INVESTIGATIVE INSIGHT TO ADMISSIBLE PROOF: BUILDING A RELIABILITY FRAMEWORK FOR BEHAVIOURAL EVIDENCE IN CRIMINAL COURTS

Kashish Anand, MIT- World Peace University, Pune

ABSTRACT

Behavioural evidence, drawn from victim selection, offender approach, control, concealment, staging, escalation, communication, post offence conduct, deception, linkage, and risk, plays an increasingly visible role in criminal investigation. Its investigative value is well documented, yet its evidentiary value is far less settled, and the two are routinely conflated in practice. This article argues that behavioural evidence requires a distinct reliability framework before it can move from investigative hypothesis to courtroom proof. Drawing on evidence law, forensic psychology, and the empirical literature on profiling accuracy and cognitive bias, the article traces the movement from behavioural clue to source, method, inference, corroboration, reliability, admissibility, and proof. It identifies thirteen diagnostic questions that judges, prosecutors, and defence counsel can apply to test behavioural evidence before it reaches the trier of fact. The article cautions against profiling overreach, diagnosis without assessment, character reasoning, predictive claims, and guilt by behavioural similarity, and proposes that behavioural evidence should assist, rather than substitute for, the court's own fact-finding function.

Keywords: behavioural evidence, offender profiling, expert testimony, admissibility, corroboration, reliability framework, criminal proof

Context And Rationale

Investigators have long used behavioural observations to narrow suspect pools, generate leads, and interpret ambiguous crime scenes. The way a body is positioned, the sequence of a burglary, the tone of a ransom-note, or the pattern across a series of assaults can all suggest something about the person who caused them. This form of reasoning, often labelled criminal profiling, behavioural crime scene analysis, or linkage analysis, has become a familiar feature of major case investigations (Turvey, 2011). Its appeal is understandable: it offers investigators a way to organise otherwise chaotic information and to prioritise limited resources.

Difficulty arises when this investigative reasoning is carried, largely unchanged, into the courtroom. A hypothesis that was useful for directing an inquiry is not automatically a reliable basis for establishing a fact in issue. Courts require more than plausibility; they require evidence that has been lawfully obtained, properly documented, methodologically sound, and tested against alternative explanations before it can support a finding of guilt. The empirical record on behavioural profiling is, at best, mixed. Meta-analytic reviews have found that self-identified profilers do not reliably outperform other groups across most predicted offender characteristics, and that the field's public reputation for accuracy rests more on selective anecdote than on demonstrated validity (Snook et al., 2007; Snook et al., 2008). This gap between investigative reputation and evidentiary substance is the central problem this article addresses.

Research Question and Aim

The article is organised around one question: how can behavioural evidence used during criminal investigation be assessed for reliability, corroboration, expert limits, and admissibility before it is presented as proof in criminal courts? The aim is not to dismiss behavioural evidence, nor to treat it as inherently illegitimate. The aim is to build a structured, defensible framework that separates behavioural material that has been tested and corroborated from behavioural material that remains, however compelling, an unproven hypothesis. The framework is intended for use by investigators preparing case files, prosecutors deciding what to tender, defence counsel testing what is tendered, and judges exercising a gatekeeping function.

Conceptual Lens

Three ideas anchor the analysis. First is the distinction between investigative utility and evidentiary sufficiency. A tool can be useful for narrowing an inquiry without being reliable enough to prove an element of an offence; the two standards answer different questions and should not be treated as points on the same scale. Second is the gatekeeping role assigned to trial judges under modern evidence law, most clearly articulated in the United States through Daubert v. Merrell Dow Pharmaceuticals, Inc. (1993) and its progeny, and in England and Wales through the common law assistance requirement established in R v Turner (1975). Both traditions ask, in different language, whether the proposed testimony rests on a sound method and whether it will genuinely help, rather than displace, the fact finder. Third is the recognition, developed extensively in the cognitive bias literature, that behavioural inference is not immune to contamination from irrelevant contextual information, and that experts who believe themselves objective can still be measurably influenced by extraneous case detail (Dror & Charlton, 2006; Dror & Kukucka, 2021). Together these ideas frame behavioural evidence as something that must earn its evidentiary status through demonstrated method and tested corroboration, not through the confidence or experience of the person presenting it.

Behavioural Evidence in Investigation

Behavioural evidence in investigation typically falls into several recognisable categories. Victim selection describes how and why an offender chose a particular target, which can narrow the pool of plausible suspects when access, opportunity, or relationship is at issue. Offender approach and control behaviour describe how the offender-initiated contact and maintained dominance during the offence, which can distinguish opportunistic from planned conduct. Concealment and staging describe efforts to hide evidence or to alter a scene to suggest a different sequence of events than actually occurred, and can point toward an offender with knowledge of the victim or scene. Escalation and communication describe changes in severity or contact across a series, relevant to risk assessment and to linking offences. Post offence conduct and deception describe behaviour after the crime, including statements to investigators, that may be probative of consciousness of guilt, though such inferences are notoriously prone to overreading. Linkage describes the attempt to connect separate offences to a common offender based on behavioural similarity, an area with a growing, though still qualified, empirical base (Woodhams et al., 2007; Bennell et al., 2014).

Each of these categories can be useful for directing investigative resources. None of them, without more, satisfies the requirements of proof. The observation that a burglar routinely disabled alarms before entry may correctly direct investigators toward offenders with a documented pattern; it does not, by itself, establish that the person in the dock is that offender. Investigative usefulness answers the question "where should we look?" Evidentiary sufficiency answers a harder question: "has it been shown, to the required standard, that this person did this thing?"

From Investigative Insight to Courtroom Proof

The movement from insight to proof passes through several distinct stages, and behavioural evidence commonly fails to complete the journey. Insight begins as an observation: something in the evidence appears patterned, distinctive, or unusual. This observation generates a hypothesis: a tentative explanation of what the pattern might mean about the offender. The hypothesis becomes evidence only once it is anchored to a documented source, examined through a defensible method, and tested against competing explanations. Reliability is the property that emerges when the method has been shown, ideally through independent validation, to produce consistent and accurate results. Admissibility is the legal threshold applied by the court, informed by reliability but also by relevance, qualification, and the risk of unfair prejudice. Proof is the final status behavioural material can achieve, but only once it has been corroborated by other evidence in the case and accepted by the fact finder as sufficiently probative in context.

At each stage, something can go wrong. An observation may not be accurately documented at the point of collection. A hypothesis may harden into an assumption before it has been tested. A method may lack any track record of independent validation. An expert may exceed the boundaries of demonstrated competence. Corroboration may be assumed rather than actually present in the case file. The framework proposed later in this article is designed to catch failures at each of these points, rather than waiting until behavioural material has already been placed before a jury.

Reliability, Method, and Expert Opinion

Two separate reliability questions arise once behavioural material is proposed for trial: whether the underlying method has any demonstrated validity, and whether the particular witness is qualified to apply it. On the first question, the empirical literature gives reason for caution. Systematic reviews of profiling accuracy have found that professional profilers do not consistently outperform comparison groups such as experienced detectives, and in some studies do not outperform lay participants, across predictions of offender age, appearance, criminal history, or cognitive processes (Snook et al., 2007). The authors of one influential review describe the widespread belief in profiling's effectiveness as resting substantially on anecdote, selective reporting of successes, and the persuasive effect of an expert label, rather than on controlled demonstration (Snook et al., 2008). This does not mean every behavioural inference is worthless; it means that global claims of expertise cannot substitute for evidence that a specific method, applied to a specific category of behaviour, has been tested and found reliable.

Linkage analysis illustrates a more favourable, though still qualified, picture. Research testing the consistency and distinctiveness assumptions that underpin case linkage has found meaningful, above chance discrimination in several offence categories, particularly for serious interpersonal crime such as sexual assault, though accuracy varies by crime type, dataset, and the specific behavioural variables used (Woodhams et al., 2007; Bennell et al., 2014; Woodhams & Labuschagne, 2012). The appropriate response is neither blanket acceptance nor blanket rejection of linkage evidence, but a demand that any party tendering it identify which specific technique was used, on what dataset it has been validated, and what error rate is associated with it in comparable conditions.

On the second question, qualification is not established by years of service or the label "profiler." Rule 702 of the Federal Rules of Evidence requires the trial court to find that testimony is based on sufficient facts, is the product of reliable principles and methods, and that those principles have been reliably applied to the facts of the case (Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993; Kumho Tire Co. v. Carmichael, 1999). English courts apply a related but distinct test, asking whether the proposed evidence would genuinely assist a jury on a matter outside ordinary experience, as established in R v Turner (1975), where psychiatric evidence about a defendant's likely emotional reaction was excluded because ordinary jurors could assess normal human reactions without expert assistance. Later cases have relaxed the rigidity of that assumption in narrow contexts, but the underlying principle endures: expert opinion is admitted to add something jurors cannot supply themselves, not to substitute the expert's judgment for the jury's (Mackay & Colman, 1991).

A further, increasingly documented risk is contextual contamination. Analysts who know that a suspect has confessed, has a prior record, or is otherwise disfavoured by the investigation can unconsciously interpret ambiguous behavioural evidence consistently with that prior belief. Experimental work has shown measurable shifts in expert judgment when irrelevant case context is introduced, even among experienced examiners who did not believe themselves susceptible to influence (Dror & Charlton, 2006). Procedural safeguards such as sequential unmasking, in which analysts are shielded from extraneous information until their independent judgment is recorded, have been proposed as practical mitigations (Dror & Kukucka, 2021).

Corroboration and Alternative Explanations

Behavioural evidence rarely stands alone, and it should never be asked to. Corroboration means independent evidence, physical, documentary, or testimonial, that supports the same conclusion the behavioural inference points toward. A pattern of staging observed at a crime scene becomes more probative when it aligns with forensic evidence of post mortem movement, timeline inconsistencies established through phone or location records, or witness accounts that independently confirm access and opportunity. Without such corroboration, a behavioural inference remains a hypothesis dressed in expert language.

Testing alternative explanations is the necessary companion to corroboration. Any behavioural pattern consistent with guilt should be checked against equally plausible innocent explanations. Concealment behaviour consistent with a controlling offender may also be consistent with panic, third party interference, or an entirely different sequence of events reconstructed after the fact. Escalation across a series may reflect a single offender's changing behaviour, or it may reflect the entry of a second offender into what was mistakenly treated as one series. A reliability framework that does not require alternative explanations to be actively sought and documented, rather than merely considered and dismissed, will tend to confirm whatever hypothesis the investigation formed earliest, a pattern consistent with the broader literature on confirmation bias in expert decision making (Dror & Kukucka, 2021).

Admissibility in Criminal Courts

Admissibility decisions rest on more than reliability alone. Relevance is the threshold requirement: the evidence must make a fact in issue more or less probable than it would be without it. Qualification requires that the specific witness has the training and demonstrated competence to apply the method in question, not merely general experience in an adjacent field. Legal relevance also requires the court to weigh probative value against the risk of unfair prejudice, undue consumption of time, or confusion of the issues, a balancing exercise codified in the United States under Federal Rule of Evidence 403 and reflected in comparable exclusionary discretion in other common law jurisdictions.

Behavioural evidence carries a distinctive prejudice risk because it can shade into character or propensity reasoning, inviting the fact finder to conclude that a person who exhibits certain behavioural traits is the kind of person who commits this kind of crime. Reasoning of that kind is generally restricted under character evidence rules, reflected in Federal Rule of Evidence 404, precisely because it risks convicting a defendant for their perceived type rather than for the specific conduct charged. Courts in several jurisdictions have expressed sustained scepticism toward profiling testimony offered as substantive proof rather than investigative background, with a pattern noted among state supreme courts favouring exclusion where such evidence suggested that a defendant matched a general offender type (State v. Fortin, 2007). Legal scholarship reinforces this scepticism on probabilistic grounds, arguing that evidence a defendant matches a broad behavioural or demographic profile is evidence about a category of people, not case specific evidence that the particular defendant committed the particular offence, and that the two should not be treated as interchangeable in a probabilistic assessment of guilt (Di Bello et al., 2024).

Behavioural evidence therefore faces a demanding, but not impossible, admissibility path. It is more likely to survive scrutiny when it is offered to explain investigative steps, to support or rebut a narrow and specifically tested hypothesis such as case linkage, or to assist the fact finder in understanding evidence genuinely outside ordinary experience, and it is more likely to fail when it is offered as a general character portrait meant to persuade the fact finder that the defendant fits the profile of an offender.

Reliability Framework for Behavioural Evidence

The framework proposed here asks thirteen questions of any behavioural material before it is treated as more than an investigative lead.

1. What is the behavioural observation, stated precisely and without inferential language. 2. What is its source, including who observed it, when, and under what conditions. 3. Is it documented contemporaneously, in a form that can be independently reviewed. 4. Is it relevant to a fact actually in issue in the proceeding. 5. Is there a method behind the inference that has a demonstrated track record, rather than an expert's unstructured impression. 6. Is expert opinion actually required, or could the fact finder assess the material unaided. 7. Is the expert qualified in the specific technique proposed, not merely experienced in an adjacent field. 8. What are the limits of the inference, stated affirmatively rather than left implicit. 9. What corroborates it, identified as specific, independent evidence in the case file. 10. What alternative explanations exist, and have they been actively investigated rather than dismissed. 11. What is its legal value, meaning its probative contribution to an element of the offence. 12. Could it unfairly prejudice the accused by inviting character or propensity reasoning. 13. Does it assist the court without displacing the court's own fact-finding function.

These questions are cumulative rather than a checklist to be passed in isolation. Behavioural material that answers most questions well but fails on documentation, for example, should not be treated as reliable simply because the underlying method is sound; a well validated technique applied to poorly recorded observations produces an unreliable result regardless of the method's general standing.

Risks, Limit and Misuse

Several recurring misuses justify the caution built into this framework. Profiling overreach occurs when a general offender type, built from group level statistical tendencies or unstructured clinical impression, is presented as though it describes a specific individual with confidence the underlying data cannot support. Diagnosis without assessment occurs when behavioural observations from a crime scene are used to assign a psychiatric or personality classification to a person the expert has never examined, a practice that conflates behavioural description with clinical diagnosis and exceeds what crime scene evidence alone can establish. Character reasoning occurs when behavioural similarity is used to suggest that the defendant is the sort of person who commits this sort of crime, a form of propensity reasoning the law generally restricts. Predictive claims occur when an expert asserts likely future conduct with a precision the underlying research does not support. Guilt by behavioural similarity occurs when a defendant's match to a general pattern is treated as equivalent to case specific proof of the offence charged, a conflation legal scholar has specifically identified as probabilistically unsound (Di Bello et al., 2024).

The empirical record on profiling accuracy provides an additional caution: even well-intentioned experts operating in good faith have been shown, across multiple studies, not to reliably outperform comparison groups on many of the predictions profiling methods claim to generate (Snook et al., 2007; Snook et al., 2008). This finding does not license dismissal of all behavioural analysis; it licenses insistence that specific claims be tied to specific, validated methods rather than to the general reputation of the discipline or the confidence of the witness presenting it.

Conclusion

Behavioural evidence occupies a legitimate and valuable place in criminal investigation, but its investigative value does not travel automatically into the courtroom. The movement from behavioural clue to admissible proof passes through identifiable stages: source, method, inference, corroboration, reliability, admissibility, and finally proof, and behavioural material can fail at any one of these stages even where it succeeds at the others. Courts, prosecutors, and defence counsel are better served by a structured reliability framework that interrogates documentation, method, qualification, corroboration, alternative explanation, legal relevance, and prejudice risk before behavioural material is allowed to influence a finding of guilt, than by treating investigative usefulness as a proxy for evidentiary sufficiency. Applied consistently, such a framework allows behavioural evidence to assist the court's fact-finding function without displacing it, preserving both the investigative value of behavioural analysis and the evidentiary integrity the criminal trial requires.

References

Bennell, C., Mugford, R., Ellingwood, H., & Woodhams, J. (2014). Linking crimes using behavioural clues: Current levels of linking accuracy and strategies for moving forward. *Journal of Investigative Psychology and Offender Profiling, 11*(1), 29-56. https://doi.org/10.1002/jip.1395

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

Di Bello, M., Cangiotti, N., & Loi, M. (2024). *Profiling vs. case-specific evidence: A probabilistic analysis* [Preprint]. arXiv. https://arxiv.org/abs/2603.00098

Dror, I. E., & Charlton, D. (2006). Why experts make errors. *Journal of Forensic Identification, 56*(4), 600-616.

Dror, I. E., & Kukucka, J. (2021). Linear Sequential Unmasking-Expanded (LSU-E): A general approach for improving decision making as well as minimizing noise and bias. *Forensic Science International: Synergy, 3*, 100161. https://doi.org/10.1016/j.fsisyn.2021.100161

Federal Rules of Evidence, Rule 403 (Excluding Relevant Evidence for Prejudice, Confusion, Waste of Time, or Other Reasons).

Federal Rules of Evidence, Rule 404 (Character Evidence; Crimes or Other Acts).

Federal Rules of Evidence, Rule 702 (Testimony by Expert Witnesses), as amended 2000, 2011, 2023.

Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999).

Mackay, R. D., & Colman, A. M. (1991). Excluding expert evidence: A tale of ordinary folk and common experience. *Criminal Law Review*, 800-810.

R v Turner [1975] QB 834 (Court of Appeal, England and Wales).

Snook, B., Cullen, R. M., Bennell, C., Taylor, P. J., & Gendreau, P. (2008). The criminal profiling illusion: What's behind the smoke and mirrors? *Criminal Justice and Behaviour, 35*(10), 1257-1276. https://doi.org/10.1177/0093854808321528

Snook, B., Eastwood, J., Gendreau, P., Goggin, C., & Cullen, R. M. (2007). Taking stock of criminal profiling: A narrative review and meta-analysis. *Criminal Justice and Behaviour, 34*(4), 437-453. https://doi.org/10.1177/0093854806296925

State v. Fortin, 917 A.2d 746 (N.J. 2007).

Turvey, B. E. (2011). *Criminal profiling: An introduction to behavioural evidence analysis* (4th ed.). Academic Press.

Woodhams, J., Hollin, C. R., & Bull, R. (2007). The psychology of linking crimes: A review of the evidence. *Legal and Criminological Psychology, 12*(2), 233-249. https://doi.org/10.1348/135532506X118631

Woodhams, J., & Labuschagne, G. (2012). A test of case linkage principles with solved and unsolved serial rapes. *Journal of Police and Criminal Psychology, 27*(1), 85-98. https://doi.org/10.1007/s11896-011-9091-1