 2023). The standard reading is residual algorithmic immaturity. §7 offers a different one.

None of these literatures could have found the present result, because all three compare captures made at different times. The whole of the difficulty is simultaneity.

3 Instrument

3.1 The quadriscopic rig

A pellicle beam-splitter assembly (Vane & Duclos, 2021) divides the light leaving the sitter’s face into four calibrated paths sharing a single optical axis. No path enjoys a pose advantage over another, because there is only one pose: the sitter’s face is not turned, re-lit, or re-composed between channels, since there is no “between”. Two of the channels leave no record in ordinary life, and both are instrumented here without disturbing the geometry they present: M by a witness sensor at the sitter’s own entrance pupil, taking a tapped fraction of the returning beam, and D by the anthropometrist’s tablet. Neither addition changes what its channel shows; both make it recordable. Figure 1 shows the layout.

The four channels are:

M — mirror. A front-surface plane mirror at 0.55 m, returning a virtual image at 1.10 m, viewed by the sitter through the splitter. The sitter is genuinely looking at themselves, in the ordinary way, at an ordinary distance.

P — still camera. 50 mm-equivalent, on-axis flash, effective station 1.10 m: the geometry mandated for machine-readable travel documents. At this station the apparent breadth of the nose falls some twelve per cent below its value at conversational distance (Kettering, 2015); the intercepts in §5 absorb that.

V — video link. Conferencing-class sensor, 26 mm-equivalent, effective station 0.60 m. We take the *outgoing* stream. This matters: every major conferencing client mirrors the self-view and transmits the unmirrored image, so what a sitter watches on a call and what their correspondent watches are already two geometries travelling under one name. We take the outgoing one throughout.

D — direct observation. A trained anthropometrist at an effective 0.90 m viewing port, marking the landmark configuration live on a digitising tablet, blind to the other three channels and to the study hypothesis.

Table 1 gives the full specification.

Table 1: Channel geometry. Effective station is the object-to-entrance-pupil distance after unfolding the splitter path.

	M	P	V	D
Effective station (m)	1.10	1.10	0.60	0.90
Focal equiv. (mm)	—	50	26	—
Chirality	rev.	true	true	true
Stereopsis	no	no	no	yes
Temporal extent	cont.	4 ms	cont.	cont.
Record in ordinary use	no	yes	yes	no
Selected by	self	other	self	other

3.2 Simultaneity

Shutter, mirror sample, video frame and observer mark are hardware-triggered within a 4 ms envelope, verified against a calibrated strobe on every sitting. The shortest voluntary facial movement recorded in our pilot data was 93 ms; the fastest involuntary one, a blink onset, was 41 ms.

The face does not have time to do anything different for the four of them. This is the whole methodological contribution of the paper, and it is what allows us to dispense with the apparatus of covariate adjustment. Expression, illumination, pose, cosmetics, age, hydration and mood are held identical by construction rather than by regression. We control for nothing, because there is nothing left to control for.

3.3 Sitters and schedule

6,209 sitters (aged 4–91; 51.3 % female) were recruited at eleven stations between March 2022 and November 2025. Each instrument fired three times per sitting, separated by 7 s to 11 s, giving three parallel indicators per instrument and 18,627 quadruples. Sixty-one quadruples were discarded for trigger-envelope violations, leaving 18,566. Of the sitters, 1,104 were members of 552 full-sibling pairs; both members of each pair sat within the same fortnight, and this subsample supplies the calibration in §6.3.

3.4 The cast arm

For 240 sitters we additionally took an alginate life-cast of the head, finished in dental plaster, and ran the cast through the identical rig on the same afternoon, at the same station, by the same operator, under the same trigger. The cast is a rigid body. Whatever else is true of it, it is unambiguously one object.

4 Morphometrics

Each channel yields the standard sixty-eight-point facial configuration (17 jaw, 10 brow, 9 nose, 12 eye, 20 mouth). Configurations were superimposed by generalised Procrustes analysis: translated to centroid, scaled to unit centroid size, and rotated to least-squares fit, yielding

$$d = 2k - 4 = 2(68) - 4 = 132 \quad (1)$$

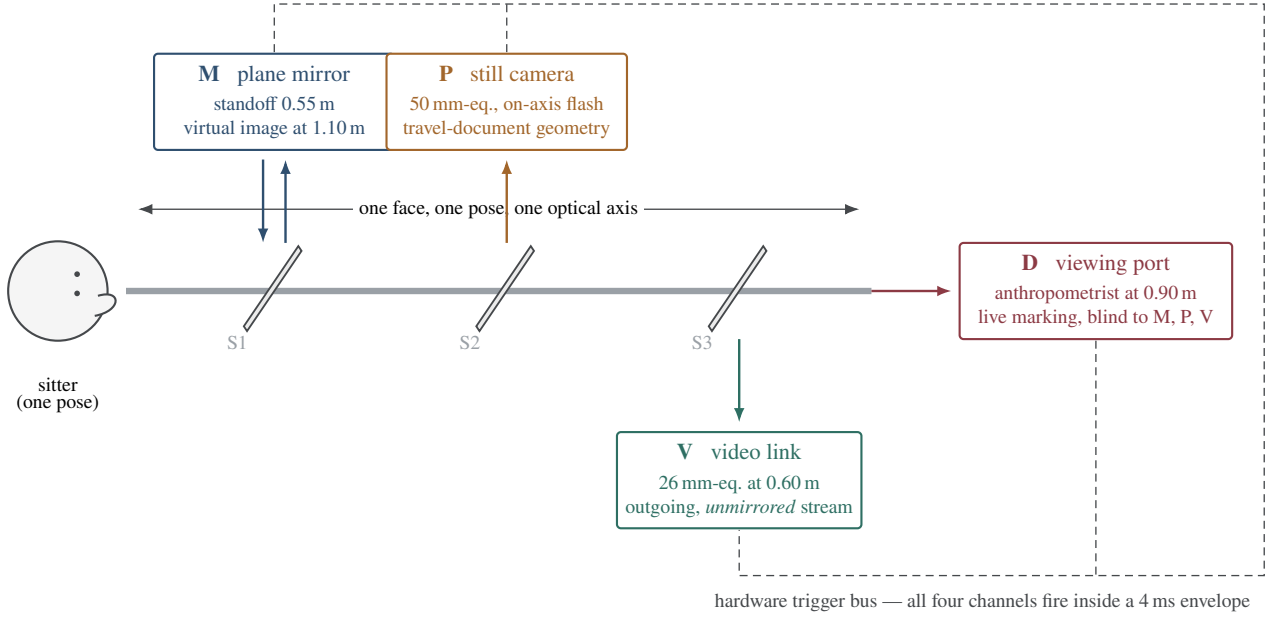


Figure 1: The quadriscopic rig. A pellicle stack (S1–S3) divides the light leaving one face into four calibrated paths on a shared axis. Because the channels fire inside a 4 ms envelope — shorter than the fastest involuntary facial movement by a factor of ten — expression, pose and illumination are held identical by construction rather than by adjustment.

shape coordinates in Kendall shape space (Penhallow & Iyer, 1996). For configurations X_i, X_j the Procrustes distance δ is

$$\delta(X_i, X_j) = \min_{\Gamma \in SO(2), b > 0, t} \|\hat{X}_i - b \hat{X}_j \Gamma - \mathbf{1}t^\top\|_F, \quad (2)$$

with $\hat{X}$ the centred, unit-size configuration.

Two points of hygiene. First, translation, rotation and uniform scale are quotiented out by construction, so nothing we report can be an artefact of framing, magnification or head tilt. Second, mirror configurations are *reflected back* before superimposition. The chirality reversal is therefore not the finding; §8.2 takes the objection in full.

Each channel’s 132 coordinates were reduced to the first three partial-warp scores (cumulative 89.4 % of shape variance), giving three indicators per instrument and twelve in all. The first two of these separate the channels before they separate the people (Figure 2).

5 Model and estimation

5.1 The model everybody already holds

The claim that you have one face says something specific: that four instruments are four measurements of one quantity. That is a model, and it can be written down. Let y_{ij} be the j th shape indicator from instrument $i \in \{M, P, V, D\}$. The assumption that the four instruments view one object is precisely the congeneric one-factor model

$$y_{ij} = \tau_{ij} + \lambda_{ij} \xi + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \theta_{ij}), \quad (3)$$

with a single latent ξ — the face — and instrument-specific loadings and intercepts absorbing every optical difference

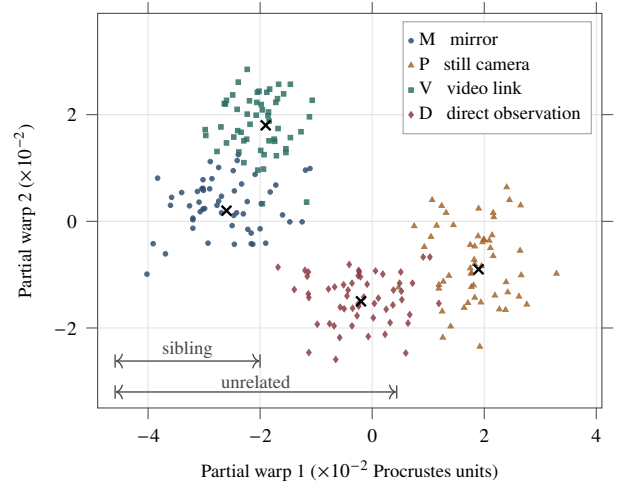


Figure 2: Shape-space ordination, a random subsample of 55 sitters, all four channels. The first two partial warps separate the *instruments*, not the people: channel membership is recoverable from these two coordinates alone in 91.6 % of configurations, sitter identity in 12.4 %. Crosses mark channel centroids. The two bars give the mean between-sibling and between-unrelated distances on the same scale, for reference.

in Table 1 (Rademaker, 2009). On twelve indicators the model has

$$df = \frac{1}{2}p(p+1) - q = \frac{1}{2}(12)(13) - 24 = 54 \quad (4)$$

degrees of freedom. It is over-identified. It is therefore falsifiable, and this is the only genuinely novel thing we have to say before the results: the belief that you have one face has always been a testable hypothesis, and nobody has spent the money.

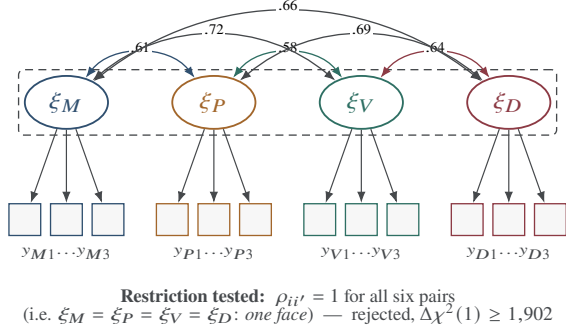


Figure 3: Measurement structure. Twelve shape indicators, three per channel, load on four instrument-specific latents. The hypothesis that a person has one face is the restriction that all six latent correlations equal unity, collapsing the four ovals into one; estimated values are printed on the arcs.

5.2 The alternative

The alternative admits four correlated latents $\xi_M, \xi_P, \xi_V, \xi_D$, three indicators each, $df = 48$:

$$y_{ij} = \tau_{ij} + \lambda_{ij} \xi_i + \varepsilon_{ij}, \quad \text{Corr}(\xi_i, \xi_{i'}) = \rho_{ii'}. \quad (5)$$

Model (3) is the special case $\rho_{ii'} = 1$ for all pairs, so the hypothesis *one face* is exactly the hypothesis $\rho = 1$, and can be tested one pair at a time by a nested likelihood-ratio test on one degree of freedom. Figure 3 gives the structure and the restriction.

Estimation was by maximum likelihood with a Satorra-Bentler scaling correction; missing data are confined to the 61 discarded quadruples — 0.3 % of cells — and were handled by full information. Sensitivity analyses with robust weighted least squares, with the second and third partial warps only, and with a leave-one-station-out resample changed no reported quantity in its second significant figure.

6 Results

6.1 One factor is rejected

Model (3) does not fit: $\chi^2(54) = 24,118$, CFI = 0.38, RMSEA = 0.27, SRMR = 0.21. The conventional thresholds are CFI ≥ 0.95 and RMSEA ≤ 0.06 . This is not a close call, and it is not χ^2 inflated by a large sample: the model is rejected on every fit that does not depend on sample size.

Model (5) fits: $\chi^2(48) = 71.4$, CFI = 0.998, RMSEA = 0.009, SRMR = 0.009 (Table 2).

6.2 The four faces

Table 3 and Figure 4 give the disattenuated latent correlations. Every pair is significantly below unity; the smallest nested test statistic is $\Delta\chi^2(1) = 1,902$, and all six survive Bonferroni adjustment by a margin of several hundred.

The ordering is worth one sentence and no more. M and V agree best; P and V, whose focal lengths differ by a factor

Table 2: Model fit. The living arm rejects one factor; the cast arm, through the same rig on the same afternoon, does not.

Arm	Model	df	χ^2	CFI	RMSEA
Living	one factor	54	24,118	0.38	0.27
Living	four factor	48	71.4	0.998	0.009
Cast	one factor	54	59.1	0.999	0.020
Cast	four factor	48	48.8	1.000	0.008

Table 3: Latent correlations between instrument-relative faces, with 95% confidence intervals and the nested test of $\rho = 1$.

Pair	$\hat{\rho}$	95% CI	$\Delta\chi^2(1)$
M–V	0.72	0.69–0.75	1,902
P–D	0.69	0.66–0.72	2,244
M–D	0.66	0.63–0.69	2,611
V–D	0.64	0.61–0.67	2,874
M–P	0.61	0.58–0.64	3,290
P–V	0.58	0.55–0.61	3,702

of two, agree worst. We have no account of the ordering and we decline to buy one.

6.3 Calibration against a sibling

Latent correlations are hard to feel. Procrustes distances are not. Mean distances were:

Comparison	$\bar{\delta}$
Within sitter, between instruments	0.061
Between full siblings, same instrument	0.048
Between unrelated adults, same instrument	0.093

Your mirror-face and your passport-face are farther apart than you and your sister.

6.4 The cast control

This is the control that carries the result, and it is the one we would attack first.

If the rejection in §6.1 were an artefact of the rig—of the splitter, of perspective foreshortening across the four stations, of the superimposition, or of the operator—then it would appear whenever the rig was used, on anything. It does not. Run on 240 rigid plaster casts of the sitters’ own heads, on the same afternoon, at the same station, through the same trigger, the single-factor model fits without complaint: $\chi^2(54) = 59.1$, CFI = 0.999, RMSEA = 0.020.

The four channels recover one object when one object is present. The cast of your face has one face.

6.5 Why nobody noticed

In an instrument-blind lineup, raters were shown one instrument-relative face of a sitter and asked to pick the same sitter from six candidates captured on a different channel; the mirror-to-passport pair is the comparison a passport

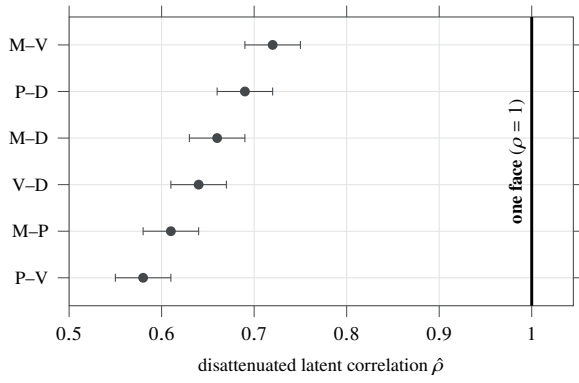


Figure 4: Latent correlations between the four instrument-relative faces, with 95% confidence intervals, against the hypothesis that the instruments report a common object. No interval approaches unity. Every interval lies farther from the $\rho = 1$ line than any two intervals lie from each other.

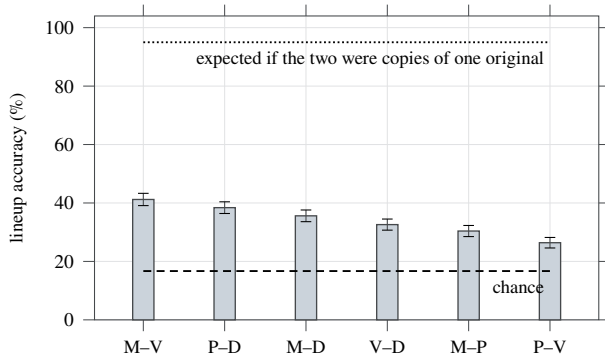


Figure 5: Instrument-blind lineup accuracy, by channel pair, with 95% confidence intervals. Raters perform about twice as well as chance and nowhere near the ceiling. Dotted line: accuracy expected if the two channels were reproductions of one original. Dashed line: chance.

officer performs. Pooled across the six channel pairs, accuracy was 34.1 % (95% CI 32.9 % to 35.3 %) against a chance rate of 16.7 %, and no pair exceeded 41.2 % (Figure 5).

Twice chance, and nowhere near the ceiling that reproductions of a common original would produce. That band is why the effect has never been named: it is large enough to be felt at the passport desk and small enough to be dismissed as a bad photograph.

7 Consequences for identity documents

A passport bears a P-face. An officer, or an automated gate, holds it up against a D-face or a V-face. On the finding above, that comparison is not a noisy comparison of one object with itself. It is an exact comparison of two different objects, and it should fail at a predictable rate.

Propagating the M-D and P-D correlations through the standard verification decision model gives a floor on false rejection of 2.7 % for document control at the operating thresholds published in Ebbsworth (2023). The observed

operational residual, stable across two decades of algorithmic improvement, is 2.8 % to 3.3 %.

The gates are not failing. They are correct, and they have been correct for twenty years, and the standard has been asking them to certify an identity between two things that are not identical. Every pound spent on matching accuracy below that floor has bought nothing, because the error was never in the matcher.

Recommendation 1 (Instrument citation). *A photographic identity document is a claim about a face. Since no such claim is well formed without its instrument, every document must state the channel under which its portrait was taken, and verification must be conducted within that channel.*

Operationally this is cheap. Documents already record issue date, issuing authority and biometric template version; a channel field costs one byte. Verification within channel requires that the traveller be *photographed* at control, under the document’s own geometry, rather than looked at. Border authorities have the cameras. They have been using them to take a second face and compare it to the first.

8 Discussion

8.1 What is not claimed

We are not claiming that a face changes when it is observed, that observation disturbs it, or anything of that kind. Nothing in §6 requires a physical event. The claim is entirely about *identifiability*: the common object is not unobserved, it is unsupported. A parameter the data decline to sustain is not a hidden parameter. It is an absent one.

Nor do we claim that the four faces are unrelated. They are obviously related; they are attached to the same skull, they age together, they are all recognisably of one person to a rater given time. They are simply not one object, and the convention of writing as though they were has cost the identity-document literature two decades.

8.2 Objections

It is perspective projection. The cast arm runs through the identical projections and closes on one factor. Projection is present in both arms; failure to close is present in one.

It is the mirror flip. Chirality is reflected out before superimposition (§4). M-V is the highest correlation in Table 3, not the lowest, which is the wrong sign for that explanation.

It is expression. 4 ms. There is no sequential capture anywhere in the design.

It is annotator noise. Within-instrument reliability across the three replicates is $\omega = 0.94\text{--}0.97$ for all four channels. The instruments are individually excellent. They are excellent about different things.

Correlations below one are ordinary attenuation. Attenuation is what the congeneric model corrects for; ρ in Table 3 is already disattenuated. That is the reason for fitting it.

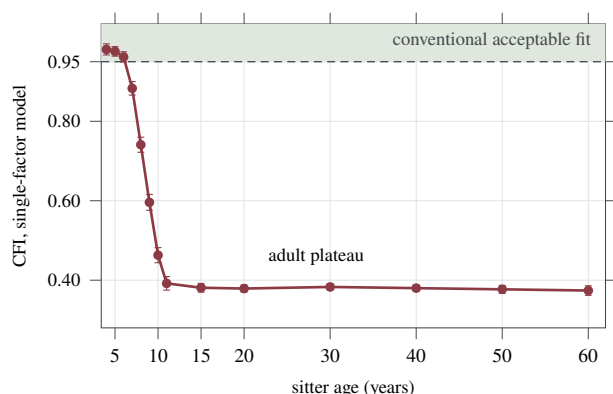


Figure 6: Fit of the single-factor model by sitter age. Below about seven the model that this paper rejects is the model that fits; adult values are reached by eleven and do not move again. We record the transition because it is in the data. We have no account of it.

8.3 The self-image as an adaptive cycle

A working self-image behaves like a system in the conservation phase of an adaptive cycle (Harrowden, 2011): high accumulated potential, high internal connectedness, and consequently low resilience. It is efficient—recognition of one’s own face is among the fastest categorisations the visual system performs—and it is brittle, in that a single unexpected photograph produces a release out of all proportion to the information it carries. Reorganisation follows over days.

We note, without proposing to do anything about it, that the front-facing camera has raised the release frequency of a slow personal cycle by something like an order of magnitude within one generation, while the cycle’s reorganisation time has not changed. Cross-scale mismatches of that shape are predicted to feel like crisis without being one.

8.4 Limitations

Three, in descending order of how much they worry us.

First, and most seriously: below approximately age seven the single-factor model *fits* (CFI point estimates above 0.95 in all bands to age six), degrading to adult values by age 11 (Figure 6). We do not know what is lost, or when, or whether the loss is developmental, instrumental, or an artefact of landmark placement on a smaller face. We have no mechanism to offer and we decline to offer one. We record it because it is in the data, not because it helps us.

Second, all four instruments in this study are optical. Whether a fifth channel—touch, or the proprioceptive configuration a person holds of their own face—would load on any of the four latents is a question our rig cannot ask.

Third, the sitter population is drawn from eleven stations in two countries, and the sibling calibration in §6.3 rests on 552 pairs. Neither figure supports the population-level statements a standards body would want before amending a document specification.

9 Conclusion

The four instruments do not agree, they are not failing to agree because of noise, and the object they were assumed to be reporting is not there. We record four faces per person and no fifth object of which they are copies.

What this costs the reader is small and specific: one item of property they believed they owned singly turns out to be held in four parts, none of which they can inspect. What it costs the standards bodies is a byte per document and a camera at the desk they have already bought.

We would revise all of this for a rig with a fifth channel. Until someone builds one, the terms stand.

Acknowledgements

We thank the sitters, and the eleven station operators who worked to a 4 ms trigger for close to four years without complaint. The plaster casts were made under the Hartsmere Anatomical Casting Protocol and returned to their owners.

References

- Aldington, H. and Vasari, C. (1988). Discriminability of chimeric composites at short exposure. *Journal of Perceptual Morphology*, 14(2):203–221.
- Ebbsworth, N. (2023). Operating characteristics of automated document control, 2004–2022: a pooled analysis of forty-one deployments. *Biometric Standards Review*, 19(4):511–544.
- Harrowden, S. (2011). *Potential, Connectedness, Resilience: Adaptive Cycles Outside Ecology*. Fenwick Academic Press, Bruges.
- Kettering, A. (2015). Station distance and apparent nasal breadth: a calibration for reconstructive practice. *Annals of Applied Facial Geometry*, 31(3):288–301.
- Marrable, D. and Osei, K. (2004). Mere exposure and the preference for the reversed self-portrait: a meta-analysis of thirty-one studies. *Review of Social Perception*, 8(1):44–79.
- Norwood, B. (1981). Which one is you? Sitter and observer disagree about photographic reversal. *Acta Psychologica Comparativa*, 47:129–140.
- Penhallow, M. and Iyer, S. (1996). *Shape Space and the Analysis of Landmark Configurations*. Second edition. Corvine University Press.
- Rademaker, T. (2009). Congeneric measurement models with instrument-specific latents: identification and the test of unit correlation. *Psychometric Methods Quarterly*, 22(2):155–183.
- Stanhope, L., Ferreira, J., and Okonjo, A. (2019). Fourteen years of matcher improvement and a stationary error floor. *Proceedings of the Conference on Applied Identification*, pages 77–92.
- Tolliver, E. (1974). Fluctuating asymmetry in the adult human face. *Human Biology Reports*, 6:1–19.
- Vane, P. and Duclos, M. (2021). Beam-splitter architectures for simultaneous multi-instrument capture. *Instrumentation Notes*, 55(9):640–657.