

Northumbria Research Link

Citation: Edmond, Gary and Wortley, Natalie (2023) Images, Investigators, Identification, Code D and the Court of Appeal. The Journal of Criminal Law. ISSN 0022-0183 (In Press)

Published by: Vathek Publishing

URL:

This version was downloaded from Northumbria Research Link:
<https://nrl.northumbria.ac.uk/id/eprint/51501/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

Images, investigators, identification, Code D and the Court of Appeal

Gary Edmond* and Natalie Wortley†

Abstract

The rapid rise in accessibility and portability of cameras has resulted in widespread reliance on the interpretation of images by analysts and investigators in criminal proceedings. Codes of practice, guidance and jurisprudence have evolved to facilitate the admission of opinions as to the identity of offenders (or persons of interest) at trial. In this article we explain why allowing investigators to give opinions as to identity on the basis of familiarity with images or suspects acquired during the course of an investigation is incompatible with mainstream scientific research and advice, and conducive to error. It rests on the flawed assumption that investigators can reliably identify or recognise persons in images, articulate and document the basis of these “identifications”, and avoid the risk of contamination (really cognitive bias) from their knowledge of, or exposure to, domain-irrelevant information. Jurors, who may be invited to conduct their own comparison between an image and the defendant in the dock, are similarly vulnerable to assuming the task is straightforward, as well as contextual and cognitive biases. Using the facts and evidence in *R v Yaryare* [2020] EWCA Crim 1314 as a case study, we show how case information available to investigators and imaging analysts both informs their interpretations of images and is (re-)presented at trial and on appeal as independent support for their opinions. We identify substantial threats to fairness, proof and rationality and propose that only witnesses with demonstrable expertise should be permitted to testify as to the identity of persons of interest in images.

Keywords:

Identification; images; expert evidence; CCTV; cognitive bias; Code D; opinion; Forensic Science Regulator; police familiars

1. Introduction: Out-PACED

In recent decades criminal justice systems have been confronted with more and more images – of crimes, preparation for crimes, and the circumstances surrounding crimes. The availability of images, and trials contested around the identity of offenders (or persons of interest) has, in England and Wales, led to widespread reliance on the interpretation of images by investigators and those recognised as experts. Police gradually developed guidelines, and courts jurisprudence, to regularise these interpretations and their admissibility.¹ The main “constraints” have been expectations that investigators will document identifications, endeavour to minimise suggestion and record derogation from norms embodied in the Police and Criminal Evidence Act 1984 (‘PACE’) Codes of Practice and *Practice Advice* issued to police officers.²

* Professor, Faculty of Law, UNSW, Sydney, Australia; Research Professor (fractional), Northumbria Law School; Fellow, Royal Society of New South Wales; Chair, Evidence-based Forensics Initiative.

† Associate Professor, Northumbria Law School.

¹ In 2011, *Codes of Practice* made under the Police and Criminal Evidence Act 1984 (PACE), were amended to include powers and procedures for presenting images to any person (including a police officer) for the purposes of obtaining evidence of recognition. Not strictly admissibility rules, Code D and parallel interventions by police chiefs have, intentionally, facilitated the admission of opinions. Section 67 of the Police and Criminal Evidence Act 1984 provides that the Codes of Practice “shall be admissible in evidence” and any provision of the Codes “shall be taken into account” if it appears to the court to be relevant to the determination of any question arising in the proceedings. Provisions of the Codes may therefore be relevant when a court is asked to consider the admissibility of identification or recognition evidence.

² Home Office, *Police and Criminal Evidence Act 1984 (PACE) – Code D: Revised Code of Practice for the Identification of Persons by Police Officers* (February 2017) (hereafter ‘Code D’); Association of Chief Police Officers, *Practice Advice on the Use of CCTV in Criminal Investigations* (National Policing Improvement Agency, 2011) (hereafter ‘Practice Advice’). The Association of Chief Police Officers of England, Wales and Northern Ireland (ACPO) was an independent strategic body which undertook responsibility for the direction and development of the police service. ACPO came under scrutiny when elected Police and Crime Commissioners were introduced nationally.

In recent years a plethora of national and international inquiries and reports have placed forensic science under “sustained scrutiny”,³ amid concerns about the role of questionable forensic methods and testimony in wrongful convictions. In England and Wales, such concerns presaged the introduction of a reliability test for expert evidence,⁴ and the creation of the office of the Forensic Science Regulator. In this context it is perhaps curious that advocates, trial judges and appellate courts apparently remain oblivious to the growing body of scientific literature in the fields of image analysis and identification. While some image analysts and even some police officers may be genuine experts, we contend that the reliability of opinions formed as a result of comparing and interpreting images ought to be subject to the same sort of critical scrutiny as scientific evidence.⁵

In this article, through close review of the evidence in the recent Court of Appeal decision of *R v Yaryare*,⁶ we show why simply allowing investigators to identify persons on the basis of “familiarity” with images and suspects acquired during the course of an investigation is unfair, conducive to error, and encourages irrationality. We draw upon mainstream scientific research and advice to explain why training and work as a police officer do not enhance the ability to accurately identify persons in images. We explain why repeatedly watching videos and studying images does not make investigators experts at face comparison or with particular sets of images. Rather, scientific research establishes that unfamiliar face comparison (or matching) is unexpectedly difficult and prone to error. Attentive scientists have repeatedly warned forensic scientists and courts about the need to employ validated methods and to carefully attend to subjective interpretations and error rates.⁷ In practice courts often disregard warnings and methodological advice, while simultaneously trivialising risks from cognitive bias on the part of police officers (and others) as a result of their participation in investigations or in suggestive viewings. The failure to manage these risks (or to recognise their magnitude) results in contaminated – i.e. biased – opinions. Investigating police officers and imaging analysts (whose evidence is admitted as a species of expertise) routinely form opinions about the identity of persons in images having been exposed to a great deal of suggestive information. This information is frequently then *re-presented* as independent support for interpretations that were informed – really *contaminated* – by it. Consequently, we will observe the same information – “strands” in evidence theory – repeatedly relied upon in the same criminal proceedings.

These substantial threats to fairness, proof and rationality are compounded where jurors are invited to make their own comparisons of images (or voices) in the incredibly suggestive conditions of the contested accusatorial trial. Through our analysis of the evidence in *Yaryare*, we will expose the cognitive dangers inherent in inviting jurors to compare images having been exposed to much of the information available to the identifying witnesses (here the investigating officer and an imaging analyst), while the defendant, said to be the person in the images, sat passively in the dock. Rather than some kind of test or confirmation, we explain why jury comparisons are actually cognitive traps.

Following the publication of an independent review (General Sir Nick Parker, *Independent Review of ACPO*, 2013) ACPO was replaced by the National Police Chiefs’ Council (NPCC) in 2015.

³ See e.g. National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (National Academies Press, 2009) 9 (hereafter ‘NRC Report’); PCAST, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature- Comparison Methods* (Executive Office of the President, September 2016) 6 (hereafter ‘PCAST Report’); Law Commission, *Expert Evidence in Criminal Proceedings in England and Wales* (Law Com No 325, 21 March 2011).

⁴ Crim PD Part 19A

⁵ See Tony Ward and Sharzad Fouladvand, ‘Bodies of knowledge and robes of expertise: expert evidence about drugs, gangs and human trafficking’ [2021] Crim LR 442.

⁶ *R v Yaryare and others* [2020] EWCA Crim 1314, [2020] 4 WLR 156 (hereafter ‘*Yaryare*’).

⁷ Much of the material discussed in this essay also applies to the comparison and recognition of voices. See David C Ormerod, ‘Sounds Familiar? Voice Identification Evidence’ [2001] Crim LR 595 and Gary Edmond, ‘Opinion, bias and double dipping’, paper presented at the Michele Taruffo Evidence Week (Girona, Spain) in June 2022.

We end by proposing a way forward for image comparison evidence that is grounded in science and cognisant of the need to facilitate the effective evaluation of testimony as to the identity of persons in images. The following analysis commences with a review of relevant jurisprudence and legislation, before using the case of *Yaryare*⁸ as a vehicle to explore their profound limitations and curious displacement from scientific research and advice.

2. Background: A-G's Ref (No 2 of 2002), PACE 1984 Code D, and the Practice Advice

In its influential review of extant authorities in *Attorney-General's Reference (No 2 of 2002)* (“A-G's Ref (No 2 of 2002)”),⁹ the Court of Appeal identified four situations in which it is permissible to admit evidence and invite a jury to conclude that a person in an image is the defendant:

- (i) where the photographic image is sufficiently clear, the jury can compare it with the defendant sitting in the dock ...
- (ii) where a witness knows the defendant sufficiently well to recognise him as the offender depicted in the photographic image, he can give evidence of this ...; and this maybe so even if the photographic image is no longer available for the jury ...;
- (iii) where a witness who does not know the defendant spends substantial time viewing and analysing photographic images from the scene, thereby acquiring special knowledge which the jury does not have, he can give evidence of identification based on a comparison between those images and a reasonably contemporary photograph of the defendant, provided that the images and the photograph are available to the jury ...;
- (iv) a suitably qualified expert with facial mapping skills can give opinion evidence of identification based on a comparison between images from the scene ... and a reasonably contemporary photograph of the defendant, provided the images and the photograph are available for the jury¹⁰

Although not presented as “an exhaustive list”,¹¹ the evidence of witnesses who purport to identify defendants from crime-related images typically falls into one of these categories. In this section we examine each in turn before turning to consider their applicability to the opinion evidence in *Yaryare*.

A. Jury comparisons: category (i)

It may seem surprising that a jury can be invited to compare images associated with a crime with the defendant in the dock (or with images of the defendant). After all, dock identifications by witnesses have long been regarded as highly unsatisfactory and impermissible, save in exceptional circumstances.¹² As long ago as 1976, the Devlin Committee acknowledged that dock identifications take place in “suggestive circumstances”, for “the root of all objections to dock identification is that it comes as an answer to ... a leading question”.¹³ The Committee’s disquiet foreshadowed contemporary concerns about cognitive bias. For example, the Committee noted that “if ... there is any sort of similarity between the man [the witness] saw and the man in the dock, [the witness]

⁸ (n 6)

⁹ [2002] EWCA Crim 2373, [2003] 1 Cr App R 21.

¹⁰ *A-G's Ref (No 2 of 2002)* (ibid) [19]. As we shall see, with respect to (iii) and (iv), real problems flow from provision of images to the trier of fact.

¹¹ *R v Ozger* [2022] EWCA Crim 1238.

¹² *R v Long* [2022] EWCA Crim 444.

¹³ Departmental Committee on Evidence of Identification in Criminal Cases, *Report to the Secretary of State for the Home Department* (‘the Devlin Report’) (HC 338, 1975-76), 2.24. The Devlin Committee was established by the Home Secretary following a series of wrongful convictions in cases involving identification evidence, including at least one contested dock identification.

naturally tends to identify the man in the dock as the criminal”.¹⁴ Similarly, a witness who lacks confidence in their ability to identify the accused “may feel differently when he finds that all he is required to do in court is to back the choice the police have already made.”¹⁵

Following publication of the Devlin Report, the Attorney-General and the Director of Public Prosecutions confirmed to Parliament that, in cases tried on indictment, the prosecution would “not invite a witness ... to make a dock identification unless the witness’s attendance at [an identification] parade was unnecessary or impracticable, or there are exceptional circumstances.”¹⁶ This statement was incorporated into, and remains part of, Crown Prosecution Service guidance on identification evidence, effectively consigning dock identifications to history.¹⁷

Modern jurisprudence acknowledges that identification by witnesses is a difficult task and that an honest witness may be mistaken. Since the decision in *R v Turnbull*,¹⁸ judges have been required to direct the jury in prescribed terms about the dangers of mistaken identification and the “special need for caution” whenever the prosecution case depends wholly or substantially on disputed visual identification evidence.¹⁹ Such warnings extend to cases in which the defendant is known to the witness, “because many people [have] experienced seeing someone in the street whom they [know], only to discover they were wrong”, giving rise to the well-known expression “I could have sworn it was you”.²⁰ Yet decisions of the appellate courts have been inconsistent as to whether a modified *Turnbull* direction is required where a jury is invited to identify the defendant from an image.

In *R v Dodson & Williams*, the Court of Appeal thought it “of utmost importance” that, when a jury is presented with images and invited to make its own identification of the defendant, the trial judge should provide guidance referencing matters such as “the quality of the photographs, the extent of the exposure, the facial features of the person photographed, evidence, or the absence of it, of a change in the defendant’s appearance and the opportunity a jury has to look at the defendant in the dock and over what period of time”.²¹ Similarly, the judgment in *R v Blenkinsop* referred to “a general and invariable requirement that the jury should be warned of the risk of mistaken identification and of the need to exercise particular care in any identification which they make for themselves”.²² This passage was cited with approval in *R v Ali*, where the appellant’s conviction was quashed because the judge had omitted to give “a careful and full direction as to the dangers of mistaken identification when he invited the jury to ‘use [their] own eyes’ when comparing the CCTV images with the appellant.”²³ Conversely, in *R v Downey* (an appeal heard a matter of days before *Blenkinsop*, by a court comprising two of the same judges), it was suggested that “[a] mandatory direction is justified in cases of identification by a

¹⁴ Devlin Report (ibid) 2.24.

¹⁵ Devlin Report (ibid) 4.92.

¹⁶ Hansard 912 Thursday 27 May 1976. Dock identifications have been permitted in summary trials for minor offences where the defendant failed to indicate, prior to trial, that identification was in issue. Otherwise, it is difficult to envisage any “exceptional circumstances” that could justify a dock identification now that it is possible for the police to covertly obtain images of a non-compliant suspect for the purposes of holding an identification procedure.

¹⁷ *Identification: Legal Guidance* (CPS, 2018) < <https://www.cps.gov.uk/legal-guidance/identification> > accessed 7 August 2022. Curiously, concerns about the suggestive propensities of dock identifications are not reflected in jurisprudence on the identification of defendants by jurors. We return to this issue below.

¹⁸ [1977] QB 224 (CA).

¹⁹ *R v Turnbull* (ibid); Judicial College, *Crown Court Compendium Part I* (June 2022) 15-1 < <https://www.judiciary.uk/wp-content/uploads/2022/07/Crown-Court-Compendium-Part-I-June-2022.pdf> > accessed 7 August 2022.

²⁰ *R v Bentley* [1991] Crim LR 620, CA. Google n-grams record a massive increase in the use of ‘doppelganger’ in recent decades.

²¹ [1984] 1 WLR 971 (CA) 979. Their Lordships declined to provide a formula upon which judicial directions should be based.

²² [1995] 1 Cr App R 7 (CA) 12.

²³ [2008] EWCA Crim 1522 [41], [2009] Crim LR 40.

witness ... [but] ... inviting the jury to consider whether the person shown in a photograph is the defendant who has appeared before them is a different process”.²⁴ This time the Court concluded that a judicial warning had not been necessary because “[t]o some extent the difficulties are obvious to any layman”.²⁵ The Court of Appeal purported to resolve this conflict in the authorities in *R v Shanmugarajah*,²⁶ preferring the judgment in *Downey* on the basis that comparing images to a defendant in the dock “is a perfectly straightforward [task], and it is a statement of the obvious to direct the jury that they must be careful and that they must bear very much in mind all the kind of warnings that [the defence advocate] has insisted should be given.”²⁷

While we would not wish to overstate the value of jury directions, judgments that acknowledge that juries ought to be explicitly warned of the need for caution are at least cognisant of some of the risks and dangers inherent in this type of evidence. To suggest that something can be omitted from judicial directions because it is so obvious that it does not need to be said embodies an apparently prevalent misunderstanding of the difficulty of image comparison (and recognition) and cognitive biases (discussed in more detail, below).²⁸ In *Shanmugarajah*, the Court suggested that defence counsel’s ability to question witnesses and to highlight the absence of an identification procedure in his closing speech were adequate safeguards.²⁹ Even if these aspects of the trial process were capable of affording satisfactory protection for defendants, it is unclear how cross-examination of a witness could protect against mistaken identification by jurors. And while it is true in theory that, as the judge pointed out in summing up, an identification procedure could have favoured either party,³⁰ the Crown is unlikely to prosecute those who do not resemble persons of interest in crime-related images. It is our contention that neither cross-examination nor jury directions are sufficient to repair the risks created by inviting a jury to engage in image comparison in the sort of conditions that worried the Devlin Committee and would be deprecated by attentive contemporary scientists.³¹

B. Recognition from prior contact: category (ii)

Although the terms “identification” and “recognition” are commonly used interchangeably, they are different processes, both legally and cognitively.³² Under Code D, recognition occurs where a person who is not an eyewitness to the alleged offence claims to know the

²⁴ [1995] 1 Cr App R 547 (CA) 555-556. Note that scientific studies have demonstrated that having an actual person standing in front of participants does not improve their ability to determine whether that person is the person in high quality images (in contrast to the comparison of two images).

²⁵ *ibid* 556.

²⁶ [2015] EWCA Crim 783, [2015] 2 Cr App R 14.

²⁷ *ibid* [29].

²⁸ See e.g. the Australian case of *R v Dickman* [2017] HCA 24. On directions and warnings, see EBFI, ‘A warning about judicial warnings’ (2023) 44 *Adelaide Law Review* (forthcoming).

²⁹ (n 26) [39].

³⁰ Empirical studies have consistently shown that where a witness is invited to participate in consecutive identification procedures, the identification they made in the first procedure will almost invariably be reproduced in the second, regardless of the accuracy of the first identification. See T Valentine, J Davis, A Memon and A Roberts, ‘Live Showups and their Influence on a Subsequent Video Line-up’ (2012) 26 *Applied Cognitive Psychology* 1.

³¹ The absurdity of permitting this form of “evidence” is apparent when one considers that a defendant facing the prospect of a jury comparing him to crime scene images might find himself in a better position, with respect to proof, if he fails to attend his trial. A jury is precluded from drawing an inference from a defendant’s absence (*Compendium* (n 15) 3-3), and would then be unable to make any comparison with the defendant in the dock (although admittedly the prosecution might then seek to adduce contemporary images of the defendant, such as custody suite images, for comparison purposes – see *R v Ozger* (n 11)).

³² *Practice Advice* (n 2). *Identification* is a formal legal process carried out in accordance with established procedures (most commonly a video identification procedure) to test whether a witness can identify a person suspected of involvement in the crime under investigation as the person the witness saw on a previous occasion. However, the term “identification” is often used to describe the results of recognition or comparison or, as in the case of *Yaryare*, attribution following a synthesis of all of the evidence.

identity of a person of interest in images based on some previous contact.³³ A broad range of indirect witnesses (i.e. non-eyewitnesses) may be permitted to testify that they recognise a person depicted in a crime scene image based on prior familiarity, just as an eyewitness may claim to recognise the person they saw committing an offence. An eyewitness will typically be required to participate in a formal identification procedure before their evidence can be adduced (unless it is accepted that the suspect knows the defendant sufficiently well that any such procedure would serve no useful purpose).³⁴ In cases involving recognition from CCTV images, appellate judgments have tended to focus on the potential to test the purported recognition before a jury.

In 2011, Code D introduced procedures for presenting images to potential witnesses for the purposes of obtaining *recognition* evidence.³⁵ Contemporaneously, the Association of Chief Police Officers (ACPO) issued *Practice Advice* on the use of CCTV in criminal investigations.³⁶ Both the 2011 iteration of Code D and the *Practice Advice* emphasise the importance of documenting the circumstances and conditions in which an individual is shown images and given an opportunity to recognise a person depicted therein. The Code and the *Practice Advice* make clear that these safeguards are equally important where the individual is a police officer.³⁷ In addition, the *Practice Advice* draws attention to the possibility that “the ... value of evidence of recognition ... may be compromised if, before the person is recognised, the witness who has claimed to know them is given[,] or is made or becomes aware of, information about the person which was not previously known to them personally but which they have purported to rely on to support their claim that the person is in fact known to them”.³⁸

In *R v Smith*, the Court of Appeal called for further amendments to Code D to take account of case law, developments in police practice, and operational concerns arising out of the increasingly widespread use of CCTV in criminal investigations.³⁹ The current version of Code D is somewhat clearer in that it delineates procedures for three main types of identification: eye-witness identification;⁴⁰ the controlled showing of images in the hope of obtaining evidence of recognition based on (prior) familiarity;⁴¹ and, the uncontrolled circulation of images to the public (e.g. via national or local media, or social media) and/or via police internal communication systems for the purposes of identifying and tracing suspects.⁴² However, the Code remains silent in relation to those who fall into category (iii) from *A-G’s Ref (No 2 of 2002)* – namely, the investigator who purports to be able to recognise persons using “specialised knowledge” acquired through “viewing and analysing ... images”.⁴³

Viewings carried out for the purposes of obtaining recognition evidence may be categorised as either “controlled” or “uncontrolled”. The former occurs whenever formal arrangements are made to show video or photographs to persons (including police officers)

³³ This is different from claims of recognition (or identification) made by eyewitnesses who know the person of interest. The categories are messy and often overlap.

³⁴ Code D (n 2) 3.12

³⁵ The 2011 edition of Code D was in force at the time of the investigation in *Yaryare* (n 6).

³⁶ (n 2).

³⁷ Code D (n 2) Notes for Guidance 3A; *Practice Advice* (n 2) 7.4.4. See also *R v Smith* [2008] EWCA Crim 1342, [2009] 1 Cr App R 36.

³⁸ *Practice Advice* (n 2) 7.4.4.

³⁹ *R v Smith* (n 37). See also *R v DS* [2014] EWCA Crim 933; Home Office, *2016 Consultation on Revising PACE Codes C, D and H: Covering Letter to Statutory Consultees* (22 March 2016)

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/510679/CONSULTATION-2016_CodesCDH_CoverLetter_22-03-2016_.pdf accessed 7 August 2022.

⁴⁰ Code D (n 2) Part A.

⁴¹ Code D (n 2) Part B.

⁴² Code D (n 2) Part C.

⁴³ *A-G’s Ref (No 2 of 2002)* (n 9).

other than eyewitnesses, in order to ascertain whether they recognise a person of interest. In response to concerns expressed by the Court of Appeal in *R v Smith*,⁴⁴ Code D now requires images to be shown on an individual basis, and a record must be made of the circumstances and conditions of the viewing and any recognition, including the reason for the recognition, the features that triggered it, the words of recognition and any expressions of doubt.⁴⁵ Implicitly, documentation is regarded as affording an “objective means of testing the accuracy” of an assertion of recognition,⁴⁶ but we contend that this is to overestimate the utility (and detail) of many records as well as the value of cross-examination as a technique capable of exploring cognitive – i.e. interpretive – processes. A particular difficulty with the recording requirements is that, while an expert might be expected to explain their methodology and accuracy to the jury, a lay witness or a police officer may be incapable of reliably explaining the basis for their opinion.⁴⁷ This may result in the creation of *ex post facto* rationalisations that do not capture the actual reasons for recognition.⁴⁸

Where an individual purports to recognise a person from an image viewed in *uncontrolled* circumstances, such as a police officer coming across an image “by accident”,⁴⁹ the circumstances of the recognition and conditions of the viewing should be documented “as soon as practicable”.⁵⁰ However, the Court of Appeal has endorsed the admission of recognition evidence despite fundamental failures to keep records, holding that “slavish adherence” to Code D is not required.⁵¹ The expectation is that jurors can be told of the breach and left to assess its impact on the reliability of the recognition evidence.

Similarly, where a witness (usually a police officer) is made aware that the suspect/defendant is believed to appear in images prior to their viewing,⁵² the Court of Appeal has held that such influences – dangers from suggestion, expectation and confirmation – can be obviated by simply reminding the jury that the officer was expecting to see the defendant when viewing the image.⁵³ If the jury is satisfied that the image is of sufficient quality, then jurors may make their own comparison between the image and the defendant.⁵⁴ Neither Code D nor the appellate courts have taken meaningful steps to manage the implication of suggested criminality or issues of cognitive bias and other forms

⁴⁴ *R v Smith* (n 37).

⁴⁵ Code D (n 2) 3.35-3.36.

⁴⁶ *R v Smith* (n 37) [67].

⁴⁷ Facial recognition is believed to be holistic in nature and studies have shown that faces are processed as integrated wholes, rather than collections of individual facial features. See James W Tanaka and Diana Simonyi, ‘The “Parts and Wholes” of Face Recognition: A Review of the Literature’ (2016) 69 Quarterly Journal of Experimental Psychology 1876. Studies subjecting police officers to face recognition tasks have shown that officers articulating a greater quantity of information about their facial recognition “strategies” often performed less well in empirical face recognition tests - see Josh P Davis, Ashok Jansari and Karen Lander, ‘I Never Forget a Face!’ (2013) 26(1) The Psychologist 726, 728.

⁴⁸ Gary Edmond and Natalie Wortley, ‘Interpreting Image Evidence: Facial mapping, police familiars and super-recognisers in England and Australia’ (2016) 3(2) Journal of International and Comparative Law 473. Where identifications are fabrications, it can be exceedingly difficult to expose such deception.

⁴⁹ As occurred in e.g. *R v Moss* [2011] EWCA Crim 252, [2011] Crim LR 560.

⁵⁰ Code D (n 2) 3.41. For additional guidance where there is an identification or recognition from social media, see Association of Chief Police Officers of England and Wales (ACPO) National Visual and Voice Identification Strategy Group, *Internet Social Media and Identification Procedures* (ACPO, July 2014).

⁵¹ *R v Moss* (n 49). See also *R v Alexander & McGill* [2012] EWCA Crim 2768, [2013] 1 Cr App R 26; *R v Lariba and others* [2015] EWCA Crim 478, [2015] Crim LR 534; *R v Byrne* [2016] EWCA Crim 2124, [2017] Crim LR 801; *R v Crampton* [2020] EWCA Crim 1334, [2021] Crim LR 309. The exception is *R v JD* [2020] EWCA Crim 1314, in which the Court of Appeal quashed the conviction of a youth on the basis of “a wholesale failure to comply with the requirements of Code D” (at [28]).

⁵² See e.g. *R v Simpson* [2019] EWCA Crim 1144.

⁵³ *R v Chaney* [2009] EWCA Crim 21, [2009] 1 Crim App R 35.

⁵⁴ *R v Tucker* [2008] EWCA Crim 3063. Although, a warning of the danger of “confirmation bias” may be appropriate: *R v Dawes* [2021] EWCA Crim 760.

of prejudice that may arise when an officer purports to have recognised a suspect as a result of the suspect's previous involvement with the criminal justice system.⁵⁵

As with jurors comparing images to the defendant in the dock, it appears that a suitably adapted *Turnbull* warning ought to be given whenever a jury is considering the evidence of a witness who purports to have recognised a defendant in an image and a judge ought to withdraw the evidence from the jury if its reliability cannot properly be assessed, for example, because the quality of the photograph or recording is poor.⁵⁶ However, once again, the Court of Appeal has tended to uphold convictions in cases where the quality of the images and the significance of any prior knowledge the witness had of the defendant were left to the jury to assess.⁵⁷

C. Non-expert image comparison: category (iii)

Identifications falling into the third category from *A-G's Ref (No 2 of 2002)* are said to be admissible on the basis that "special knowledge" can be obtained by the repeated viewing *and analysis* of images, usually by police officers as part of their investigation into an offence. The reference to "analysis" and "special knowledge" is suggestive of an attempt to address one of the core criteria for the admission of expert evidence, where there is an expectation that the witness possesses an advantage over the jury.⁵⁸ Although not engaging in comparative analysis in a technical sense, these indirect witnesses are sometimes referred to as experts *ad hoc*.⁵⁹ Their opinions might be treated as identification or recognition (based on repeat viewings),⁶⁰ and the latter might even be treated as evidence of fact, such that exclusionary opinion rules are not applied (though this approach sits awkwardly with *A-G's Ref (No 2 of 2002)*). Importantly, there is little empirical evidence to support the proposition that extended exposure to CCTV or other images is capable of producing "special knowledge" or significant improvements in performance relative to others, such as jurors.

Revealingly, the Forensic Science Regulator does not consider the study of images by so-called super-recognisers to be a species of forensic science for three reasons:

- (a) the work is generally carried out within an operational policing unit, with no separation to ensure independence and impartiality;
- (b) photographs of known suspects or offenders are studied prior to watching the footage containing unknown individuals, without implementing safeguards against cognitive bias; and
- (c) ... attempting to match faces from photographs against CCTV footage is not based on scientifically validated methodology, nor are error rates known.⁶¹

These considerations would seem to apply where it is claimed that an investigator has acquired special knowledge through extended or repeated viewing. Such *ad hoc* expertise, which is not really expertise at all, is therefore outside the remit of the Forensic Science

⁵⁵ Edmond and Wortley (n 48) 477.

⁵⁶ *R v Ali* (n 23).

⁵⁷ *R v Doherty* [2016] EWCA Crim 246; *R v DS* (n 39).

⁵⁸ Expert evidence is only admissible if it concerns a matter that would be outside the knowledge of the jury: *R v Turner* [1975] QB 834 (CA). Earlier, Munday suggested that "the matters on which [such] witnesses are testifying are matters which the jury, given time and patience, could judge for themselves": Roderick Munday, 'Videotape Evidence and the Advent of the Expert *Ad Hoc*' (1995) 159(33) JP 547.

⁵⁹ But see *R v Flynn* [2008] 2 Cr App R 20, [14] and Gary Edmond and Mehera San Roque, 'Quasi-Justice: Ad hoc experts and identification evidence' (2009) 32 Criminal Law Journal 8.

⁶⁰ When it is suggested that repeated viewings result in an investigator developing the ability to *recognise* a suspect, categories (ii) and (iii) of *A-G's Ref (No 2 of 2002)* (n 9) appear to overlap. Therefore, the warning in the *Practice Advice* that the value of recognition evidence may be compromised if a witness becomes aware of information about the suspect that was not previously known to them (see section 4B, below) ought also to apply in category (iii) cases.

⁶¹ Dr Gillian Tully, *Annual Report November 2016-2017* (Forensic Science Regulator, 19 January 2018), para 1.8. See Edmond and Wortley (n 48) and the discussion below in section 5F.

Regulator. In practice, while often referring to such witnesses as “expert” (or a species of expert), the courts have treated their testimony as non-expert evidence, with the consequence that none of the safeguards regulating the introduction of expert opinion are applied. In particular, there is no need for the witness to demonstrate any expertise in the comparison of images,⁶² or for the prosecution to show that there is a sufficiently reliable basis for admitting their evidence.⁶³ Judges are not required to consider the factors set out in the Criminal Practice Direction to which a court ought to have regard when determining the admissibility of expert evidence, such as: the validity of the method used by the officer; whether she took account of all relevant information; whether her examination, technique, method or process was properly carried out or applied; and, whether her evidence relies on an inference or conclusion that has not been properly reached.⁶⁴ There is no assessment of the witness’ actual ability to perform the task to which their testimony relates. Furthermore, almost all of the reported cases involve investigators who were already aware, or became aware during the course of their repeated viewings, that the defendant was a suspect, so their “special knowledge” was acquired in highly suggestive circumstances.

Code D is said to be applicable to image comparisons by all police officers, but investigators routinely fail to keep the sort of records envisaged by the Code and indeed, in some cases (including *Yaryare*), fail to make any records of their identifications or “recognitions”. This may be a consequence of the piecemeal and protracted nature of many investigations and the nature of the cognitive activity involved, which makes it difficult to determine (let alone to record) a precise moment at which familiarity with one or more images crystallises into a belief that it depicts a particular individual. Investigators may also consult or liaise with other police officers and/or experts (as in *Yaryare*), thereby breaching the requirements in Code D and the *Practice Advice* that images should be shown on an individual basis and without suggestion.⁶⁵ Both trial and appellate courts routinely excuse derogation from the provisions of Code D, rendering its provisions otiose in category (iii) cases.⁶⁶

D. Expert image comparison: category (iv)

According to the Forensic Science Regulator (FSR), forensic scientists must possess demonstrable expertise “acquired through a documented process of training and testing”, such that their opinions are capable of assisting the jury.⁶⁷ They must utilise validated methods in their analysis and any limitations should be identified and disclosed. With respect to imaging analysis and comparison, the FSR has acknowledged that “[t]here has been very limited scientific investigation into whether or not forensic facial image comparison is a reliable form of identification evidence”, describing the techniques and processes involved as “still fairly novel”.⁶⁸

⁶² *R v Turner* (n 58); *R v Stockwell* (1993) 97 Cr App R 260 (CA).

⁶³ *R v Dlugosz and Others* [2013] EWCA Crim 2, [2013] 1 Cr App R 32; Criminal Procedure Rules 2020 (hereafter CrimPR) 19.4(h).

⁶⁴ CrimPD 19A.5-6. The extent to which these factors, which derived from the Law Commission’s report, *Expert Evidence in Criminal Proceedings in England and Wales* (n 3) have in practice heralded “a more rigorous approach” to expert evidence is doubtful given the dearth of case law concerning their application.

⁶⁵ In practice, therefore, some officers appear to regard Code D as inapplicable to investigators who are to be called to give an opinion as to the identity of an individual in an image under the auspices of category (iii).

⁶⁶ *R v Moss* (n 49); *R v Alexander & McGill* (n 51); *R v Lariba and others* (n 51); *R v Byrne* (n 51); *R v Crampton* (n 51); cf. *R v JD* (n 51).

⁶⁷ *Forensic Image Comparison and Interpretation Evidence: Guidance for Prosecutors and Investigators* (Forensic Science Regulator, National Crime Agency, CPS & Metropolitan Police, 2nd February 2016) p 3

⁶⁸ *ibid* p 8. Though, see Gary Edmond et al, ‘Facial recognition and image comparison evidence: Identification by investigators, familiars, super-recognisers and algorithms’ (2021) 45 Melbourne University Law Review 99.

In addition, in the 2019 Annual Report, the FSR expressed concerns that “image analysis experts are, on occasion, failing to stay within the bounds of their expertise and failing to communicate effectively to the courts the limitations of work carried out”.⁶⁹ She emphasised that “[e]xpertise in CCTV, video, imaging or enhancement does not equate to expertise on the content of the image”. Guidelines issued by way of a Regulatory Notice in July 2019 effectively preclude an imaging expert whose expertise is based upon qualifications and/or experience in image quality, compression and enhancement, from expressing an opinion as to whether a person shown in a crime scene image is the same person depicted in a reference image of a suspect.⁷⁰

While the FSR has not sought to proscribe the use of facial mapping techniques, the previous Regulator cautioned that “facial image comparison evidence based on poor quality video or images (such as CCTV) should be used as supporting evidence only”.⁷¹ The Court of Appeal’s willingness to allow the “subjective opinions” of facial mappers and image analysts to be presented as “conventional expressions arranged in a hierarchy”, from “lends no support” through to “lends powerful support”,⁷² is a further source of controversy.⁷³

Notwithstanding early concerns over the validity and reliability of facial mapping evidence and its presentation,⁷⁴ and the Forensic Science Regulator’s observations, none of the relevant appellate judgments refers to a defence advocate attempting to mount a scientifically-informed reliability challenge or referencing the Regulator’s concerns. The introduction of the Law Commission’s reliability standard for expert evidence therefore does not appear to have resulted in any scientifically-informed challenges to facial mapping (or imaging analysis) as a species of expertise, to the application of a particular practitioner’s methods, or to the conclusions reached in any individual case.

Our case study considers the treatment of the opinions of the primary investigator (DC B), the opinions of an imaging analyst (P), and an invitation to jurors to compare the images themselves. It therefore enables us to explore the operation of Code D in the context of *A-G’s Ref (No 2 of 2002)* categories (i), (iii) and (iv), as well as highlighting hitherto unrecognised problems with the interpretation of CCTV and other images.

3. *R v Yaryare and others* [2020] EWCA Crim 1314

In 2018, a jury convicted Liban Yaryare of two counts of conspiracy to commit violent disorder and one count of attempted murder. Yaryare’s was a joint trial (with Hassan and Osman), itself part of a larger set of prosecutions arising from a series of gang related incidents that occurred in Leicester on the 30 April 2015. The following discussion primarily focuses on Yaryare’s trial and appeal. For our purposes the important features of the case concern Yaryare’s alleged participation in a series of altercations between two groups at three locations, culminating in a person being stabbed 11 times (the attempted murder).⁷⁵ Apart from the testimony and condition of the victim (which supported the

⁶⁹ Forensic Science Regulator, *Annual Report* (15 March 2019).

⁷⁰ Forensic Science Regulator, *Image Enhancement and Image Comparison: Provision of Opinion* (FSR Regulatory Notice 01/2019, 17 July 2019)

⁷¹ *Guidance for Prosecutors and Investigators* (n 67)

⁷² *R v Atkins* [2009] EWCA Crim 1876, [2010] 1 Cr App R 8

⁷³ It should be made clear to the jury that this sliding scale is not based upon a database of facial or body features but this terminology nevertheless creates the impression that there is some sort of empirical basis for the strength of the expert’s conclusions. See Gary Edmond et al, ‘*Atkins v The Emperor: The “Cautious” Use of Unreliable “Expert” Opinion*’ (2010) 14 *International Journal of Evidence and Proof* 146.

⁷⁴ *R v Gray* [2003] EWCA Crim 1001.

⁷⁵ One group was said to be composed of ‘men mainly of Somali descent’ and the other ‘of men of either African or Asian heritage’: *Yaryare* (n 6) [11]-[13]. See also the discussion at [76] and [80], citing *R v Chaney* (n 53) [14], [21]-[25] and *R v JD* (n 51) [40].

assault), the prosecution cases against Yaryare and his co-accused were based upon “four strands of evidence”.⁷⁶

First, given each of the incidents was captured by CCTV recordings, *recognition evidence* from the officer in the case, DC [B]. Second, evidence from a facial mapping expert, [P], *which provided support for DC [B]’s identifications*. Third, evidence showing the movements of some of the relevant mobile telephones and the details of contact between some of those involved. Fourth, various exhibits seized from the respective defendants.⁷⁷

There was no mobile telephone location evidence against Yaryare,⁷⁸ but security staff at a local college testified that he “had an association with a number of the offenders”, some of whom appeared in his mobile telephone contacts list.⁷⁹ Yaryare was also said to have “lied in his interview under caution about his friendship and association with Hudur,” (who appears to have been convicted in relation to the same events in separate proceedings).⁸⁰ At trial and on appeal, Yaryare and his co-defendants/appellants protested that they had been incorrectly identified.

The primary evidence implicating Yaryare was DC B’s identification. She had obtained images of Yaryare from his Facebook page, his visit to the police station, and his time in custody for comparison purposes.⁸¹ She also interviewed Yaryare and reviewed the CCTV with him.⁸² On this basis she was said to have “had extensive opportunities to study his appearance”, enabling her to compare Yaryare with “a number of sightings of person A on the footage which provided a clear basis for making comparisons”.⁸³ DC B gave what the Court of Appeal described as “unequivocal evidence” identifying Yaryare as person A. The “imaging analyst”, P, “found ‘strong support’ for the suggestion that Yaryare and person A were the same individual.”⁸⁴ The jurors were also encouraged to undertake their own comparisons of the persons of interest in the CCTV with images of Yaryare (who had been sitting in court for the duration of the trial).⁸⁵ Without the identification evidence of DC B and P, the case against Yaryare was largely dependent on his associations, a single lie to investigators about his friendship with Hudur, and whatever the jurors might make of the images.⁸⁶

⁷⁶ (n 6) [24]. Use of the strands metaphor is routine in circumstantial cases, where the strands are said to be susceptible to combination, making a stronger cable (or case). As we shall see, it is important that the strands are kept independent – free from cross-contamination.

⁷⁷ (n 6) [24] (italics added). This version understates the centrality of DC B and her identification. Movement based on mobile phone records was not adduced against Yaryare. And, no clothes resembling those worn by ‘person A’ (said to be Yaryare) were recovered from a search of Yaryare’s property.

⁷⁸ Although, it is presented as ‘fact’ that Yaryare was at one stage in a BMW at one of the scenes, (n 6) [18], see also [15], [17], [19], [22]. Yaryare was implicated through circumstantial evidence and the opinions of other police officers, before he was identified by DC B and P. DC B worked with “the telephone analyst” [32] as well as the imaging analyst.

⁷⁹ (n 6) [63].

⁸⁰ (n 6) [63]. The Court of Appeal states that ‘[i]n fact Hudur had been with the appellant at the time of the offences and he was convicted by the jury.’ This seems to have been based on Hudur’s conviction, because it was in issue in Yaryare’s trial and appeal.

⁸¹ (n 6) [33].

⁸² Other suspects, such as Hasan, were not shown the CCTV during their interview because ‘he had not been identified by that stage’: (n 6) [34], [50].

⁸³ (n 6) [62].

⁸⁴ (n 6) [62], [64]. P’s evidence was weaker with respect to Hasan, where he could only provide ‘limited support’, [65]. An imaging analyst (Z) called by the defence, in contrast, “found . . . only “some support”” for that conclusion. It might be considered revealing that the defence expert provides a weaker conclusion rather than the kind of critique that familiarity with the relevant scientific research might have enabled.

⁸⁵ We have used the detective’s initial to provide a semblance of anonymity. However, it is important to make clear that our analysis is aimed at the limitations of Code D and jurisprudence concerning the use of images and the interpretation of images. We have no reason to believe that DC B was doing anything that was insincere. Indeed, the difficulty of effectively challenging the sincere opinions of a police officer appears to be a significant part of the problem, as is the inability to determine sincerity. See *R v Smith* (n 37) [67].

⁸⁶ We are critical of jury comparisons for reasons developed below.

Yaryare's counsel raised three grounds of appeal. First, he challenged "the judge's refusal to exclude the recognition evidence of DC [B] pursuant to section 78 Police and Criminal Evidence Act".⁸⁷ Second, he appealed "the judge's refusal to direct the jury as to the dangers identified by the Court of Appeal in *R v Smith*; and to correct a false impression created by DC [B] as to the lack of an obligation to take notes" as required by Code D and the ACPO *Practice Advice*.⁸⁸ Third, with the exclusion of the identification evidence there was said to be no case to answer. Our analysis focuses on grounds one and two, but with obvious implications for the no case submission and the cases against Yaryare's co-defendants, who were also identified by DC B and P.

Looking more closely at the image evidence, DC B explained that, as the investigation progressed and having "studied" the CCTV recordings, she recognised several of the suspects from the video of events on 30 April 2015.⁸⁹ The appellant accepted "that DC [B], as a matter of law, was entitled ... to give evidence *as an expert* on the CCTV footage."⁹⁰ Rather, the complaint was that she had not complied with Code D or the *Practice Advice*. In particular, she provided her initial witness statement "setting out why she had arrived at her conclusions as to the identification of the suspects considerably after she commenced viewing the CCTV footage and she had failed to maintain a satisfactory log".⁹¹

DC B made three written statements in total. In each she referred to viewing the CCTV images on numerous occasions over the course of the three year investigation. She was unable to quantify the number of hours but said she had watched the videos "hundreds of times" and over "hundreds of hours".⁹² She described focussing on the appearance, actions and movements of particular individuals with whom she had "become familiar" and, as a result, she claimed to be able to recognise and identify them as they moved from one video to another. DC B also referred to "receiving assistance from", "liaising with" and "working with" an imaging analyst.⁹³ The reasons DC B offered for identifying Yaryare as a person of interest (Person A) in the CCTV images were his "height, age, ethnicity, walk and posture ('stands with his shoulders rounded and his hips forward'), hairline, shape of his lips, ear position and skin colour"; in testimony she "referred also to the shape of his eyes."⁹⁴ As a result, she was "satisfied" that she had "identified" him at the three key locations.⁹⁵

Following applications to exclude her evidence at trial, DC B produced 21 investigation workbooks and 3 interviewing workbooks.⁹⁶ Covering the entire investigation, entries were described by the Court of Appeal as "hasty ... far from complete ... a handwritten record of particular matters that attracted the officer's attention". There was no record as to how individuals had become suspects or when she had "recognised" particular individuals in the video. DC B "indicated in cross-examination that she relied to a significant extent on her memory ('it's all up here')".⁹⁷ The first documented reference to Yaryare appears to have been in a statement dated almost two years after the investigation commenced.

⁸⁷ (n 6) [6]. Section 78 of the Police and Criminal Evidence Act 1984 enables a court to exclude prosecution evidence if it would be unfair to admit it.

⁸⁸ (n 6) [6] references deleted. See *R v Smith* (n 32).

⁸⁹ (n 6) [26].

⁹⁰ On the basis of category (iii): (n 6) [26]. Although, this is a curious expression. What does it mean to be an expert on "the footage"? More importantly, does this bear on the ability to accurately identify persons in the images? Elsewhere the judge refers to DC B's "familiarity with the materials" (rather than the persons), [45] para. 9 (trial judge).

⁹¹ (n 6) [27].

⁹² (n 6) [26], [69].

⁹³ (n 6) [28], [29], [32].

⁹⁴ (n 6) [33].

⁹⁵ (n 6) [29].

⁹⁶ (n 6) [31].

⁹⁷ (n 6) [36].

The Court of Appeal considered that the circumstances in which DC B's identification took place were "broadly in line with" *A-G's Ref (No 2 of 2002)* category (iii). The trial judge had regarded it as "unarguable" that DC B had "'acquired special knowledge' as envisaged in the authorities". Accordingly, her "familiarity with the materials" allowed her to "give 'expert' evidence of this nature". The defendants' concerns about DC B's identifications and record keeping were for the trial process and the jury.⁹⁸ Dismissing the defence advocate's criticism of the records as a "counsel of perfection", the trial judge indicated that the "mischief identified ... in previous cases can easily be met ... with guidance given to the jury in the usual way." That is, with "[a]ppropriate directions" the "defence can have a fair trial."⁹⁹ On appeal, Yaryare's counsel submitted that it was "unclear whether this was an 'identification' or a 'recognition' case,"¹⁰⁰ but that, in any event, the trial judge should have directed the jury as to the need for detailed record keeping "in accordance with *Smith*".¹⁰¹ He also highlighted the fact that "two other officers who had viewed a still from the CCTV footage had suggested to DC [B] that [person] A might be Yaryare, prior to her viewing".¹⁰² As to whether there had been a case to answer, he raised the "marked absence of other evidence" implicating Yaryare,¹⁰³

Counsel for the respondent submitted that this was "not a case where Code D strictly applied", as DC B was never instructed to watch the CCTV in order to identify suspects. He conceded that it would have been "preferable" for DC B to have documented her *recognition* once it "crystalised" during the investigation, but submitted the "failure to do so did not render her recognition unreliable."¹⁰⁴ Similarly, he argued that the trial judge's failure to draw the non-compliance with Code D clearly to the attention of the jury, or to correct DC B's misleading testimony which implied that there was no requirement to document her activities, did not make the trial unfair or the conviction unsafe. The "key question" for the respondent was "whether DC [B]'s recognition of the suspect was reliable and whether the jury were in a position to test the reliability of the recognition during the course of the trial."¹⁰⁵

Drawing heavily on the cases of *Smith*,¹⁰⁶ *JD*,¹⁰⁷ *Chaney*,¹⁰⁸ and *Lariba*,¹⁰⁹ the Court of Appeal accepted that the "germane provisions of Code D ... should have been followed ... particularly as regards creating a contemporaneous record",¹¹⁰ but added that whether "a failure to follow Code D renders the verdict unsafe will depend on the particular facts of the case", particularly "the extent and significance of any breaches ... and any consequential unfairness that could have been caused".¹¹¹ The Court detected two "notable strands" in the authorities. One, where there was no contemporaneous record and the "recognition evidence [is] inherently poor". Another, where:

⁹⁸ (n 6) [35], paras 7-9.

⁹⁹ (n 6) [45] paras 10-12.

¹⁰⁰ (n 6) [47].

¹⁰¹ (n 6) [55]-[57].

¹⁰² (n 6) [50].

¹⁰³ (n 6) [50], [61] and [64].

¹⁰⁴ (n 6) [53].

¹⁰⁵ (n 6) [54].

¹⁰⁶ (n 37).

¹⁰⁷ (n 51).

¹⁰⁸ (n 53).

¹⁰⁹ (n 51).

¹¹⁰ (n 6) [86].

¹¹¹ (n 6) [87].

notwithstanding the failure to apply Code D ..., if a detailed explanation is given of the basis for the recognition, *particularly when the jury is in a position to view the relevant material itself*, it may – depending always on other factors – be fair to admit the recognition evidence.¹¹²

Applying this analysis to the present case, the Court held that the admission of DC B’s evidence had been fair. In her testimony, she maintained that she had “observed the finer details of the movements and other relevant characteristics of those involved”, she explained that she “became familiar” with some of them and “began to recognise some, but not all, of them”.¹¹³ She explained the reasons for her recognitions “by reference to relevant features of their appearance” and the jury was able “to see exactly what [DC B] saw” by viewing the images, which were of “relatively good quality”. As to potential contamination of DC B’s identifications, the jury “would have been well aware for Yaryare that other officers had suggested he may be one of the suspects.”¹¹⁴

For the Court of Appeal, non-compliance with the Code and Advice could be addressed by: making relevant imagery available to the jury; disclosing the videos to the defence; making witnesses available to be cross-examined; and, permitting the defence to adduce their own recognition evidence (as Yaryare chose to do, by calling imaging analyst Z); all of which had taken place.¹¹⁵ “[F]undamental difficulties” with DC B’s evidence, including “insufficient” explanation for the identifications and substantive breaches of Code D, became issues of weight rather than admissibility.¹¹⁶

As for the directions, the Court of Appeal expressed some concern that “the judge failed to direct the jury that the law required DC [B] to maintain a sufficient log whilst she was viewing the CCTV” and “failed to correct DC [B]’s assertion that the approach provided in *Smith* and Code D did not apply in this case.”¹¹⁷ Nevertheless, the judge “did remind the jury of the main criticisms that were made of DC [B]’s failings ... invit[ing] the jury to make appropriate allowances if they considered there was force in those contentions”.¹¹⁸ The Court considered that, “in assessing the officer’s reliability, the jury would have undoubtedly focussed on whether the notably detailed reasons she provided for recognising each of the appellants was undermined, principally, by the absence of a contemporaneous record”. However, in view of “the overall quality of the images and the footage, and the detail provided by DC [B]”, the Court concluded that the recognition evidence did not fall to be excluded as “poor” under the *Turnbull* guidelines.¹¹⁹

The Court of Appeal was assured by the presence of “*some other evidence*” implicating each of the accused.¹²⁰ For Yaryare, the Court pointed to the imaging analyst (P) finding “strong support” for Yaryare being person A. In addition, Yaryare’s associations with other “offenders” and the lie about his acquaintance with an individual convicted in related proceedings were “relevant background evidence”.¹²¹

4. A scientifically inflected review of the “identification” evidence in *Yaryare*

¹¹² (n 6) [88] (italics added).

¹¹³ (n 6) [89].

¹¹⁴ (n 6) [91].

¹¹⁵ (n 6) [84], citing *R v Lariba and others* (n 51) [46]. See also *Yaryare* [78], [79] and [88], respectively citing *R v Henry McGrath* [2009] EWCA Crim 1758 [15] and *R v Moss* (n 53) [23]-[25].

¹¹⁶ (n 6) [47], [49]. Where an investigator is not called upon to make an identification but is deemed to have watched *sufficient* “footage” they are usually allowed to testify regardless of compliance with Code D or the Practice Advice.

¹¹⁷ (n 6) [92].

¹¹⁸ (n 6) [93]. The jury were presumably obliged to speculate about this.

¹¹⁹ (n 6) [94].

¹²⁰ (n 6) [95] (italics added).

¹²¹ (n 6) [97].

The Court of Appeal affirmed the trial judge’s decision to treat the opinions of DC B as relevant and admissible. For the reasons developed in this and the following sections, it is our contention that the opinions of investigators should rarely be admitted. In this section, drawing on scientific research which, in large measure, is incompatible with the various assumptions, procedures and rules regulating the admissibility and use of this particular species of opinion evidence, we explain why trial procedures and safeguards – such as Code D, cross-examination and jury directions – are *not* capable of identifying and repairing the dangers created when investigators are allowed to express their inexpert impressions on facts in issue. Indirectly, we endorse some of the wisdom embodied in the Devlin Report to explain why criminal prosecutions that rely heavily on the opinions of investigators are generally problematic and will very often be substantially unfair.¹²²

The following analysis of *Yaryare* also endeavours to explain why: identification from images is much more difficult than courts imagine; cognitive bias and context effects create intractable problems for opinions and trial procedures; investigators and some of those admitted as imaging analysts are not relevant experts; and, inviting jurors to undertake their own image comparisons compounds risks. Rather than some kind of independent evaluation (or confirmation), jury comparisons of images (and voices and other types of evidence, such as fingerprints, shoeprints and handwriting) suffer from many of the problems confronting inexpert investigators. For jurors these dangers are accentuated by the suggestive context of the criminal trial and their inevitable exposure to other evidence. In addition, our discussion casts light on the tendency to allow evidence to be over-valued when investigating officers, who are exposed to all the information available to the juries (and appellate courts), are allowed to express *synthetic* opinions. We will observe how trials bias decision-makers and encourage irrationality, where non-image evidence is presented as separate from, and so capable of providing *independent support for*, identifications. In our case study, the same evidence appears to be informing decisions about identity or guilt – colloquially being “used” or “counted” – up to four times.

A. The opinions of investigators: DC B and the imaging analysts

We begin with the identification of *Yaryare*, by DC B. Appellate courts have long endorsed the view that juries are entitled to take into account that police officers are “professionally trained to carry out observation”.¹²³ Guidelines for prosecutors state that “honest police officers are likely to be more reliable than the general public” and “likely to have a greater appreciation of the importance of identification, and so to look for some particular identifying feature”.¹²⁴ However, as the editors of *Archbold Magistrates’ Courts Criminal Practice* note, “there is a lack of empirical evidence to support the proposition that police training makes officers better at identification than untrained civilians”.¹²⁵ The third of the categories in *A-G’s Ref (No 2 of 2002)* is based on a similar misconception, namely that police officers, by virtue of their office, either have or acquire special skills

¹²² The unfairness is consistent with anxieties identified in the Devlin Report (n 13) 2.24 and 4.92

¹²³ *R v Ramsden* [1991] Crim LR 295 (CA); *Powell v DPP* [1992] RTR 270 (DC); *R v Tyler* (1996) Cr App R 332 (CA); *R v Spittle* [2008] EWCA Crim 2537, [2009] RTR 14.

¹²⁴ Police are sometimes invested with special abilities beyond those possessed of jurors – “inevitably deeper or greater or more considered than could have been arrived at by the jury”, see *Yaryare* (n 6) [75], citing *R v Abnett (Gary)* [2006] EWCA Crim 3320, [20]. Claims about special abilities are inconsistent with the results of scientific studies. When tested, police, as a group, tend to be *less accurate* than other participants. See Annelies Vredeveldt, Peter J van Koppen, ‘The Thin Blue Line-Up: Comparing Eyewitness Performance by Police and Civilians’ (2016) 5(3) *Journal of Applied Research in Memory and Cognition* 252 and Anna Lvovsky, ‘The judicial presumption of police expertise’ (2017) 130 *Harvard Law Review* 1995.

¹²⁵ Stephen Leake (ed), *Archbold Magistrates’ Courts Criminal Practice 2022* (18th edn, Sweet & Maxwell 2022).

that a layperson does not possess or is unlikely to acquire. It is in this vein that DC B was presented as a species of expert.¹²⁶

In responding to the contention that investigating officers possess, or are capable of acquiring, either expertise or specialised knowledge in image comparison, the first point to make is that attempting to determine the identity of an unknown person by comparing their appearance with images of a known person is a cognitively demanding and unexpectedly error-prone task.¹²⁷ In favourable conditions, with high quality close-up images and plenty of time, ordinary adults tend to make errors around 20-30% of the time. We are much more accurate, though nowhere near error-free, when the faces we are asked to compare or recognise are very familiar to us, such as when we are asked to identify images of family members, friends and colleagues –i.e. *A-G's Ref (No 2 of 2002)* category (ii).¹²⁸

This last point is important because the judgment treats DC B as someone who is familiar with the images and with some of the “unknown” persons in the images.¹²⁹ She is presented as having acquired sufficient familiarity with person A to be able to *recognise* him as Yaryare. The following extract is one of several references in the judgment:

... during her evidence DC [B] set out that, whilst viewing the footage over the course of hundreds or thousands of hours, she observed the finer details of the movements and other relevant characteristics of those involved, and *over time she became familiar with some of the people who feature in this case*. Further, she began to recognise some, but not all, of them. She viewed the footage on a variety of different types of equipment, from ordinary laptops to “high tech machines”.¹³⁰

But this is not how genuine familiarity is acquired. Ordinarily, we become familiar with a person through a range of in-person encounters. Exposure tends to be interactive and dynamic, and extends beyond appearances. Usually, we hear, often touch and sometimes smell (or even taste) those with whom we are familiar. Our exposure occurs across moods, social contexts, and time. And, importantly, we almost always know with a high level of certainty who the person is, and whether it is always the same person with whom we are becoming familiar. These conditions are quite different to the repetitive watching of segments of CCTV video (or examining selected frames) featuring persons of interest.

Studies suggest that there is no significant advantage to be gained from a non-expert (like DC B) repeatedly studying a limited set of photographs or videos.¹³¹ (And, this is before we consider the suggestive circumstances in which she watched and purported to recognise suspects.) Passively watching an unknown person in a limited set of recordings, even if the quality is reasonable or good, does not generate familiarity.¹³² This kind of exposure is not known to enhance accuracy with a specific face or the ability to compare and match faces – or bodies or movement.¹³³ DC B was not familiar with Yaryare outside the context of the investigation, and so we cannot be confident that she possessed any

¹²⁶ The trial judge suggested DC B had “acquired special knowledge” (n 6) [45](7) and had developed “expertise” [46](5). The Court of Appeal referred to DC B being “entitled... to give evidence as an expert on the CCTV footage” [26].

¹²⁷ This also applies to comparing a physical person with images, and the memory of a person (in recognition cases).

¹²⁸ Vicki Bruce, Z Henderson, C Newman et al, ‘Matching Identities of Familiar and Unfamiliar Faces Caught on CCTV Images’ (2001) 7(3) *Journal of Experimental Psychology: Applied* 207.

¹²⁹ Consider the extracts reproduced in the previous paragraphs.

¹³⁰ (n 6) [89] (italics added), see also [29] (DC B’s second statement), [32], [45] paras 9-11 (trial judge).

¹³¹ R Clutterbuck and RA Johnston, ‘Exploring Levels of Face Familiarity by Using an Indirect Face-Matching Measure’ (2002) 31(8) *Perception* 985; David White, A Mike Burton, Rob Jenkins and Richard I Kemp, ‘Redesigning photo-ID to improve unfamiliar face matching performance’ (2014) 20(2) *Journal of Experimental Psychology: Applied* 166.

¹³² We are to assume that it is the same person across the images, based apparently on similarities in clothing and appearance. This may be reasonable as a working hypothesis, but it is hardly sufficient for categorical identification.

¹³³ See e.g. David White, Richard Kemp, Rob Jenkins et al, ‘Passport Officers’ Errors in Face Matching’ (2014) 9(8) *PLoS ONE* 1 and Royal Society and Royal Society of Edinburgh, *Forensic gait analysis: A primer for courts* (London, 2017).

“specialised knowledge” (*A-G’s Ref (No 2 of 2002)* category (iii)) or heightened abilities associated with familiarity (category (ii)) as it is typically understood.¹³⁴

Secondly, underlying these identifications is the issue of cross-racial identification – also known as the “other race effect”.¹³⁵ We are not provided with the demographic details of DC B or the imaging analysts but we are told about the suspects. Yaryare was a young man who was part of a group described as “mainly of Somali descent”, some of whom were said to be related to one another.¹³⁶ Most people are better at identifying faces that share their ethnicity. In a meta-analysis, Meissner and Brigham found that those participating in studies were 1.4 times more likely to recognise persons who shared their ethnicity as opposed to other-ethnicity faces.¹³⁷ To put this another way, we are even more prone to error when trying to identify persons of different ethnicities to our own.¹³⁸

DC B was not an expert in face comparison. Nor was she an expert “on the CCTV footage” or in the appearance of the defendants, for there is no such expertise.¹³⁹ She nevertheless made confident assertions about her abilities (in terms that resembled the applicable legal standards), claiming to have “become familiar” with the people depicted in the images and to be “able to recognise and identify” them as they (re-)appeared in the various videos.¹⁴⁰ An expert would be required to provide the court with information enabling it to evaluate the reliability of her opinions.¹⁴¹ It does not appear that DC B knew anything about relevant – though critical and often counter-intuitive – scientific literatures and knowledge. She was not conversant with research methods and error rates, or risks from context and cognitive bias (of which, more below). The upshot is that cross-examination (and examination-in-chief) could not place the jury in a position to evaluate her evidence.¹⁴² Rather, DC B was engaged in a difficult and unexpectedly error-prone task and, as explained below, her interpretations were compromised by her exposure to information and insights beyond the images themselves.

As for the imaging analyst (P) called by the prosecutor, he was qualified and admitted as an expert witness, but no information is provided about the validity and accuracy of his method(s) or his actual abilities.¹⁴³ The judgment documents the strength of P’s conclusion as “strong support” for Yaryare being person A, but we are not told how often P is wrong when he expresses such an opinion or whether other imaging analysts would (or did) disagree with the strength he allocates, *and why*.¹⁴⁴

¹³⁴ Our criticisms here concern category (iii) and related jurisprudence, rather than the integrity of DC B.

¹³⁵ Christian Meissner and John Brigham, ‘Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review’ (2001) 7 *Psychology, Public Policy, and Law* 3.

¹³⁶ (n 6) [12]- [13].

¹³⁷ Meissner and Brigham (n 135).

¹³⁸ Of course, features attributed to race will not always be observable, especially if the “crime” images are of low quality; See Ahmed M Megreya, David White and A Mike Burton, ‘The Other-Race Effect does not Rely on Memory: Evidence from a Matching Task’ (2011) 64(8) *Quarterly Journal of Experimental Psychology* 1473.

¹³⁹ (n 6) [26]. DC B seems to have compared persons in the images with Yaryare, once other information pointed to his possible involvement. This is not recognition based on familiarity and not a reliable way to undertake image comparison.

¹⁴⁰ (n 6) [29].

¹⁴¹ *Davie v Edinburgh Magistrates* [1953] SC 34.

¹⁴² While scientifically-informed cross-examination could, in theory, draw attention to serious methodological oversights and dangers, it would still be for the jury to make sense of what these “omissions” or limitations might mean for the evidence or even their own comparisons. But how are jurors to make sense of unexplicated and counter-intuitive problems with “expert” opinions and comparisons?

¹⁴³ While some imaging analysts are better than ordinary persons, not all are. See P Jonathon Phillips, Amy N Yates, Ying Hu et al, ‘Face Recognition Accuracy of Forensic Examiners, Superrecognizers, and Face Recognition Algorithms’ (2018) 115(24) *Proceedings of the National Academy of Sciences* 6171. According to the PCAST Report (n 3) 6, and the results of empirical studies, when it comes to comparisons and comparison forensics, we should not rely on a person’s experience doing something as evidence of an ability. See also White et al (n 133).

¹⁴⁴ We know that Z was less confident, but there are no references to him raising scientific research or proffering methodological criticisms of P’s (or DC B’s) methods and conclusions.

The form of the imaging analyst's opinion is, nevertheless, illuminating. Whereas P sought to place his conclusion in a subjective probabilistic framework, limiting it to "strong support" for it being Yaryare rather than someone else, DC B actually invoked the language of guilt; as the sole investigator she was "satisfied" with her categorical identification. Ironically then, the police officer was both willing and allowed to express her inexpert impression in terms that were far stronger than those used by the legally recognised expert. Neither the advocates nor the judiciary appear to have commented on this curious discrepancy. Though, to be clear, shifts in nomenclature or strength would not be capable of saving DC B's (and perhaps P's) opinions for, as we shall see, hers was not a reliable, uncontaminated recognition based on images.¹⁴⁵

In the absence of fundamental information about the validity of their methods, their performance and the general difficulties inherent in the identification of persons from images, how is a jury to rationally evaluate the opinions of DC B and P? Are they to rely on the number of times each had done this before (e.g. proffered opinions during investigations or testified about identity in cases where convictions were obtained), legal recognition of the opinions, their employment by the state, the plausibility of their "methods", their confidence and demeanour, impressions of resilience under cross-examination, or something else?¹⁴⁶ Furthermore, to treat the opinions of DC B and P (and Z) as implicitly independent and mutually supporting is misconceived and liable to mislead the jurors and subvert fact-finding.

B. The effects of context and unconscious bias on interpretations of the images

The circumstances in which DC B came to *recognise* Yaryare and the information available to her were raised during the trial and on appeal, albeit without reference to the Forensic Science Regulator's Guidance on cognitive bias effects or any other scientific literature.¹⁴⁷ Consideration of the issues was framed by the terms of Code D and case law, such as *R v Smith*. In this section, drawing on scientific research on the dangers posed by the environment of interpretation (or context) and cognitive biases, we explain why prevailing admissibility jurisprudence, Code D, the ACPO *Practice Advice* and trial safeguards were incapable of capturing, let alone mitigating, the very serious risks to perception, cognition and decision-making created during the investigation, prosecution and appeal.

In this section we draw upon recent reports prepared by peak scientific organisations to provide a succinct introduction to some of the apparently unrecognised but nevertheless very real dangers introduced by the context and the failure to guard against cognitive biases (such as suggestion, expectation and confirmation). The explanations and recommendations in these reports were primarily directed towards forensic science evidence, particularly feature comparison evidence, which involves comparing two or more samples (or traces) in order to evaluate similarities, usually to develop inferences

¹⁴⁵ Rather, DC B came to identify Yaryare following her participation in the investigation and exposure to (really contamination from) a great deal of admissible and inadmissible information, including the impressions of other police officers and interaction with the imaging analyst (P). While DC B and courts might characterise her evidence as recognition, we cannot retrospectively ascertain the factors or circumstances that led her to her conclusion. We cannot even be confident, notwithstanding her presumably sincere testimony, that it was primarily based on the images.

¹⁴⁶ (n 6) [45] para. 10 (trial judge): "What they make of DC [B]'s evidence in terms of quality and reliability and the other evidence is a matter for them." At [33]: "She explained her reasons for identifying him, namely by reason of height, age, ethnicity, walk and posture ('stands with his shoulders rounded and his hips forward'), hairline, shape of his lips, ear position and skin colour. Subject A had been wearing the same clothing in all the footage. In evidence DC [B] referred also to the shape of his eyes." The claimed ability to identify facial features or provide a plausible sounding explanation for recognition does not overcome the underlying problems. In addition, all of the suspects were young, and height is correlated with ethnicity, as are skin colour and to varying degrees hairline and shape of lips.

¹⁴⁷ Forensic Science Regulator, *Guidance: Cognitive Bias Effects Issue 2* (July 2020), 11.1-11.4.

about their source or identity.¹⁴⁸ Historically, many of the forensic sciences involved in feature comparison used terms such as “match” to describe similarities between a relevant trace (e.g. a latent fingerprint or bullet cartridge) and a reference sample (from a known suspect or recovered weapon). Opinions tended to be expressed in categorical terms, linking the trace to a specific person (or thing). With the emergence of DNA profiling and stabilisation of its empirically-based statistical platform, attentive scientists became increasingly critical of categorical identifications and the lack of method validation among the historical forensic sciences. Recall that DC B expressed her inexpert opinion in categorical terms.¹⁴⁹

So, what advice have mainstream scientists offered about context and cognitive biases for those engaged in feature comparisons such as trying to identify persons in images? In its seminal report in 2009, the United States’ National Academy of Sciences (NAS) insisted on the need for “a reliable scientific methodology” so that results are valid and accuracy understood.¹⁵⁰ It placed a conspicuous and unprecedented emphasis on human factors.¹⁵¹ Those endeavouring to evaluate opinions must consider “the extent to which practitioners in a particular forensic discipline rely on human interpretation that could be tainted by error, the threat of bias, or the absence of sound operational procedures and robust performance standards.”¹⁵² These fundamental questions about methods and human interpretation, particularly bias and error, led the authors to conclude that “it matters a great deal whether an expert is qualified to testify about forensic evidence and whether the evidence is sufficiently reliable to merit a fact finder’s reliance on the truth that it purports to support.”¹⁵³

More recently, President Obama’s Council of Advisers on Science and Technology (PCAST) endorsed the concerns expressed in the NAS Report, again with respect to the vulnerability of forensic scientists:

Subjective methods require particularly careful scrutiny because their heavy reliance on human judgment means they are especially vulnerable to human error, inconsistency across examiners, and cognitive bias. In the forensic feature-comparison disciplines, cognitive bias includes the phenomena that, in certain settings, humans may tend naturally to focus on similarities between samples and discount differences and may also be influenced by extraneous information and external pressures about a case.¹⁵⁴

Drawing on “hundreds of scientific studies across dozens of domains” these reports focus attention on the vulnerability of human perception, interpretation and decision-

¹⁴⁸ “Traces” include fingerprints, firearms and tool marks, handwriting, shoe and tire marks, blood spatter, DNA profiling, voices, as well as persons of interest in images.

¹⁴⁹ (n 6) [62]. Unqualified opinions tend to be common among those lacking expertise. Interestingly, P and Z expressed opinions in forms influenced by Bayesian norms. P’s subjective impression of similarities, was expressed through a qualitative hierarchy (also subjectively selected) and was presented in the not entirely helpful language of “strong support” for the proposition “that Yaryare and person A were the same individual.”

¹⁵⁰ NRC report (n 3) 9. We note that the Law Commission placed emphasis on reliability (n 3), but English courts have not required much in response. The revised Criminal Practice Directions and Criminal Practice Rules for expert witnesses have not been applied to investigators, nor led to the exclusion of their opinions. Though, see Tony Ward, ‘Explaining and trusting expert evidence: What is a “sufficiently reliable scientific basis”?’ (2020) 24 E & P 233.

¹⁵¹ Interest in “human factors” is oriented to understanding human performance, especially vulnerabilities. Studies are frequently concerned with performance in the workplace, in order to make systems such as flight safer or athletes more competitive.

¹⁵² NRC report (n 3) 9 (italics added).

¹⁵³ The passage continues, “Unfortunately, these important questions do not always produce satisfactory answers in judicial decisions ...” See also Expert Working Group on Human Factors in Latent Print Analysis, *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach* (National Institute of Standards and Technology, 2012) (hereafter ‘NIST Report’).

¹⁵⁴ PCAST Report (n 3) 31.

making.¹⁵⁵ They caution about the vulnerability of investigators to selectivity, searching for confirming information, discounting differences, seeing what is expected and focusing on apparent similarities, along with the influence of extraneous (or domain irrelevant) information and external pressures.¹⁵⁶ Significantly, cognitive biases can act unconsciously and are not obviated by training, experience, professionalism, the best of intentions, or by acting with integrity and conscious effort. In most cases neither training and experience, nor relevant expertise, protect against these invidious threats to perception, interpretation and decision-making.¹⁵⁷ This last point is important because it applies to the experiential opinions of DC B (notwithstanding any training she might have received), as well as the opinions of P (and Z).

A series of studies demonstrates what all of the previous extracts imply: that being an expert in some domain does not necessarily protect against bias.¹⁵⁸ One of the most notorious demonstrations of the vulnerability of experts involved a study conducted by Itiel Dror and his colleagues involving certified and experienced fingerprint examiners. Each participant was invited to compare sets of fingerprints to determine whether they matched, did not match, or were unsuitable for matching. Along with the prints, Dror provided (domain or task-irrelevant) information which suggested that the prints did *not* match – e.g. another suspect confessed or other examiners had concluded they did not match. What the examiners were not told is that the pairs of prints they were each asked to review were taken from their personal case files and they had previously matched them as part of their routine case work. Against expectations, a large majority of the examiners unwittingly offered different opinions about whether the prints matched when comparing them in the shadow of the suggestive information.¹⁵⁹ These studies are revealing because forensic scientists (here fingerprint examiners) with undoubted expertise came to inconsistent conclusions about whether two prints matched when they were presented with additional information that was not necessary to carry out the comparison. Exposure to *domain irrelevant information* – which is not required for the analysis – has a demonstrated tendency to bias comparisons.¹⁶⁰ For experts and non-experts alike, the effects of context and cognitive biases are more pronounced where interpretations are unusual or challenging or appear to be confirmed by other information (e.g. knowing the suspect is associated with a gang or lied during a police interview). Following Dror’s lead, in something of a forensic “cottage industry”, other groups of forensic scientists, doctors and anthropologists have been shown to be vulnerable to domain irrelevant information.¹⁶¹

The idea that all humans, including police officers, experts, jurors and judges, are vulnerable to context effects and cognitive biases should not be surprising. We are all aware of highly trained biomedical researchers using double-blind randomised trials for the clinical evaluation of pharmaceuticals. In the standard version, researchers “blind” the patients, treating clinicians and others as to whether they are receiving/dispensing the active ingredient or a placebo to avoid the notorious risks posed by suggestion, expectation

¹⁵⁵ NIST report, (n 153) 11 (Box 1.2).

¹⁵⁶ See NRC report (n 3) Recommendation 5; NIST report (n 153) Recommendations 3.3 and 7.1; PCAST report (n 3) Recommendation 1.

¹⁵⁷ Significantly, notwithstanding the over-reaction to the mere suggestion of bias in DC B’s cross-examination (see Section 4D), vulnerability to cognitive biases is not an ethical issue or a problem of personal integrity. Though, ethics might be raised with respect to failing to address problems as they become known.

¹⁵⁸ Glinda S. Cooper and Vanessa Meterko, ‘Cognitive bias research in forensic science: A systematic review’ (2019) 297 *Forensic Science International* 35: “As demonstrated in this set of studies, case-specific information can be influential, even when the information is wrong. ... so it is important to carefully consider whether, and when, information is provided to the analyst.”

¹⁵⁹ The expectations of the fingerprint examiners that is; not those conducting the experiments. See R Rosenthal and D Rubin, ‘Interpersonal expectancy effects: The first 345 studies’ (1978) 3 *Behavioral & Brain Science* 377.

¹⁶⁰ PCAST report (n 3) 5, 10, 14, 31ff.

¹⁶¹ Cooper and Meterko (n 158).

and confirmation. Experience as a doctor or researcher does not somehow enable these participants, some unquestionably expert, to avoid subtle cues and threats to cognition. A large range of modern scientists use blinding to avoid self-deception and error. Many forensic science institutions, in the wake of Dror's studies, recommendations from peak scientific bodies and the Forensic Science Regulator, have begun to manage domain irrelevant information.¹⁶²

In terms of identifying person A from CCTV recordings, all that is required is the crime-scene images (or video) and a set of reference images (or video) of the suspects.¹⁶³ Everything else is gratuitous, domain irrelevant information with the demonstrated potential to mislead. Yet, DC B's exposure to domain irrelevant information and suggestion was absolute. As the primary investigator over three years, she was exposed to interviews, criminal records, suggestions, hearsay, innuendo and speculation. We are not told about P's exposure but there are no references to his use of methods to protect against or mitigate similar risks or relevant disclosure. Rather, the judgement indicates that he worked closely – "liaised" with and provided "assistance" to – DC B.¹⁶⁴

Of course, while domain irrelevant information is not required for – and should be withheld from those involved in – feature comparisons, we do not suggest that it is irrelevant to proof of guilt. Strands of domain irrelevant information might be combined, and all of the domain irrelevant information might be combined with the results of valid image comparisons, but *only* if the comparisons were not influenced by the domain irrelevant information. If those engaged in comparisons were exposed to domain irrelevant information (or steeped in suggestive contexts), we cannot be sure about the bases for their opinions. Are they based on the images or are they informed, even determined, by other information or the context of the comparison? We cannot resolve the basis or bases of interpretation following exposure to domain irrelevant information; which is why prospective blinding is so fundamental.¹⁶⁵

5. Implications for the case against Yaryare

A. DC B's "recognition"

Prior to DC B's "recognition", the judgment implies that Yaryare was suggested to DC B by other police officers (who did not testify). She knew of his associations with others who had already been placed at the scenes (or arrested or charged) and knew they appeared in his telephone contact list.¹⁶⁶ What was presented as "recognition" was inextricable from DC B's exposure to other suggestive information. Rather than recognition, her identification seems to have been some kind of deduction from the limited pool of suspects available to her and the totality of the information she had acquired. The lie about Hudur may have strengthened or confirmed her beliefs. DC B did not manage, and no investigator could be expected to manage, the suggestive context and biasing information to which she was exposed during the

¹⁶² Dan E Krane et al, 'Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation' (2008) 53 *Journal of Forensic Sciences* 1006; C Robertson and A Kesselheim (eds), *Blinding as a Solution to Bias: Strengthening Biomedical Science, Forensic Science, and Law* (Academic Press, 2016). In some areas scientists have sought to reduce human subjectivity, as in algorithms (e.g. STRMix and TrueAllele) now widely used to calculate statistics for DNA matches.

¹⁶³ In most cases it will be appropriate to include a small pool of credible foils.

¹⁶⁴ DC B is reported to have received "assistance from" P: (n 6) [28]. P's opinion is presented as "separate" (at [47]) and supporting (at [64]), but his evidence was not independent if either they worked together *or* drew upon the same domain irrelevant information.

¹⁶⁵ This is why pro-active blinding is so important. We need to prevent rather than try to retrospectively identify or manage it. Cognitive bias can be very subtle, such that documentation is generally inadequate.

¹⁶⁶ (n 6) [91].

investigation. Nevertheless, the lack of blinding prevents us from confidently attributing her “recognition” to the images and her “familiarity” with person A and/or Yaryare.¹⁶⁷

B. Irreparably biased juror comparisons

Jurors were encouraged to consider the images in order to evaluate the opinions of DC B, P (and Z) and to determine whether Yaryare was the offender (and therefore guilty).¹⁶⁸ These non-expert, non-familiar witnesses undertook comparisons in some of the most suggestive conditions imaginable. The jury considered the images after the primary investigator had categorically identified Yaryare as person A *and* an imaging analyst, who was presented as an expert, claimed the images lent “strong support” to Yaryare being person A. Jurors were also encouraged to consider all of the other evidence implicating Yaryare, including his associations with other “offenders”. They undertook these (potentially other-race) comparisons without meaningful alternatives – other young men who might resemble the person of interest – during the course of an accusatorial proceeding.¹⁶⁹

Although routine in criminal proceedings involving CCTV and other images, where identity is in issue, juror comparisons are fraught in the extreme.¹⁷⁰ They are unnecessary, unexpectedly unreliable (and so likely, in most trials and appeals, to mislead jurors and judges) and very likely to result in image evidence being misunderstood and over-valued. Jury comparisons, as we shall see, are incurably biased toward conviction.

Already the reader should have an emerging sense of just how problematic jury comparisons might be. Although ordinary persons are unexpectedly error prone when comparing images, they tend to be over-confident about their abilities.¹⁷¹ Impressions of competence or confidence seem to be extrapolations from our performance identifying faces that are familiar to us, but identifying unfamiliar faces is a fundamentally different activity.¹⁷² The jurors may have seen Yaryare during the course of the trial but they were not familiar with him. Rather, he was presented as the offender and this suggestion and their expectations inexorably influenced their perceptions and interpretations of the images.¹⁷³

What seems likely, and there are few scientifically credible alternatives, is that the kinds of information that contaminated the opinions of DC B and P, also (or again) biased the jurors when they came to consider the reliability of the opinion evidence and to make their own interpretations of the images. They were not only exposed to, but encouraged to consider, domain irrelevant information as support for identity when making their comparisons. This means that some of the strands of evidence that informed the opinions of DC B and P were also, presumably unwittingly, mis-represented to the jury as

¹⁶⁷ Her beliefs were also vulnerable to confirmation bias – the post-“recognition” *confirmation* afforded by police and prosecutors who accepted her opinions and the strength of the case, and were willing to charge offenders and pursue convictions.

¹⁶⁸ Though, the basis for any evaluation of her biased interpretation of the images (and consideration of the overall circumstantial case) is unclear. Moreover, once we recognise that DC B’s opinions were not based on the images, the redundancy and constitutional illegitimacy of the exercise becomes clear.

¹⁶⁹ Note that foils are used for eyewitnesses but not for investigators and jurors.

¹⁷⁰ We have used “juror” rather than “jury” because this is both an individual and collective exercise. Its collective nature means that the results are unavoidably contaminated. See David White, Mike A Burton, Richard I Kemp et al, ‘Crowd Effects in Unfamiliar Face Matching’ (2013) 27 *Applied Cognitive Psychology* 769.

¹⁷¹ Romina Palermo, Bruno Rossion, Gillian Rhodes et al, ‘Do People have Insight into their Face Recognition Abilities?’ (2017) 70 *Quarterly Journal of Experimental Psychology* 218.

¹⁷² A. Megreya and M. Burton, ‘Unfamiliar faces are not faces: Evidence from a matching task’ (2006) 34 *Memory and Cognition* 865.

¹⁷³ (n 6) [63].

independent strands of evidence. The jurors undertook their own comparisons, and evaluations of DC B, P and Z, in a miasma of suggestion and expectation.¹⁷⁴

C. Contamination of the circumstantial case

The prosecutor, DC B, the judge and the appellate court each presented the various strands of evidence in the case against Yaryare as independent support for identity and guilt. That approach is misguided because the strands were not independent but rather irreparably contaminated and cross-contaminating. The same evidence was used (or acting) more than once and so was unwittingly over-valued (or over-valuation was likely and could not be prevented). Some of the strands appear to have been used, or were available to influence interpretations, on as many as four separate occasions.

This misunderstanding is perhaps clearest in the Court of Appeal's treatment of the no case to answer submission (Ground 3). Reviewing the case against Yaryare, the Court enumerated the strands of evidence:

DC [B] had had extensive opportunities to study [Yaryare's] appearance and there were a number of sightings of person A on the footage which provided a clear basis for making comparisons. *Furthermore*, [P], the imaging analyst, found "strong support" for the suggestion that Yaryare and person A were the same individual.

Additionally, there was unchallenged evidence from the security staff at Leicester College ...that the appellant had an association with a number of the offenders,.... *There was also* evidence from his mobile phone contacts list that he had close contact with [other] offenders. *Finally*, he lied in his interviews under caution about his friendship and association with Hudur. In fact Hudur had been with the appellant at the time of the offences and he was convicted by the jury.¹⁷⁵

The evidence against Yaryare consisted of the "recognition" evidence of DC B. "Furthermore" the "strong support" for identity provided by P. "Additionally", there was evidence of Yaryare's association with other "offenders" and one of his co-accused (Osman) based on associations and the contact list on his mobile phone.¹⁷⁶ There was also a "lie" about his friendship with Hudur.¹⁷⁷ The strands are treated by the Court as being implicitly independent and mutually supporting and the strength of the case is said to derive from the combination of these independent strands.¹⁷⁸

Similarly, when DC B was challenged about possible bias during her cross-examination, she emphatically rejected the imputation, contending that "[t]he identifications *are supported by* other evidence that the jury have heard."¹⁷⁹ Consider also the trial judge's references to supporting evidence when directing the jury:

So, you have to look at DC [B]'s evidence, considering what has been said by different people and see if you think there is mileage in that, or do you think that you can accept her identifications. The prosecution say: Of course, *it's not just DC [B]'s evidence you're looking at in isolation, there is other evidence to support...*¹⁸⁰

¹⁷⁴ These dynamics extend beyond the interpretation of images (or voices), but with respect to comparisons they are acute. See Gary Edmond, 'Against jury comparisons' (2022) 96 Australian Law Journal 315.

¹⁷⁵ (n 6) [62]-[63] (italics added).

¹⁷⁶ It is unclear how many offenders or (aside from Hudur) what the proof of their being 'offenders' actually was. If it is based on their convictions then there is a danger of tautology; or convicting Yaryare on the basis of his association with others who were convicted. Conviction, as a litany of historical errors and the continuing work of the Criminal Cases Review Commission make clear, is not ground truth.

¹⁷⁷ This was treated as a lie based on the "fact" of Hudur's conviction: (n 6) [63]. There may have been evidence supporting Yaryare's presence (apart from CCTV images), but we are not told what it is.

¹⁷⁸ (n 6) [24]: "Second, evidence from a facial mapping expert, [P], which provided support for DC [B]'s identifications"; [46] (trial judge).

¹⁷⁹ (n 6) [69](7) (trial judge, original emphasis removed, italics added).

¹⁸⁰ (n 6) [46] (italics added), [69] para. 5 (trial judge): "The prosecution put her forward as a reliable witness and they say there is supporting evidence from other sources."

Table 1 reflects the basic case against Yaryare as described by prosecutor, trial judge and Court of Appeal.

DC B's "recognition" + P's "strong support" + Associations + Lie + Jury comparison
(strand) (strand) (strand) (strand) (real evidence)

Table 1: Circumstantial strands (and real evidence) presented as *independent* by prosecutor and courts.

However, we submit that the characterisation of the prosecution case as an assemblage of mutually supporting strands is fundamentally mistaken for, as we have seen, they were not in fact independent. When DC B made her identification, she knew all of the other information. We cannot determine what actually led her to make the identification (characterised as recognition) but we cannot be sure, or even confident, that the non-image information was not determinative (or a significant factor). Similarly, the imaging analyst "liaised with" DC B and was exposed to some, perhaps all, of the same information, as well as DC B's impression.¹⁸¹ Their opinions as to the identity of Person A were not independent and did not support one another as the Court suggested. Rather they were contaminated and contaminating. Evidence that informs an interpretation – regardless of whether the interpreter is conscious of the influence – cannot be used to confirm that interpretation. The presentation and representation of the various strands as independent and mutually supporting involves circular reasoning.

These problems were compounded when the CCTV images and images of Yaryare, the opinions of DC B and P, and the non-image evidence were presented to jurors and they were encouraged to make comparisons of Yaryare and Person A. The jurors were presumably also unaware that the strands of evidence with which they were presented were not independent but irreparably cross-contaminated and, for their comparisons, contaminating.

Table 2 might help the reader to visualise some of the contamination, particularly the potential for double-, triple- and quadruple-dipping. Where the same "strand" appears in multiple columns it is available and presumably informing the interpretive task. It can no longer be considered to be independent of, or capable of providing support for, the interpretations in the other columns. It does not matter if the witness, decision-maker or judge believes (or testifies) that it was not used or did not make a difference. Following exposure (or admission) it is impossible to guard against or manage influences – including the serious risk of unconscious bias.

Type of evidence or proof	DC B's opinion	P's opinion (imaging analyst)	Jurors' (and judges') Comparisons	Beyond reasonable doubt – evidence presented as independent 'support' for identification and guilt
Based on (or informed by)	IMAGES and <ul style="list-style-type: none"> • Everything about the case • Police suggestion about Y • P's "strong support" (1) • Association with others (1) • Lie (1) 	IMAGES and <ul style="list-style-type: none"> • DC B's "recognition" (1) • Association with others (2) • Lie (2) • Any suggestive information, including inadmissible things, disclosed 	IMAGES and <ul style="list-style-type: none"> • Context of trial (or appeal) • DC B's "recognition" (2) • P's "strong support" (2) • Association with others (3) • Lie (3) • Evidence implicating co-accused • Osman's criminal record (for Court of Appeal: conviction(s) and criminal records) 	IMAGES and <ul style="list-style-type: none"> • DC B's "recognition" (3) • P's "strong support" (3) • Association with others (4) • Lie (4) • (Jurors' comparisons)

¹⁸¹ DC B refers to developing chronologies of the CCTV to assist the analyst (n 6) [29]. "[W]orking with experts, she had compiled an edited CCTV compilation of each of the defendants, which she continually reviewed" [32].

Table 2: Information available to various participants when comparing images and evidence presented as independent strands (or support) for guilt. The strands in bold type in the “Jurors’ comparisons” column are all used elsewhere. The numbers in parentheses refer to the number of times the various strands appears and are available. Table 2 indicates how these strands are irreparably cross-contaminated. The lie and associations, for example, informed the interpretations by DC B, P and the jurors’ comparison, but were also presented as available to support each of their identifications in the determination of guilt.

Table 2 enables us to see how information available to DC B (and potentially P) was also available to inform the jurors’ comparisons. But things are even more complicated and invidious because the judge, prosecutor and DC B presented the opinions of these witnesses and the other evidence (association with others and the lie) as independent support for the opinions of DC B and P as well as the jurors’ own irreparably contaminated comparison (and verdict). All the evidence in the final column of Table 2, was presented as providing independent support on the question of identity and guilt – as in Table 1. Table 2 illustrates why they are not independent.¹⁸²

There do not appear to have been any references to the threats from double-, triple- or quadruple-dipping (i.e. ‘(2)’, ‘(3)’ and ‘(4)’ in Table 2) in the judge’s directions to the jury or the Court of Appeal’s judgment. Rather, the jurors were invited to treat cross-contaminated strands as independent pieces of evidence each implicating Yaryare as the offender. The various strands of evidence should only appear once (see Table 3, in the following section) and should only be relied upon once in evaluating the case against Yaryare. Where strands appear more than once, particularly in relation to challenging and contested identification from CCTV, they are being used illegitimately – that is, irrationally.

D. Directions (and other trial safeguards)

In conjunction with the other trial safeguards (e.g. prosecutorial restraint, admissibility rules, cross-examination, adducing the testimony of a defence imaging analyst (Z), and the burden and standard of proof) jury directions were presented as making Yaryare’s trial fair and his conviction safe.¹⁸³ The jury was directed in conventional terms that, “even where there is a single, unchallenged expert opinion, it remains a matter for ... the jury to decide whether or not [to] accept it”. They were told that the experts’ opinions were subjective and that the ways in which they expressed the strength of their conclusions (e.g. “strong support”) were merely “labels” and did not represent a “numerical ... scale”. They were told that DC B was not an expert witness, but “she [had] developed expertise by an extensive viewing of the visual footage and stills of images.” This was “not unusual”, but it remained for the jury “to consider whether [to] accept her evidence, or not”. The directions included a reminder that DC B’s notetaking had been criticised, particularly “on her lack of detail about initial reactions to the recording when she made recognitions, what features stood out and so forth”.¹⁸⁴ The jury was reminded that it could “make the appropriate allowances when ... consider[ing] her evidence” if it thought “there [wa]s (force) in [those criticisms]”.¹⁸⁵ The judge also gave a *Turnbull* warning in broadly conventional terms, including references to eyewitnesses picking out suspects on identification procedures (which was irrelevant in the circumstances) but provided very little guidance as to how the

¹⁸² In addition, we might reflect on how inadmissible information (e.g. beliefs and innuendo, criminal records, evidence of prior convictions, and the results of related proceedings), was also available to DC B (and possibly P) and so could not be prevented from informing, really biasing, their opinions and confidence.

¹⁸³ Though we are not told much about the cross-examination of DC B and P, there are glimpses in the appeal judgment.

¹⁸⁴ (n 6) [46].

¹⁸⁵ They were then told on several occasions that DC B’s conclusions had “support from other sources of evidence”: (n 6) [24], [46], [62], [64], [69] paras 5 and 7, [95], [96] and [97].

risk of mistake and the need for caution was applicable to the evidence before the jury, saying merely that, “this case doesn’t involve that kind of identification”.¹⁸⁶

On appeal, the challenge to the summing up and directions focused on conformity with the decision in *R v Smith* and Code D rather than the more fundamental and corrosive issues raised in the previous sub-sections. The directions warned of the risk of mistakes in abstract terms. Where the trial judge endeavoured to assist with the evidence, fundamental issues were repeatedly misconceived or overlooked. The directions presented DC B as a species of expert – an individual with an advantage – and, the judge reminded the jurors, more than once, of “support from other sources”, describing the other available evidence as “added support for [DC B’s] opinion of identification”.¹⁸⁷ The jury was told it was entitled to “use one witness’s evidence on identification, if you’re sure it’s correct, as partial support for the other.”¹⁸⁸

The directions privileged the jurors’ evaluation but did not identify the main problems with image interpretations, cognitive bias or cross-contamination of the evidence, including their own irredeemable comparisons. It is our contention that there was nothing the judge could say to address the problems with DC B’s (and presumably P’s) identifications, or to repair the suggestive jury comparison or the unmanageable potential for double and triple-dipping into evidence that was re-presented (and understood) as independent support.¹⁸⁹ Jurors engaged in their own comparisons having seen and heard a great deal of suggestive evidence and with a black man said to resemble person A sitting in the dock. These dangers cannot be managed through directions but it is revealing that no attempt was made to do so. Courts appear to be insensitive to serious limitations with their procedures and their facilitation of irrational approaches to the presentation and evaluation of evidence in these (and other) criminal proceedings.

Persistent lack of engagement with scientific knowledge suggests systemic problems. It is also a failure of admissibility standards that do not require those recognised (or treated) as experts to demonstrate their abilities or familiarity with relevant scientific literatures and procedures. Why did P and Z not insist upon blinding? Why was no challenge made to their evidence on the basis of the reliability factors now contained in CrimPD 19A and the FSR’s guidelines? We would urge practitioners (both legal and forensic) and the courts to attend to scientific knowledge and advice that bears directly on the value of evidence and the effectiveness of legal procedures, particularly those characterised as trial safeguards.¹⁹⁰

Not only were the dangers from context and cognitive bias substantially overlooked, but when counsel for Osman (one of Yaryare’s co-accused) raised the very real possibility of bias in the cross examination of DC B, it had serious consequences:

Q: Do you accept, against that background, you are anything but independent?

DC B: *I don't get what you mean.*

Q: Do you accept you have invested a lot in this case and you want a result?

DC B: *Only if it is the right result.*

Q: That you may have approached the identification evidence with at least subconscious [sic] bias?

¹⁸⁶ (n 6) [46].

¹⁸⁷ (n 6) [33], [46], [69].

¹⁸⁸ It is not clear, in terms of the Courts’ own logic, why this is characterised as only “partial”, but if the reason is some recognition of contamination, then this raises issues.

¹⁸⁹ (n 6) [69](7). See DC B’s reply to questions during cross-examination which were said to render Osman’s bad character admissible, below.

¹⁹⁰ Evidence cannot possess any value a jury may care to assign. Where there are decades of mainstream scientific research which suggests (or indicates) that some types of evidence may not be very probative, and warns against dangers posed by context and cognitive biases, these should not be overlooked, and jurors and judges cannot simply ignore them.

DC B: *No, absolutely not. The identifications are supported by other evidence that the jury have heard...*¹⁹¹

These scientifically-legitimate questions were treated by the trial judge and Court of Appeal as “a serious attack on the integrity and the honesty of DC [B]”.¹⁹² According to the Court of Appeal it was suggested that she had “behaved, or is disposed to behave, in a reprehensible way”.¹⁹³ These questions were said to enable the prosecutor to introduce Osman’s criminal record into the trial.¹⁹⁴ We suggest this response was disproportionate, unfair, and inconsistent with scientific research and advice. In advancing an alternative reading, we revert to the NAS report where “psychologists have identified many sources of cognitive bias, including ... motivational factors”.¹⁹⁵ (While the questions might have been expressed more dispassionately, they draw attention to serious risks and oversights.)

DC B’s – presumably sincere – response to these questions was also deeply problematic. Ignorant of dangers she unwittingly misled the court. Consequently, rather than reject the insinuation, indirect witnesses (such as DC B) should be able to explain the procedures used to protect their interpretation from biasing information and contexts. Lacking specialised knowledge (and relevant expertise), DC B could not assist. Legitimate questions about her vulnerability to human factors were misconstrued as an attack on her personal integrity. It was not unreasonable to ask DC B, as the sole investigator in a multi-year investigation, whether she was personally invested in the outcome and whether she may have been unconsciously affected by extraneous information to which she had been exposed.¹⁹⁶

In theory, the admission of Osman’s previous convictions for offences of violence could only be used by the jury when considering the case against Osman. It was not to be used against the associates who were on trial with him. The trial judge presumably directed the jury in these terms. Here again we can observe what might be considered an optimistic approach to human cognitive abilities (and trial safeguards). We also obtain a sense of misplaced confidence in “the opportunities to test and assess the reliability of the recognition evidence”.¹⁹⁷

E. Code D, the Practice Advice and admissibility jurisprudence

Unfortunately, the admissibility regime outlined in *A-G’s Ref (No 2 of 2002)* and endorsed in subsequent cases, is misconceived and largely incompatible with scientific research and advice.¹⁹⁸ Together with Code D and the *Practice Advice*, its categories rest on the assumption that investigators can reliably identify persons in images and can document potentially contaminating exposures and the reasons for their interpretations.¹⁹⁹ The jurisprudence assumes that, where practices are explained to jurors, any implications for

¹⁹¹ (n 6) [69](7) (trial judge, emphasis in original). Note the reference to “subconscious”, which suggests the approach is informed by fairly popular ideas. Most experimental psychologists and cognitive scientists do not use such Freudian inflected terminology.

¹⁹² (n 6) [102].

¹⁹³ (n 6) [101].

¹⁹⁴ Enabled, but did not require. The prosecutor appears to take, indeed push for, technical advantages without apparently understanding the implications for the fairness of the proceedings. See Gary Edmond, ‘(Ad)Ministering justice: Expert evidence and the professional responsibilities of prosecutors’ (2013) 36 *University of New South Wales Law Journal* 921.

¹⁹⁵ NRC Report (n 3).

¹⁹⁶ It is not easy to conceive of a police officer more personally invested in the outcome, than a detective who spent three years working in an effectively solo investigation. Her professional credibility, and perhaps career prospects, were inseparable from her opinions and the securing of convictions.

¹⁹⁷ (n 6) [91].

¹⁹⁸ (n 6) [84], citing *R v Lariba and others* (n 51) [39].

¹⁹⁹ Emily Pronin, Daniel Y Lin, and Lee Ross, ‘The Bias Blind Spot: Perceptions of Bias in Self Versus Others’ (2002) 28(3) *Personality & Social Psychology Bulletin* 369.

accuracy can be sensibly “factored in” or “taken into account”. However, even where there has been compliance with Code D and the *Practice Advice*, orthodox directions will rarely place a decision-maker – judge or jury – in a position to understand the basis of an investigator’s identification or the risk of error.

Treating bias as something that can be resolved by the jury is likely to generate irrational approaches to image evidence, such as treating the non-independent and mutually biasing opinions of DC B and P as supporting one another (and being independent of the domain irrelevant information contaminating them). These problems are compounded when jurors are themselves presented with contaminating evidence – such as Yaryare’s associations and the lie – as independent of both the identification evidence and their own comparisons.

For the reasons already provided, we contend that investigators should not be allowed to express their opinions about the identity of offenders and jurors should not be enabled, let alone encouraged, to make comparisons during the course of a trial. And there are other reasons, at least as compelling as the science-based criticisms, favouring proscription on investigator opinions and juror comparisons. These concern issues of relevance and redundancy. As to the first of these, if the investigator does not have a demonstrable advantage over the jury, then their opinion is irrelevant.²⁰⁰ In such circumstances (following the exclusion of irrelevant or biased opinions), it is no solution to give the images to the jurors to compare because such comparisons are unexpectedly error-prone, and will be biased by the accusatorial setting and the other evidence. If images are to be interpreted, with few exceptions the Crown should solicit the opinion of a genuine expert – discussed in Section F.

Redundancy is an even more fundamental problem. As we have seen, identification or recognition by investigators is frequently compromised by their participation in the investigation and exposure to domain irrelevant information. The result is that the opinions of investigators are (in)formed by their knowledge and exposure to information and environments beyond the images. Their opinions are a holistic synthesis – that is, interpretations based on all the information to which they are exposed. This is a problem because we do not typically allow witnesses, particularly those without relevant expertise, to synthesise evidence the jury will be asked to consider, evaluate and combine. Admitting the *synthetic* opinions of an investigator or imaging analyst trespasses on the fact-finding prerogatives and constitutional responsibilities of the jury.

It is with these considerations in mind, that we re-visit what the trial judge and Court of Appeal concluded with respect to PACE s 78:

It is to be regretted that DC [B] failed to apply the relevant parts of Code D and the guidance given in *Smith*, and we accept that the appellants and the jury were put at a disadvantage as a result, but given the extent of the opportunities to test and assess the reliability of the recognition evidence, and in the particular circumstances of this case, we consider that this expert evidence which it is accepted was *prima facie* admissible did not fall to be excluded under section 78. The evidence did not have such an adverse effect on the fairness of the proceedings that the court ought not to have admitted it.²⁰¹

Here we can observe judges struggling to get beyond compliance with the Code. From this limited perspective, the opportunity to take advantage of trial safeguards that are (or have been) incapable of exposing or correcting serious problems were characterised as sufficient to “test ... the recognition evidence”.²⁰² In a declaratory tone, unwittingly displaced from relevant scientific knowledge, the Court assures the community that DC B’s recognition

²⁰⁰ That is how a majority of the High Court of Australia reasoned in *Smith v The Queen* (2001) 206 CLR 650.

²⁰¹ (n 6) [91] (trial judge).

²⁰² See also *Yaryare* (n 6) [39] and *Smith* (n 37) [66] ff, esp [69].

evidence did not exert a particularly “adverse effect on the fairness of proceedings”.²⁰³ The broader epistemological issues and their profoundly deleterious implications for fact-finding, fairness and institutional credibility, pass without comprehension or comment.

F. Where are the demonstrable experts?

Just as research exposed the unexpected difficulty of identifying unfamiliar persons in images, and the counter-intuitive lack of advantage conferred by many types of experience and exposure, and notwithstanding concerns expressed by the FSR, recent research confirms the existence of individuals who are demonstrably good at image comparison for the purposes of identification.²⁰⁴ Three “groups” warrant our attention. Two are much better than judges and jurors. Let us consider the other group first. This group is composed of persons, like DC B (and possibly P and Z), who appear to possess expertise but, against expectations, often do not. Scientists studying abilities, in conditions where the correct answers were known, found that most of those involved in image comparison work – which presumably includes quite a few imaging analysts and facial mappers regularly used in investigations and prosecutions – are not significantly better than ordinary persons (such as jurors and judges).²⁰⁵ They found that interpreting and comparing images as part of a person’s routine work *does not enhance ability*.²⁰⁶

There is, however, good news. Scientists found that *some* of the analysts employed by police and security agencies were significantly better than ordinary persons. We might call this talented sub-set *specialists*.²⁰⁷ This finding is important because it confirms that there are genuine experts, but it also signals the need for caution and discrimination because not all of those employed in such roles performed better than ordinary persons.²⁰⁸ We cannot rely on experience in performing comparisons or with specific sets of images, nor other proxies such as job titles or convictions, as the basis for expertise.²⁰⁹ Studies confirm the need to attend to evidence of actual ability, derived from performance in tests featuring images resembling those in the specific proceedings but in circumstances where the ground truth is known.²¹⁰ Prosecutors should only adduce the evidence of specialists whose abilities have been demonstrated through rigorous testing.²¹¹

The final group is *super-recognisers*. This group is composed of those at the extreme end of a natural curve of ability. These individuals appear to have, on the basis of a battery of standardised facial memory and comparison tasks, an innate ability to accurately compare and recognise persons.²¹²

Investigative agencies, prosecutors and courts therefore have access to two groups with demonstrable expertise: specialists and super-recognisers. One of the benefits of using genuine experts, particularly specialists, is that they can present their opinions in ways that

²⁰³ (n 6) [91].

²⁰⁴ Discussed in Edmond and Wortley (n 48) and Edmond et al (n 68).

²⁰⁵ Their opinions are actually irrelevant.

²⁰⁶ By way of example, Australian passport examiners turned out to be no more accurate than university students. See White et al (n 133).

²⁰⁷ We do not yet know why they outperform others, but the important point is to focus on evidence of demonstrable ability rather than misleading proxies such as experience, job titles, plausibility, legal acceptance and so.

²⁰⁸ Scientists are trying to ascertain what makes some analysts better than others. In the meantime, many agencies have begun to select analysts from among those who already have relative advantages.

²⁰⁹ *Guidance for Prosecutors and Investigators* (n 67) 3: “simply citing years of experience and stating that the type of evidence has been previously accepted by the courts is insufficient to ensure that [an] expert is competent and their analysis and conclusions are reliable”. Yet DC B’s (non-expert) evidence was admitted on precisely this basis.

²¹⁰ Scientific studies and not criminal proceedings.

²¹¹ This is consistent with the advice of the FSR.

²¹² See Richard Russell, Brad Duchaine and Ken Nakayama, ‘Super-Recognizers: People with Extraordinary Face Recognition Ability’ (2009) 16(2) *Psychonomic Bulletin and Review* 252.

provide some indication of reliability and accuracy.²¹³ Another advantage conferred by genuine experts, particularly specialists, is that there is no need to engage jurors in speculative comparisons in incredibly suggestive conditions. Instead, the jury can be provided with the opinions of a genuine expert, along with insights into the limitations, uncertainties and error rates of those opinions. While the abilities and accuracy of genuine experts are affected by the quantity and quality of images (and some other factors), we can obtain a reasonable idea of their performance across a range of conditions based on standardised tests. This is the very information that enables fact-finders to evaluate opinions. And, the blinded, and therefore unbiased, opinions of a genuine expert can be safely – i.e. rationally – combined with other evidence because the expert’s perception and interpretation are not contaminated (and will not be contaminating). The opinions of a genuine expert, developed through a blinded process, are independent.

Before concluding, it is useful to make a few comments about face recognition (or comparison) using algorithms (and AI). With moderate to good quality images, modern algorithms tend to outperform humans, including the best performing humans. Depending on how they are deployed, algorithms can avoid dangers from context effects and cognitive biases.²¹⁴ And, depending on their validation testing, we can obtain a reasonably clear idea of their accuracy in specific types of conditions and with specific population groups. There can be dangers with algorithms, such as where algorithms are trained or designed to compare and identify particular groups – whether Caucasians or men, as opposed to African American women.²¹⁵ When using algorithms there is a need for their performance with the suspect’s demographic to be studied, disclosed and considered. Many algorithms have reached the stage where their results could be admitted and relied upon at trial. Unlike DC B, they can be impervious to context, afford insight into their limitations and frailties, and present the conclusion in statistical terms, along with an indication of how often they make errors in similar sorts of comparisons.²¹⁶

Where experts and algorithms are incapable of making meaningful comparisons, we suggest that images should not be given to the jury to resolve identity. Jurors should not be invited or enabled to make comparisons where the best humans and algorithms cannot. We should not forget that human perception and cognition is most vulnerable to biases where the interpretation is difficult and the available information (or trial context) suggests an answer (or identity).

G. What would our case look like in a rational system of criminal justice?

Now, conversant with the difficulty of identifying persons from images, the dangers posed by context, cognitive bias and the cross-contamination of the different strands of the circumstantial case, we are in a position to reflect on what the prosecution of Yaryare ought to have looked like. If the police or prosecutor obtained the services of a genuine expert, then the circumstantial case against Yaryare would be based on the (strands of) evidence outlined in Table 3. Without the unbiased opinion of a genuine expert, the only admissible evidence against Yaryare is his association with “offenders” and lying about a friendship with a particular “offender”.

²¹³ Cf the opinions of DC B, P and Z, *Yaryare* (n 6). See David Faigman, John Monahan and Christopher Slobogin, ‘Group to Individual (G2i) Inference in Scientific Expert Testimony’ (2014) 81(2) *University of Chicago Law Review* 417.

²¹⁴ See David White et al, submission to Justice and Home Affairs Committee, *Human oversight of Facial Recognition Technology in forensic applications* (NTL0012, 20 October 2021).

²¹⁵ Clare Garvie, *A Forensic Without the Science: Face Recognition in U.S. Criminal Investigations* (Center on Privacy & Technology at Georgetown Law, 2022).

²¹⁶ See Richard I Kemp, Gary Edmond and David White, “A Proposed Solution to the Problem of Identifying People from CCTV and Other Images” in Andrew M Smith, Michael P Togliola and James Michael Lampinen (eds), *Methods, Measures, and Theories in Eyewitness Identification Tasks* (Routledge, 2021) 13.

Genuine expert's opinion + Association with "offenders" + Lie
 (not exposed to domain
 irrelevant information)

Table 3: Circumstantial case against Yaryare. The "Genuine expert's opinion" is italicised because it needs to be solicited and we do not know the conclusion.

In this prosecution, the speculative impressions of the non-expert investigator (DC B) are not adduced because they are irrelevant or their reliability cannot be demonstrated. They cannot be extricated from the evidence contaminating them, they cannot be managed (even with trial safeguards), and they cannot be sensibly evaluated in accusatorial proceedings. Their admission had the tendency to bias jury decision-making toward identifying Yaryare (and the other defendants) regardless of whether DC B's identification was correct.²¹⁷ Here, the imaging analyst is replaced by an expert whose ability is demonstrable, who uses a validated method, who does not work closely with DC B, and who is blinded to the anticipated result and any other information that is not required for her comparisons. We do not know what such an expert might report but her interpretation would be exclusively based on the images because nothing else would be available to her. Her opinion would embody what the images (rather than other information) are capable of supporting. In this scheme, jurors are not invited or enabled to undertake their own comparisons. Rather, they are presented with the considered opinion of a genuine expert, an indicative error rate, and the other (relatively limited) admissible evidence implicating Yaryare to consider and evaluate.

We can see how this case against Yaryare is more straightforward and does not require jurors to somehow unravel and evaluate contaminated strands of evidence or resist suggestive comparisons. In this prosecution there is no need to cross-examine in detail on bias, less chance of losing the character shield, and directions would be simpler. Each of the various strands of evidence appears only once and does not contaminate any of the others. This case may be less superficially persuasive than a prosecution where an investigator is allowed to express her categorical opinions about the identity of Person A, but it is much more likely to embody the evidential value of the images and to be fair(er). Any resulting conviction will be safe(r) because it will be based upon rational assessment of independent strands of evidence.

6. Conclusion

The foregoing analysis points to deficiencies of a fundamental kind with arrangements currently regulating the opinions of investigators and those called as experts. Drawing upon mainstream scientific research and advice, we have endeavoured to explain how rules of admissibility, Codes of Practice and policing guidance, and s 78, have not been developed or used in ways that are likely to produce fair trials or verdicts that are safe. Scientific ignorance has led lawyers and judges to misunderstand the independence, value and strength of evidence, and to over-estimate the efficacy of their own procedures and abilities.

Our study and analysis of *Yaryare* suggest that appellate courts are not being presented with the kinds of scientific knowledge required to fulfil their important constitutional responsibilities. Rules are applied, and evidence is represented and evaluated in ways that are incompatible with mainstream scientific understanding and advice. To that end, some

²¹⁷ That is those who did not directly witness the events. Here, police (and other investigators) are proscribed from expressing opinions on the identity of those whom they investigated.

of the practices outlined above, such as reliance on the contaminated opinions of an investigator and allowing the jury to make their own comparisons during the trial, are either irrational or have the propensity to encourage irrationality. They are apt to mislead fact-finders and judges, cannot be corrected through trial safeguards, and are much more likely to produce an identification than instil appropriate doubts, even when the wrong person is on trial.

Our courts must be much more interested in the effectiveness of their rules and procedures. This article might be understood as a warning against excessive reliance on common sense impressions, legal experience (individual or collective), tradition or authority, for legal proceedings to be fair and convictions safe. We have identified two species of genuine expert whose opinions could legitimately be sought and presented as forms of independent, expert evidence. Notwithstanding judicial declarations to the contrary, we have serious reservations about the outcome of *Yaryare* and other cases where the identification of defendants was dependent on the biased opinions of investigators and/or the contaminated comparisons by jurors. It is our contention that the public cannot be confident about such convictions or the institutions overseeing and upholding them.