Dissecting a Criminal Investigation

D. Kim Rossmo[1] ⬡

Abstract

Despite the considerable attention criminologists devote to the study of policing, they tend to ignore one of its most important functions. Detectives comprise approximately 16% of law enforcement personnel and play a major role in the public's image of the police through their successes or failures. This scholarly lacuna is even more surprising given the gateway position held by police investigators; unless a crime is solved and an individual arrested, the entire remainder of the criminal justice system—prosecutors, defense attorneys, judges, juries, probation, prisons, parole, rehabilitation—fails to come into play. What research that has explored this function has been primarily interested in organizational and technical aspects of detective work. Here, I take a different approach by dissecting criminal investigations to expose their underlying structure—what they involve, how they fail, and ways they might be improved. Specific areas of interest include the functional phases of an investigation, the definition and nature of evidence, and the systemic structure of criminal investigative failures. The Gail Miller-David Milgaard murder investigation in Saskatoon, Saskatchewan, is used as a case study.

Keywords Criminal investigations · Detectives · Policing · Criminal investigative failures · Wrongful convictions

Introduction

In the popular conception of a criminal investigation, an intrepid detective discovers a number of puzzling clues, logically analyzes the evidence, and brilliantly exposes the murderer. The iconic image is that of Sherlock Holmes, with his magnifying glass, following a series of footsteps to the dénouement of the mystery.[1] Or C. Auguste Dupin, in *The Murders in the Rue Morgue*, reading his associate's mind by rationally tracing his line of thought over the previous fifteen minutes.

While real-world investigations are typically less dramatic and more chaotic than fictional detective stories, both require a framework centered on information (Willmer 1970). The clues methodically uncovered by Holmes, Dupin, and their colleagues possess an information content. Investigators are particularly interested in a special type of information known as evidence. Flowing from the crime scene to the arrest to the trial, evidence is the lifeblood of any investigation.

Previous research found that many police managers, along with many academics, did not properly understand the nature of the investigative function (Eck and Rossmo 2019; Horvath et al. 2003). In response to this void, the present article seeks to provide novel and useful insights by dissecting the anatomy of criminal investigations—the structure of criminal evidence, police detective work, and investigative failures. Most investigations are short and perfunctory; the offender is identified by the victim, a suspect is arrested by patrol officers, or an administrative decision is made to stop pursuing the case because of a lack of leads (Chaiken et al. 1976). The focus in this article is on criminal investigations of substance and complexity, often involving true mysteries[2] and "whodunnits."

✉ D. Kim Rossmo
krossmo@txstate.edu

1 Center for Geospatial Intelligence and Investigation, School of Criminal Justice and Criminology, Texas State University, 601 University Drive, San Marcos, TX 78666, USA

---

[1] As Sherlock Holmes explained to Dr. Watson in *A Study in Scarlet*, "There is no branch of detective science which is so important and so much neglected as the art of tracing footsteps." This novel also fictionally debuted the magnifying glass as an investigative tool. In later stories, in addition to human footsteps, the famous detective followed horse, cow, hound, carriage, and bicycle tracks.

[2] While such cases are typically murders, the investigative issues discussed here are general in nature and apply to most crime types. The term detective is used throughout this article, but it is recognized that investigative work is done by a variety of police ranks and positions (such as patrol officers), depending on the specific law enforcement agency and its jurisdiction.

Table 1 Criminal investigative phases

Evidence	Suspects			Offender
	Generate	Prioritize	Assess	
Collection	Records	Description	Motive	Witness
Evaluation	Public	Behavior	Means	Confession
Analysis		Geography	Opportunity	Physical evidence

Functional Phases of an Investigation

If evidence is the lifeblood of an investigation, the body it flows through consists of three different functional phases—obtaining evidence, processing suspects, and convicting the offender (see Table 1). The first phase includes evidence collection, evaluation, and analysis. The second phase can be divided into subparts involving suspect generation, prioritization, and assessment. The third phase uses case evidence to identify the offender and prove his or her guilt—a goal that can only be accomplished through a witness, a confession, or physical evidence.

While Table 1 presents these phases in chronological order, in actuality they overlap and may even operate independently of each other. Ideally, detectives would first find and process all the available evidence connected to a crime and carefully analyze its implications before moving on to suspects. However, the chaos and attention surrounding major crimes often result in a torrent of information and tips flowing into the police long before all the witnesses have been interviewed, the neighborhood canvassed, and the lab results returned (Rossmo 2020). Detectives may also suffer from premature judgment and tunnel vision, concluding a particular suspect is guilty prior to considering all the evidence in a case. The premature shift from an evidence-based to a suspect-based investigation has led to a number of wrongful convictions (Rossmo and Pollock 2019).

Suspect and offender phases are sometimes temporally reversed. For example, detectives might have a good idea of who the offender is but lack the evidence necessary for an arrest. Alternatively, they may have the ability to establish guilt but not know who the offender is. This last scenario occurs when there is forensic evidence such as DNA that definitely links to an individual, but that person is unknown and therefore must first be found before a comparison is possible. Collecting, prioritizing, and assessing suspects then become the investigative focus.

Research has shown that most crimes are solved by information obtained from the public (Chaiken et al. 1976; see also Leovy 2015). Police interview witnesses, informants, family and friends of the victim, neighbors in the area of the crime scene, ask for tips, and question informants

in an effort to generate suspects. Sex offender registries, parolee lists, and other police and criminal databases are searched, along with motor vehicle department and various government record systems.

For serious stranger crimes, such as murder or rape, it is not uncommon for investigations to suffer from information overload problems after amassing hundreds or even thousands of suspects; prioritizing suspects then becomes necessary (Rossmo 2009). However, there are only a limited number of methods for doing so. A generic physical description (race, sex, height, weight, hair color) is one option, though this requires a witness to the crime. Similarities between known offender *modus operandi* and crime scene behaviors provide another approach. Finally, the ubiquity of addresses in various databases makes geography a useful method. A behavioral or geographic profile specific to the crimes under investigation can increase the precision of these last two techniques.

After suspects have been generated and prioritized, investigators have to assess them – or at least the top prioritized. This task is significantly more time-intensive than prioritization. Suspects have to be interviewed, along with their families, friends, and neighbors, alibis checked out, *modus operandi* compared, and so on. A suspect can also be assessed on the basis of motive, means, and opportunity. These three classic crime requirements directly relate to the basic 5Ws + 1 H of investigations: suspects (who) can be assessed by considering motive (why), means (what, how), and opportunity (where, when).

Proving guilt in a criminal trial can only be accomplished through a witness, a confession, or physical evidence[3] (Klockars and Mastrofski 1991). This requirement connects back to the first, evidence-focused phase of an investigation. The dynamics of the interconnected phases depicted in Table 1 are outlined in the crime-evidence-suspects-offender (cESO) schema shown in Fig. 1.

- A crime is discovered.
- Police search for and collect evidence from the crime scene, which may lead to new investigative avenues and additional evidence. All evidence needs to be evaluated and analyzed.
- Guided by the evidence, investigators generate suspects, prioritize them, and then conduct assessments of the most likely possibilities.
- Identifying the offender requires "proof beyond a reasonable doubt" (Boyd 2019). Such proof is established

---

[3] Evidence is considered here in the context of investigative information and police decision-making; the various types of legal evidence, while related, are not always directly equivalent (Ho 2015).
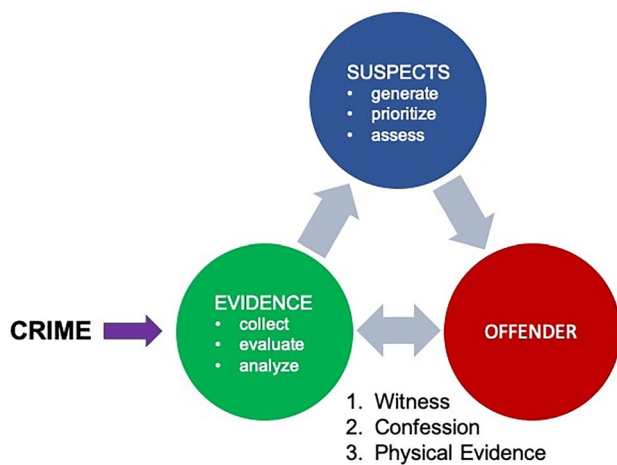
Fig. 1 Crime-Evidence-Suspects-Offender (cESO) schema

through evidence in the form of a witness, a confession, and/or physical evidence.

- This process of identifying the offender and establishing guilt operates in both directions. Evidence found at the crime scene (e.g., fingerprints) or from follow-on investigative efforts may lead to a suspect; alternatively, evidence obtained from investigating a suspect (e.g., a confession) may link back to the crime.

Evidence Structure

Information—intelligence and evidence—is needed to strategically guide the collection, prioritization, and evaluation of suspects. It is important to understand what distinguishes evidence from intelligence (Rossmo 2020). Evidence is a recorded fact relevant to the crime, the origin of which can be identified (e.g., witness statement, crime scene photograph, laboratory report, etc.). Theories, assumptions, and intuition are not evidence. Intelligence is a broader category of information that may not rise to the standard of evidence or be admissible in court. While criminal intel is often useful in the early stages of an investigation to help guide police decision-making, evidence is required to legally solve a crime.

As evidence is collected, suspicion may build around a particular suspect. If this suspicion becomes strong enough, the suspect will be arrested. Conversely, a suspect may be cleared if exculpatory evidence, such as an alibi witness, is uncovered. New evidence may also point towards an alternative suspect. Investigations typically assemble multiple items of evidence, some confirming and others conflicting; it is therefore necessary to know how to integrate different items of evidence to build a rational basis for judgment. At some point, the investigation will

hopefully reach an approximation of a stable probability of guilt regarding the correct suspect (Cox 1946).

Bayesian Inference

Bayes' theorem provides a method for updating the probability of a hypothesis following the addition of new evidence (Eddy 2004; Iversen 1984). The theorem is

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

where:

$P(A|B)$ = probability of event A given event B
$P(B|A)$ = probability of event B given event A
$P(A)$ = probability of event A
$P(B)$ = probability of event B.

In Bayesian inference, the updated (posterior) probability is calculated by (1) first converting the initial probability (the base rate) to a prior odds (likelihood) ratio, (2) multiplying the prior odds by the evidence likelihood ratio to calculate the posterior likelihood ratio, and (3) converting these posterior odds to a posterior probability:

$$LR_{prior} = \frac{P_{prior}}{1 - P_{prior}}$$
$$LR_{post} = LR_{prior} \times LR.$$
$$P_{post} = \frac{LR_{post}}{1 + LR_{post}}$$

where:

P_{prior} = prior probability
LR_{prior} = prior odds ratio
LR = evidence likelihood ratio
LR_{post} = posterior odds ratio
P_{post} = posterior probability.

When a new item of evidence is added, the posterior probability becomes the new prior probability, and so on. The evidence likelihood ratio is formed from the quotient of the conditional probability of the evidence given a hypothesis (e.g., a suspect is guilty) to the conditional probability of the evidence given the compliment of the hypothesis[4] (e.g., a suspect is innocent):

$$LR = \frac{P(E|H_1)}{P(E|H_0)}$$

where:

---

[4] Equation 1 defines the positive likelihood ratio, which is applied when evidence is present. The negative likelihood ratio is used when evidence is absent. The two ratios are related as follows: LR+ = sensitivity/(1 − specificity), while LR− = (1 − sensitivity)/specificity.

LR = likelihood ratio.

$P(E|H_1)$ = probability of the evidence given hypothesis H_1

$P(E|H_0)$ = probability of the evidence given hypothesis H_0 (the compliment of hypothesis H_1).

Evidence has two structural properties that determine its importance in an investigation—strength and reliability. The stronger the evidence, the more it corroborates the guilt (or innocence) of a suspect or supports a particular theory of the crime, in comparison with other suspects or theories. Reliability is the accuracy or truthfulness of the evidence. Even strong evidence has little probative value if it is wrong.

Strength

The likelihood ratio provides a formal measure of evidence strength. This quantity is known as information gain (Wells and Olson 2002). A likelihood ratio > 1 means the evidence contributes to a conclusion of guilt, while a likelihood ratio < 1 means the evidence contributes to a conclusion of innocence; the larger or smaller the likelihood ratio, the greater that contribution (Robertson and Vignaux 1995). A likelihood ratio of 1 supports neither guilt or innocence (in other words, it is neutral and has no impact on either hypothesis).[5] The numerator, which varies between 0 and 1, can only reduce the probability of guilt, while the inverse of the denominator, which varies between 1 and ∞, can only increase the probability of guilt. It is therefore not how well the evidence matches the suspect that determines guilt—though a low match can establish innocence—but rather how unlikely it matches anyone else.

With a few exceptions, such as DNA, it is not possible to precisely determine evidence probabilities. However, it is still important to understand the different contributions made by the numerator and the denominator in the equation. For example, suppose a witness describes a burglar as a white male wearing a black T-shirt and jeans. Police then see someone who fits this description. The description match is evidence, but how strong is it?

The first part that needs to be determined is the probability of the evidence (description) given that the suspect is guilty. All parts of the description match; however, unlike sex and race, people may change their clothing. The value of the numerator is therefore somewhat < 1 (its maximum potential value).

The second part that needs to be determined is the probability of the evidence (description) given that the suspect is innocent; in other words, how likely is it that police might come across a white male in a black T-shirt who had nothing to do with the robbery? The answer is quite probable, given the demographic frequency of white males, the prevalence of black T-shirts and jeans, and the many individuals police encounter on a daily basis. Consequently, the denominator is reasonably high. Overall, the likelihood ratio in this scenario would be slightly over 1 (i.e., weak). While the police may want to question the individual, the reasonable and probable grounds for an arrest are lacking.

Now suppose the burglar was described as a white male wearing a purple Berlin Philharmonic T-shirt and jeans, with a Sasquatch tattoo on his right forearm. The police stop a similar suspect. In this case, the denominator probability is small because Berlin Philharmonic T-shirts are not common (at least in North America) and arm tattoos of Sasquatches even less so. The low denominator results in a high likelihood ratio, meaning the evidence is much stronger than in the first scenario, and justifies a detention.

It is necessary to remember that both the numerator and the denominator influence the strength of the evidence. In some wrongful conviction cases, detectives made the mistake of focusing only on suspect-offender similarities, while failing to consider how common the evidence was overall. A perfect match of a suspect to the evidence is only meaningful if the evidence is rare in the general population. This is the reason DNA is powerful evidence; it typically involves extremely low random match probabilities that establish guilt beyond any reasonable doubt.

Reliability

The strength of even the strongest evidence is eroded if it is not reliable. A number of factors can undermine reliability; for example, a witness might have something to gain, hold animosity towards the accused person, been drunk at the time, and so forth. After Thomas Sophonow's arrest for the murder of a teenage girl in Winnipeg, 11 different jailhouse informants rushed to volunteer their services to police in the hopes of having their current charges dismissed. Crown counsel chose to call the three "best" of these to testify at trial. One, who had a conviction for perjury, had testified as a jailhouse informant in nine other cases. A provincial inquiry later drily noted: "He seems to have heard more confessions than many dedicated priests" (Cory 2001).

Confessions can also suffer from reliability problems. Detectives coerced a confession from 14-year-old Michael Crowe for the murder of his sister, Stephanie, in Escondido,

---

[5] A likelihood ratio is a measure of evidence diagnosticity. During the trial of Guy Paul Morin for the murder of his neighbor, Christine Jessop (see below), Crown counsel argued that Morin's failure to attend Jessop's funeral was evidence of consciousness of guilt. However, his attendance could have equally been regarded as *indicative* of guilt; police commonly monitor—as they did in the Jessop case—those who attend a murder victim's funeral in case the killer shows up. A provincial inquiry later noted, "Mr. Morin's failure to attend the funeral or funeral home was worthless evidence and ought not to have been admitted" (Kaufman 1998, p. 34). As the numerator and denominator probabilities for the strength of this evidence were similar, it had no diagnosticity.

California. He was grilled for hours, promised leniency if he confessed, and lied to about the recovered evidence in the case (McCrary 2009). Michael's "confession" included such statements as: "I'm so sorry that I can't even remember what I did to you…. I never meant to hurt you and the only way I know I did it is because they told me I did." Despite such dramatic warnings of unreliability, police charged Michael with murder. He was later released after blood discovered on the clothing of a homeless man who had been trying to enter houses in the neighborhood the night of the murder was matched through DNA to the victim.

Finally, certain forensic science techniques have come under scrutiny for exaggerated claims, problematic analyses, and the use of outright junk science (National Research Council 2009). Guy Paul Morin was accused of murdering nine-year-old Christine Jessop who lived next door to him in Queensville, Ontario (Kaufman 1998). Forensic scientists testified at trial that hairs and fibers seized from Morin matched those recovered from the murder scene. However, this claim was overstated and the forensic analyses were unreliable. Moreover, the fact that some of the evidence had been contaminated was not communicated to either police investigators or Crown counsel. Morin was convicted but later exonerated by DNA testing of semen stains found on the victim's underwear. The Ontario Centre of Forensic Sciences subsequently underwent a major overhaul as a result of the scandal. In 2020, a genealogical DNA test identified the real offender.

The relationship between the influence of evidence and its reliability can be expressed as follows:

$$P_{\text{post}} = r P_{\text{postE}} + (1 - r) P_{\text{prior}}$$

where:

P_{post} = final posterior probability of the hypothesis adjusted for the reliability of the evidence

R = reliability of the evidence (expressed as a probability)

P_{prior} = prior probability of the hypothesis (before the evidence)

P_{postE} = posterior probability of the hypothesis (after the evidence, assuming its complete reliability).

Assessing the reliability of trustworthy witnesses and valid forensic techniques is also necessary as honest mistakes are possible (and even probable). For example, a witness might describe a criminal as younger than he actually is, or a pathologist could estimate the time of death in a murder as later than it really occurred. If a description involves a range, say "25 to 35 years of age" versus "late 20s," then it will usually be more accurate

(reliable) than a point estimate, albeit, at the cost of precision and investigative utility.

Strength, Reliability, Independence, and Patterns

In addition to strength and reliability, evidence must also be independent as derivative evidence does not contribute anything new to an investigation. Finally, it is important to understand how all the evidence fits together into a holistic pattern; cherry picking must be avoided. However, evidence integration often suffers from imbalance. People tend to place disproportionate importance on strong evidence even when its reliability is low (Stubbins and Stubbins 2009). "The extensive experimental literature on judgment under uncertainty indicates that people do not combine strength and weight [reliability] in accord with the rules of probability and statistics" (Griffin and Tversky 1992, p. 413). Initially, people focus on how strong the evidence is, and then make some adjustment in response to reliability. But that adjustment is usually insufficient, so evidence strength tends to dominate (Tversky and Kahneman 1974). The focus on evidence strength also leads to the deemphasis of base rates. As explained below, if detectives suffer from confirmation bias, there is a risk that evidence supporting the dominant investigative theory will be overemphasized while conflicting evidence will be downplayed.

In summary, detectives need to consider questions of strength, reliability, independence, and patterns (SRIP) to fully understand the probative value of an item of evidence (Eck and Rossmo 2019):

- strength (strength of the evidence = ratio of guilt to innocence probabilities)
- reliability (accuracy/truthfulness of the evidence)
- independence (is it a unique contribution?)
- pattern (holistic consideration of all the information— avoid cherry picking).

Criminal Investigative Failures

Understanding the anatomy of a criminal investigation helps to both focus police efforts and identify risks of failure. While evidence flows through all phases of an investigation, it takes on varying complexions and vulnerabilities at different points. Problems with evidence at any stage can lead to an investigative failure. It should be noted, however, that not all unsolved crimes are investigative failures as the tactical decisions of

the offender also influence the probability of solving a crime (Beauregard and Martineau 2014; James and Beauregard 2020). Police cannot recover evidence that does not exist in the first place.

A detective must conduct a thorough investigation while avoiding a number of potential traps. There are three types of criminal investigative failure: (1) ignored crimes[6], (2) unsolved crimes that should have been cleared, and (3) wrongful convictions (Rossmo 2009). While the second type of failure is the most common, the third is the most damaging. All three failure types share common causes (Rossmo and Pollock 2019).

Using a scheme derived from Reason's (1990) failure domains, Rossmo and Pollock (2018, 2019) analyzed in detail a number of criminal investigative failures ($N = 50$) for murder and rape/sexual assault, from 1969 to 2014, most of which were wrongful convictions; 84% of these cases occurred in the USA, 10% in Canada, and 6% in Europe.[7] Causal factors were identified for each failure and categorized as personal, organizational, or situational in nature. Personal problems were within the individual control of a detective (e.g., rush to judgment, tunnel vision, logic errors). Organizational issues were within the control of the police agency (e.g., lack of resources, groupthink, interagency communication failures). Situational influences were outside the control of the criminal justice system (e.g., intense media attention, deceitful witnesses). The causal factors were also analyzed by proximity and interconnection. Finally, concept maps were constructed for each case outlining the various causal factors and their relationships with each other.

Of the 363 identified causal factors ($\bar{x} = 7.3$), 61% were personal in nature, 21% were organizational, and 18%

were situational. These were coded into 40 categories, the top eight (20%) of which accounted for half of all causes: confirmation bias, tunnel vision, intense media attention, management/supervision issues, careless/incompetent investigation, improper interrogations, rush to judgment, and improper forensics. These were further grouped into nine meta-categories, the most dominant of which involved cognitive biases. Particular causal factor combinations tended to cluster together; for instance, high-profile crimes often caused a rush to judgment, followed by the premature shift from an evidence-based to a suspect-based investigation. These problems then led to tunnel vision and confirmation bias, ultimately leading to evidence failures (Rossmo and Pollock 2019).

Rush to Judgment

Horrific crimes and high-profile cases put great pressure on police to achieve a timely resolution (O'Brien 2009). Detectives may jump to conclusions before their investigation is complete because of public fear, media attention, organizational stress, or a compelling desire to apprehend a dangerous criminal. Strong evidence that emerges early in an investigation can also produce a rush to judgment, even if its reliability is uncertain. In such situations, detectives risk prematurely moving from an evidence-based to a suspect-based investigation (Stelfox and Pease 2005). In the former mode, they are still gathering information to establish what happened during the crime and to help identify possible suspects. In the latter mode, they have concluded they know who the offender is and their investigation shifts to gathering evidence for the prosecution.

Premature judgment[8] may also be influenced by intuition, or what is commonly known as "gut instinct" (Wright 2013). We use two types of decision-making—the rational and the intuitive (Kahneman 2003). Intuition is automatic, fast, and powerful, and often involves cognitive shortcuts or heuristics (Risinger and Loop 2002). However, intuition is based on experience and takes time to properly develop (Gilovich et al. 2002; Myers 2002). The process is implicit and therefore difficult to control, and may often be wrong (intuition exists for our survival, not for its accuracy). It is most useful in stable environments with consistent rules or when data are missing (Klein 2003). But intuition is not a replacement for the rational analysis and logic required in complex tasks such as those in criminal investigations or legal prosecutions.

---

[6] Ignored crimes are those police fail to recognize as such despite the extant evidence due to incompetence or negligence. Theresa Allore, a 19-year-old student, disappeared in November 1978 from the small rural village of Compton in the Eastern Townships of Québec where she had been attending college. Five months later, Allore's body, clothed only in her underwear, was found lying face down in a creek far from the road a mile outside of Compton (Allore and Pearson 2009). Concerned about protecting the reputation of the local college, police wrote off the 19-year-old's death as an accidental drug overdose, despite a lack of pathological evidence to support their theory. Her death is now considered a murder, possibly linked to other similar homicides of young women in the region. However, the case remains unsolved because of the failure to properly collect and retain evidence from the original scene. In the mid-1990s, a large number of street sex trade workers reported missing from Vancouver's Downtown Eastside were written off by the police department's major crime inspector as simply a matter of the women leaving town. The Vancouver missing women investigation eventually metastasized into the Pickton Pig Farm serial murder case, the worst in Canadian history (Oppal 2012).

[7] While some jurisdictional differences exist in investigative procedures (mainly legal in nature), their underlying structures, the nature of their failures, and how they might be improved are similar.

[8] A rush to judgment, if accurate, does not necessarily lead to failure by itself; however, it can become a problem if it results in confirmation bias because the detective is now less willing to change his or her opinion; consequently, case evidence may not be properly evaluated.

Major crime investigations are complicated processes that, unlike their fictional depictions, often progress in a chaotic and confusing manner. They may suffer from information overload, high noise-to-signal ratios, and red herrings, all of which require time and effort to evaluate and sort (Doney 1990). The temporal order of evidence discovery is usually unrelated to the actual timeline of the crime; items are found after the initial crime scene search, witnesses come forward days after the incident, and laboratory resource limitations delay forensic testing. This haphazard information flow is not conducive to logical analytic thinking. Moreover, the random order and time lags can contribute to the risk of thinking errors.

Strong "convincing" evidence uncovered early in an investigation may distort evaluations of subsequent evidence. Mathematically, the order of the presentation of the evidence items should not change detectives' final estimate of the probability of a suspect's guilt; what happens in reality, however, is a different question.[9] Research in experimental psychology suggests people do not sequentially evaluate items of evidence; rather, they integrate them into larger cognitive structures of stories and narratives (Stubbins and Stubbins 2009). In holistic theories, evidential reasoning is bi-directional rather than linear, moving backwards and forwards between evidence and conclusions. However, this process can produce subconscious revisions of evidence, shaped by an individual's prior beliefs and biases. Consequently, some scholars suggest subjecting evidential reasoning to Bayesian scrutiny to help mitigate unintended distortions (Griffin 2013; Ho 2015).

Confirmation Bias

Systematic thinking errors in a police investigation can result in an unsolved crime or a wrongful conviction. Like all decision-makers, detectives are susceptible to the influence of cognitive biases operating at a below-conscious level (Kahneman et al. 1982; Stubbins and Stubbins 2009). A rush to judgment often leads to tunnel vision and confirmation bias, distorting the interpretation and evaluation of other case evidence.

Tunnel vision occurs when there is a narrow focus on a limited range of alternatives:

Tunnel vision is insidious…. It results in the [police] officer becoming so focussed upon an individual or incident that no other person or incident registers in the officer's thoughts. Thus, tunnel vision can result in the elimination

of other suspects who should be investigated. Equally, events that could lead to other suspects are eliminated from the officer's thinking. (Cory 2001, p. 37).

However, tunnel vision has not been well defined by psychological researchers, and the term is more commonly used in legal literature as a metaphor for the reluctance to consider alternatives (Snook and Cullen 2009). Confirmation bias, on the other hand, has an operational definition, making it more amenable to proper research and analysis.

Confirmation bias is a type of selective thinking. We have a tendency to confirm rather than refute our hypotheses by looking for supporting information, minimizing inconsistent evidence, and interpreting ambiguous information in a manner consistent with our beliefs (Nickerson 1998). Types of confirmation bias include (1) the biased search for evidence, (2) the biased interpretation of information, and (3) a biased memory. Explanatory mechanisms for confirmation bias include motivated processes (the need for consistency, the influence of desire on belief) and cognitive processes (a limited ability to handle complex tasks) (Oswald and Grosjean 2004). A rush to judgment leading to tunnel vision is often the catalyst (Findley and Scott 2006; MacFarlane 2010; Martin 2002).

Confirmation bias, with the support of faulty assumptions, probability errors, and groupthink, plays a key role in the causal structure of wrongful convictions (Rossmo and Pollock 2019). Detectives suffering from confirmation bias are more likely to interpret evidence in a prejudicial manner, uncritically accepting what supports their investigative theory while cursorily rejecting what does not (such as alibi information). They often fail to engage in differential diagnoses by considering other hypotheses of the crime or alternative suspects.

As more resources—time, money, emotions, reputation—are invested in a particular theory of the crime, sunk costs may prevent investigators from rethinking their case. Whether through psychological lethargy or organizational momentum, belief perseverance can survive powerful contradictory evidence (Brockner 1992; Coleman 2010; Staw 1981). Some detectives have refused to abandon their original suspect, justifying their obstinacy through the most bizarre of reasoning (Rossmo 2009). Cognitive biases can distort our interpretations of probability (beliefs of what is likely or not), producing errors in rationality and logic. If exculpatory evidence has been uncovered, detectives must reconsider their investigation and explore alternative explanations. Unfortunately, as critical thinking requires effort, and organizational shifts are challenging, this does not always happen.

Evidence Failures

As evidence flows through all phases of a criminal investigation, any integral evidentiary problems risk derailing

---

[9] Confirmation bias in a criminal investigation can be tested for by altering the temporal order of evidence discovery; if doing so influences any of the case conclusions, there is likely a problem with the investigative logic (Rossmo 2016).

Fig. 2 Evidence failure modes
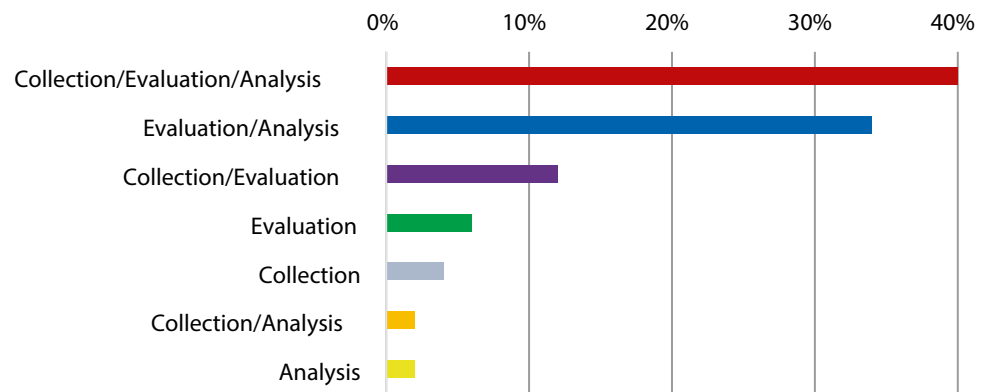
a successful outcome. There are three types of evidence error that can impede or bias detective decision-making:

1. evidence collection—a failure to collect all the relevant evidence necessary to thoroughly investigate the case (e.g., crime scene evidence, neighborhood canvass, interviews);
2. evidence evaluation—a failure to assess evidence reliability (the probability an item of evidence—e.g., a confession, a witness statement, a lab analysis—is accurate or true); and
3. evidence analysis—a failure to logically analyze the evidence (e.g., strength, reliability implications, connections, patterns).

These errors often originate from a rush to judgment, followed by tunnel vision, confirmation bias, and/or groupthink. Research on the systemic causes of criminal investigative failures has found most involve multiple issues, 92% of which include evidence evaluation problems (Rossmo and Pollock 2018). Figure 2 illustrates this breakdown.

Evidence collection problems likely play a more significant role in criminal investigative failures involving unsolved crimes that should have been solved. An investigation is a form of historical science (Cleland 2002). Classical experimental sciences, such as physics and chemistry, observe and test in the effort to validate or falsify hypotheses (Popper 1965). The process is inductive and forward-looking in that the results of an experiment are, by definition, only known in the future. By contrast, the evidential reasoning of historical sciences, including paleontology, astronomy, and geology, follows a more deductive process. Traces of past events are exposed and explanatory hypotheses formulated. The classical sciences reason from causes to effects, while the historical sciences—investigations included—reason from effects to causes. For the latter, the search for evidence takes on a primary role as traces generated by past events may exist but remain undiscovered. Any undiscovered evidence will

never be evaluated or analyzed. In a criminal investigation, the probability of discovery is a function of elapsed time as most evidence eventually decays. Cold case detectives "face problems of faded memories, missing witnesses, retired police officers, misplaced evidence, and altered crime scenes. Their criminal investigative expertise must be combined with historical research skills" (Rossmo 2017, p. 560).

Case Study

The infamous wrongful murder conviction of David Milgaard provides a useful case study for examining the functional phases of a police investigation, the role of evidence strength and reliability, and the causal structure of a criminal investigative failure.

On the morning of January 31, 1969, 19-year-old nursing assistant Gail Miller was attacked while walking to the bus stop on her way to work at City Hospital in Saskatoon, Saskatchewan (Boyd and Rossmo 2009). She was raped, stabbed, and left in a snowbank in an alley less than a block from her home in the neighborhood of Riversdale. The temperature with wind chill that morning was 42° degrees below zero Fahrenheit.[10]

David Milgaard, then a 16-year-old "hippie," became a suspect in the crime. He had set out from Regina on a road trip with two other teenagers, Ron Wilson and Nichol John, to buy drugs. They drove to Saskatoon to pick up another friend, Albert "Shorty" Cadrain (who had the money to fund their enterprise), but ended up lost and stuck in the snow. Milgaard and Wilson separated to look for help, while John stayed in the car because of the cold. The teenagers were eventually able to free their vehicle and find Cadrain's home.

---

[10] Imperial measurements were used in Canada in 1969; − 42 °F is equivalent to −41 ° C.

Police recovered items from Miller's purse leading from the murder scene to the Cadrain house, a block and a half south of the crime scene. They subsequently interviewed the three teenagers, who all denied knowing anything about the crime. However, Albert Cadrain later told detectives he saw blood on Milgaard's clothing that morning, though no one else observed this. Investigators repeatedly questioned the teenagers. Eventually, and only after spending a night in jail, John told police she saw Milgaard stab Miller and provided detectives a statement describing the attack. He was then arrested, charged with murder, and subsequently convicted.

Milgaard spent 23 years in federal prison. He was denied parole because of his refusal to admit guilt. In 1997, DNA testing conducted by a forensic laboratory in England determined that semen stains on the victim's uniform matched Larry Fisher, a convicted serial rapist, who caught the same bus at the same transit stop as Miller did every morning, and coincidentally lived in the basement suite of the Cadrain house. Milgaard was exonerated and received $10 million from the Saskatchewan government. Fisher was convicted of first-degree murder and sentenced to life in prison, where he died in 2015.

cESO Schema

The framework of the cESO (crime-evidence-suspects-offender) schema can be used to analyze the Milgaard case.[11]

Evidence

Upon discovery of the murder, police investigators took photographs, recovered two pale yellowish frozen clumps from the snow by Miller's body, conducted an area search (in which they found the victim's purse and contents, wallet, sweater, and right boot), and interviewed people within a four-block radius of the crime scene, including those who used the same bus stop as she did. The frozen clumps consisted of seminal fluid and pubic hairs. The seminal fluid contained type A blood antigens, present in about 40% of the population. Semen was recovered from Miller's vagina during her autopsy, but for some unknown reason was not subjected to any laboratory tests. Semen stains on her white nursing uniform were completely overlooked by the crime laboratory and consequently were not submitted to forensic analysis.

Suspects

Police initially thought that a serial rapist in the Riversdale neighborhood might be the killer; however, the rapes stopped after Miller's murder and this lead dried up. The victim's boyfriend was investigated and cleared. A total of 208 suspects were eventually checked, with sex offenders and criminals known to have used a knife in the past prioritized. The items from Miller's purse led police towards the nearby Cadrain house, focusing suspicion on Milgaard and his associates. The Regina teenagers were literally in the wrong place at the wrong time. It likely did not help that they were petty criminals involved in drug trafficking in a former temperance colony.

Offender

Detectives concentrated on the four youths and began to pressure them in an effort to generate evidence linking Milgaard back to the murder. Police thought they finally had their break when John gave a statement claiming to be an eyewitness to the crime. Detectives subsequently arrested Milgaard and charged him with murder. However, John's evidence was flawed; police failed to properly consider the circumstances under which they had coerced a statement from her and what this meant for the reliability of her evidence (see below). "Evidence" derived from a witness is not really evidence if the witness is lying, and a strong-arm approach is risky if investigators suffer from cognitive biases and ignore any reliability problems.

The real murderer was Larry Fisher, the Riversdale serial rapist police had first suspected. However, when Fisher's wife told police her husband might be the killer of Gail Miller, they dismissed her warning because Milgaard had already been convicted. In 1997, DNA obtained from the semen stains on Miller's nursing uniform (originally missed by the crime laboratory) established a positive link between Fisher and the murder.

Evidence Strength and Reliability—John's Statement

Key excerpts from Nichol John's statement to the police follow (*Commission of Inquiry Into the Wrongful Conviction of David Milgaard* 2005):

> After we got to Saskatoon we drove around for about 10 or 15 min. Then we talked to this girl. This was in the area where Sgt. Mackie drove me around.
> Ron was driving the car at this time. He drove to the curb where Dave spoke to this girl.

---

[11] The following discussion focuses on only key parts of the investigation; for more detail, see "Investigation into the Death of Gail Miller" (Chapter 8), *Commission of Inquiry Into the Wrongful Conviction of David Milgaard* (2005).

Dave was on the outside passenger side of the front seat. Dave opened the door to talk to this girl as she approached along the sidewalk.

Dave asked this girl for directions to either down town or Pleasant Hill. He offered to give her a ride to where ever she was going. She refused the ride.

Dave closed the door and said 'The stupid bitch'.

We started to drive away and only went about half a block when we got stuck. We ended up stuck at the entrance to the alley behind the funeral home.

Ron and Dave got out and they tried to push the car. They couldn't get it out.

I recall Dave going back in the direction we had spoke to the girl. Ron went the other way past the funeral home.

The next thing I recall is seeing Dave in the alley on the right side of the car. He had a hold of the same girl we spoke to a minute before. I saw him grab her purse. I saw her grab for her purse again. Dave reached into one of his pockets and pulled out the knife. I don't know which pocket he got the knife from. The knife was in his right hand. I don't know if Dave had a hold of this girl or not at this time. All I recall seeing is him stabbing her with the knife.

The next I recall is him taking her around the corner of the alley. I think I ran after that. I think I ran in the direction Ron had gone. I recall running down the street. I don't recall seeing anyone. The next thing I knew I was sitting in the car again. I don't know how I got back to the car.

John's statement positively identifying Milgaard as the murderer was highly probative. But it should have been evident, despite its strength, that its reliability was low. John was a 16-year-old street youth and drug user. She continued to travel with Milgaard after the murder. The stories she told police varied greatly. She was under pressure by detectives to talk about the Miller case, and the day before her statement she was arrested and jailed overnight. John later claimed she could not remember anything about the attack.

Moreover, John's description of the location of the attack was not consistent with the route Miller typically took to the bus stop; unsurprisingly, this was the most direct path, a 3-min walk (Rossmo 2016). On the morning of her murder, Miller left her home 5 min before the bus was due—what you would expect given the extreme cold that morning—and there is no reason she would have been at the place where John said the attack occurred.

Most critical, the physical evidence failed to support John's description of events. No blood was found in the snow where she claimed the attack happened or anywhere other than by Miller's body. Even more problematic, while the pattern of cuts in the victim's coat matched the pattern of her stab wounds, there were no cuts in her nursing uniform. This meant that, at some point, Miller's coat had to have come off, her uniform pulled down, and her coat put back on again. None of this was described by John to the detectives.

There were significant inconsistencies between Nichol John's statement and the other evidence in the case, and the risk of her lying simply to get the police to leave her alone should have been considered. Despite its unreliability, the statement had great strength and the police and prosecutor believed her. Unfortunately, so did the jury. This flawed piece of evidence ended up playing a key role in the wrongful conviction of David Milgaard.

Concept Map

The Milgaard wrongful conviction can be deconstructed by identifying its causes and their relationships. The primary cause of this criminal investigative failure was confirmation bias on the part of police detectives, which resulted in a biased interpretation of evidence. However, as in most other such failures, several causal factors were involved and interacted to produce the outcome. Specific factors included:

- Unsolved sex murder—Police were under some pressure because they had not solved a high-profile sex murder.
- Fisher-Cadrain house coincidence—Fisher lived in the basement suite of the Cadrain house. Items from Miller's purse and clothing were recovered between the murder scene and the nearby Cadrain house.
- Cadrain's deceitful witness claim—Albert Cadrain told police he saw blood on Milgaard the day of the murder, something no one else noticed. Cadrain may have been jealous of Nichol John's interest in David Milgaard or motivated by the reward money.
- Police pressure on juveniles—Police detectives put pressure on Milgaard's teenage friends because of Albert Cadrain's deceitful claim that he had seen blood on Milgaard, and because of the proximity of the Cadrain house to the murder scene.
- John's untruthful witness statement—Nichol John gave detectives a statement, after considerable police pressure (including being arrested and spending a night in jail), claiming she witnessed Milgaard attack Miller on the street, stabbing and then dragging her into an alley. However, her description of events was inconsistent with much of the physical evidence, including the location of the attack, the lack of blood stains in the snow, and the coat cuts and wound pattern.
- Confirmation bias—Police detectives suffered from confirmation bias, which resulted in the biased evaluation and interpretation of evidence. John's statement was taken at face value and its unreliability was ignored; the

Fig. 3 Milgaard case concept map
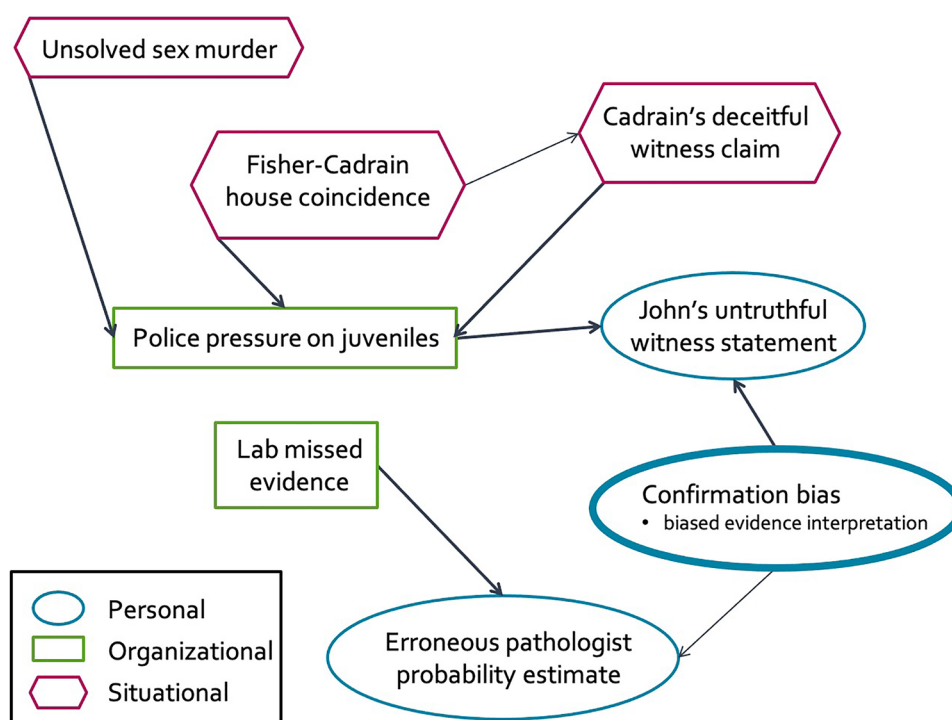
limited opportunity for Milgaard to commit the murder given the known timing of events was not considered; exculpatory witnesses were discounted; Milgaard's "hippie" lifestyle and drug use were considered relevant.

- Erroneous pathologist probability estimate—A trace amount of semen found in the snow at the murder scene was determined to have originated from someone with type A blood. Tests showed Milgaard was a non-secretor, meaning there would have been no antigens in his semen. The pathologist tried to explain this discrepancy with a highly improbable scenario involving a genital injury.
- Lab missed evidence—The RCMP forensic laboratory failed to find the killer's sperm on Miller's white nursing uniform. Consequently, physical evidence that would have excluded Milgaard in 1969 was ignored.

The concept map for the Milgaard case is presented in Fig. 3:

Conclusion

Police investigators are routinely expected to solve a case quickly and move on to the next crime—often within 48 hours (Simon 1991). But production pressures can undermine detective work. Undiscovered evidence cannot be analyzed, plays no role in the search for suspects, and will never be used to convict an offender. Time constraints jeopardize accurate reliability assessments and encourage cognitive biases. Resource limitations translate into evidence that is not fully analyzed in the forensics laboratory or in the minds of detectives.

These issues are particularly problematic in large cities where the preponderance of crime occurs. Most offenses are never cleared and even many serious crimes such as murder remain unsolved by police (Eck and Rossmo 2019). An understanding of the anatomy of a criminal investigation, however, may help suggest remedies for stalled or cold cases by highlighting specific problem areas. Some questions that might be considered include:

- Was any evidence missed; if so, can it still be recovered? Has reliability been evaluated and its impact on strength determined? Was the evidence logically analyzed and holistic patterns developed?
- Are failure risk factors present; if so, how can they be mitigated and managed?
- What type of evidence will most likely establish guilt in this particular case? Has sufficient effort been devoted to developing it?

Evidence is the *sine qua non* of any investigation; it originates from the crime itself and flows through every phase of the process. Each of these stages involves different objectives and requires distinct tactics. A better

appreciation of the structure and nature of the criminal investigative function is necessary if we want to maximize the probability of success while minimizing the risk of failure.

Declarations

Conflict of Interest The author declares that he has no conflict of interest.

References

Allore J, Pearson P (2009) What happened to Theresa Allore? Rossmo DK. Criminal investigative failures. Taylor & Francis, Boca Raton, FL, pp 219–252

Beauregard E, Martineau M (2014) No body, no crime? The role of forensic awareness in avoiding police detection in cases of sexual homicide. J Crim Justice 42:213–220

Boyd N (2019) Canadian law: An introduction, 7th edn. Nelson Publishing, Toronto

Boyd N, Rossmo DK (2009) Milgaard v The Queen: understanding a wrongful conviction for sexual homicide. Rossmo DK, Criminal investigative failures. Taylor & Francis, Boca Raton, FL, pp 179–204

Brockner J (1992) The escalation of commitment to a failing course of action: toward theoretical progress. Acad Manage Rev 17:39–63

Chaiken JM, Greenwood PW, Petersilia J (1976) The criminal investigation process: a summary report. Rand, Santa Monica, CA

Cleland CE (2002) Methodological and epistemic differences between historical and experimental science. Police Pract Res 69:474–496

Coleman MD (2010) Sunk costs and the commitment to medical treatment. Current Psychol 29:121–134

Commission of inquiry into the wrongful conviction of David Milgaard. (2005, March 7). Public documents. Retrieved December 10, 2013, from http://www.justice.gov.sk.ca/milgaard/pubdocs/march7/Nichol%20John/018589.pdf

Cory PC (2001) The inquiry regarding Thomas Sophonow. Queen's Printer, Winnipeg, MB

Cox RT (1946) Probability, frequency and reasonable expectation. Am J Phys 14:1–13

Doney RH (1990) The aftermath of the Yorkshire Ripper: the response of the United Kingdom police service. In: Egger SA (ed) Serial murder: An elusive phenomenon. Praeger, New York, pp 95–112

Eck JE, Rossmo DK (2019) The new detective: rethinking criminal investigations. Crim & Pub Policy 18:601–622

Eddy SR (2004) What is Bayesian statistics? Nat Biotechnol 22:1177–1178

Findley KA, Scott MS (2006) The multiple dimensions of tunnel vision in criminal cases. Wisconsin Law Rev 2:291–397

Gilovich T, Griffin D, Kahneman D (eds) (2002) Heuristics and biases: the psychology of intuitive judgment. Cambridge University Press, Cambridge

Griffin D, Tversky A (1992) The weighing of evidence and the determinants of confidence. Cogn Psychol 24:411–435

Griffin LK (2013) Narrative, truth, and trial. Georget Law J 101:281–335

Ho HL (2015, Winter) The legal concept of evidence. In Zalta EN (ed), The Stanford encyclopedia of philosophy. Stanford, CA:

Metaphysics Research Lab, Stanford University. Retrieved, April 13, 2020, from https://plato.stanford.edu/archives/win2015/entries/evidencelegal/

Horvath F, Meesig RT, Lee YH (2003) National survey of police policies and practices regarding the criminal investigation process: twenty-Five years after Rand (NCJRS 202902). U.S. Department of Justice, Washington, DC

Iversen GR (1984) Bayesian statistical inference. Sage university paper series on quantitative applications in the social sciences, 43. Beverly Hills: Sage

James J, Beauregard E (2020) Murder vs investigator: factors influencing the resolution of sexual homicide cases. Police Pract Res 21:172–186

Kahneman D (2003) A perspective on judgment and choice: mapping bounded rationality. Am Psychol 58:697–720

Kahneman D, Slovic P, Tversky A (eds) Judgment under uncertainty: heuristics and biases. Cambridge University Press, Cambridge

Kaufman F (1998) The commission on proceedings involving Guy Paul Morin: Report. Ontario Ministry of the Attorney General, Toronto

Klein G (2003) The power of intuition. Doubleday, New York

Klockars CB, Mastrofski SD (eds) (1991) Thinking about police: contemporary readings, 2nd edn. McGraw-Hill, New York

Leovy J (2015) Ghettoside: a true story of murder in America. Spiegel and Grau, New York, NY

MacFarlane BA (2010) Wrongful convictions: the effects of tunnel vision and predisposing circumstances in the criminal justice system. Retrieved May 16, 2017, from https://www.attorneygeneral.jus.gov.on.ca/inquiries/goudge/policy_research/pdf/Macfarlane_Wrongful-Convictions.pdf

Martin D (2002) Lessons about justice from the "laboratory" of wrongful convictions: tunnel vision, the construction of guilt and informer evidence. Univ of Missouri-Kansas City Sch of Law Rev 70:847–864

McCrary GO (2009) Who killed Stephanie Crowe? In: Rossmo DK (ed) Crim investig failures. Taylor & Francis, Boca Raton, FL, pp 143–178

Myers D (2002) Intuition: its powers and perils. Yale University Press, New Haven, CT

National Research Council (2009) Strengthening forensic science in the United States: a path forward. The National Academies Press, Washington, DC

Nickerson RS (1998) Confirmation bias: a ubiquitous phenomenon in many guises. Rev Gen Psychol 2:175–220

O'Brien B (2009) Prime suspect: An examination of factors that aggravate and counteract confirmation bias in criminal investigations. Psychol, Public Policy, and Law 15:315–334

Oppal WT (2012) Forsaken: the report of the Missing Women Commission of Inquiry. Executive summary. Victoria: British Columbia Provincial Government

Oswald ME, Grosjean S (2004) Confirmation bias. In: Pohl RF (ed) Cognitive illusions: a handbook on fallacies and biases in thinking, judgement and memory. Psychology Press, Hove, UK, pp 79–96

Popper KR (1965) Conjectures and refutations: the growth of scientific knowledge. Harper and Row, New York

Risinger DM, Loop JL (2002) Three Card Monte, Monty Hall, modus operandi and "offender profiling": some lessons of modern cognitive science for the law of evidence. Cardozo Law Rev 24:193–285

Robertson B, Vignaux GA (1995) Interpreting evidence: evaluating forensic evidence in the courtroom. John Wiley & Sons, Chichester

Rossmo DK (2009) Criminal investigative failures. Taylor and Francis, Boca Raton, FL

Rossmo DK (2016) Case rethinking: a protocol for reviewing criminal investigations. Police Pract Res 17:212–228

Rossmo DK (2017) Geographic profiling in cold cases. In: Walton RH (ed) Cold case homicides: practical investigative techniques, 2nd edn. CRC Press, Boca Raton, FL, pp 559–582

Rossmo DK (2020) Anatomie d'une enquête criminelle (The anatomy of a criminal investigation). Criminologie 53(2):17–42

Rossmo DK, Pollock JM (2018) Case deconstruction of criminal investigative failures. Final summary overview submitted to the National Institute of Justice, Washington, DC

Rossmo DK, Pollock JM (2019) Confirmation bias and other systemic causes of wrongful convictions: a sentinel events perspective. Northeastern University Law Review 11:790–835

Simon D (1991) Homicide: a year on the killing streets. Ballantine Books, New York

Snook B, Cullen RM (2009) Bounded rationality and criminal investigations: has tunnel vision been wrongfully convicted? In: Rossmo DK (ed) Crim investig failures. Taylor & Francis, Boca Raton, FL, pp 71–98

Staw BM (1981) The escalation of commitment to a course of action. Acad Manag Rev 6:577–587

Stelfox P, Pease K (2005) Cognition and detection: reluctant bedfellows? In: Smith MJ, Tilley N (eds) Crime science: New approaches to preventing and detecting crime. Willan Publishing, Cullompton, Devon, pp 191–207

Stubbins D, Stubbins N (2009) On the horns of a narrative: Judgment, heuristics, and biases in criminal investigation. In: Rossmo DK (ed) Crim investig failures. Taylor & Francis, Boca Raton, FL, pp 99–140

Tversky A, Kahneman D (1974) Judgment under uncertainty: heuristics and biases. Science 185:1124–1131

Wells GL, Olson EA (2002) Eyewitness identification: Information gain from incriminating and exonerating behaviours. J Exp Psycho: Applied 8:155–167

Willmer MAP (1970) Crime and information theory. Edinburgh University Press, Edinburgh

Wright M (2013) Homicide detectives' intuition. J Investig Psycho and Offend Prof 10:182–199