European Journal for Philosophy of Science (2026) 16:37
https://doi.org/10.1007/s13194-026-00733-0

PAPER IN GENERAL PHILOSOPHY OF SCIENCE

The relevance principle from a causal perspective

Alexander Gebharter[1] · **Michał Sikorski**[1]

Received: 21 November 2024 / Accepted: 1 March 2026

Abstract
The relevance principle plays a central role in the methodology of forensic science. Recently, it has been argued that it should also be applied in other scientific disciplines. The principle rules which information experts should use for evaluating evidence. A precise formulation has been given in terms of probabilistic relevance. In this paper, we focus on this probabilistic version and put it to the test by applying it to different causal scenarios and by discussing it to the background of two different assumptions about the trier of fact: the trier of fact as a causal reasoner vs. the trier of fact being agnostic about the causal structure. Finally, we point to problems with the probabilistic version of the relevance principle and propose improved versions that capture the principle's underlying idea better than the purely probabilistic version does.

Keywords Relevance principle · Evidence evaluation · Causal structure · Bias · Double-counting fallacy

1 Introduction

Scientific evidence is sought after by both prosecutors and defense attorneys. Given its prevalence and the critical role it plays in legal proceedings, it is paramount to ensure that such evidence fulfills its promise and is, in fact, reliable. Broadly speaking, scientific evidence can be divided into two categories: general scientific findings (e.g., from psychology or sociology) or models (e.g., mathematical or

✉ Alexander Gebharter
alexander.gebharter@gmail.com

Michał Sikorski
michalpsikorski@gmail.com

[1] Center for Philosophy, Science, and Policy (CPSP), Department of Biomedical Sciences and Public Health, Marche Polytechnic University, Via Tronto 10A, 60126 Ancona, Italy

Published online: 28 April 2026

economic models) that are relevant to the case at hand on one side, and results from established forensic disciplines, such as fingerprint or DNA analysis, on the other. The admissibility of both types of evidence is governed by the so-called Daubert standard (see, e.g., Chin, 2014, Peruzzi & Cevolani, 2024). In practice, the admissibility of established forensic methods, such as fingerprint or DNA analysis, is rarely challenged in Daubert hearings, as their compliance with the standard was established at the time of their adoption.[1] In contrast, discussions about the reliability of established forensic disciplines often center on the evidentiary basis and the reliability of forensic procedures themselves.[2] Concerns about the reliability of forensic science were prominently raised by the influential report of the National Research Council (2009), which suggested that the results of forensic analyses might be vulnerable to bias from external information (e.g., knowledge of a suspect's confession or results from other forensic tests). This conjecture has since been the subject of extensive empirical investigation, with most studies finding that such external factors can indeed influence forensic judgments (see Cooper & Meterko, 2019 for a systematic review). In response, a number of scholars and organizations have proposed the implementation of strict guidelines regarding the types of information that forensic practitioners should and should not be exposed to (e.g., National Research Council, 2009; Quigley-McBride et al., 2022; Thompson, 2011). One of the most developed and widely discussed proposals is what we call the *relevance principle* in this paper, saying that "[f]orensic scientists should rely solely on task-relevant information when performing forensic analyses" (National Commission on Forensic Science, 2016, p. 1). Task-relevant information is often informally characterized, but can also be analyzed in a precise formal way as information that would still have a probabilistic impact on the evidence given the hypothesis (e.g., whether the suspect was the source of the fingerprints) were true or false (ibid., p. 9). In this paper, we focus on this precise version of the relevance principle.

The relevance principle has been the subject of heated debate. Proponents argue that it helps mitigate forensic bias and makes forensic results more useful for resolving questions of fact at trial (see, e.g., Kukucka, 2020; Thompson, 2011). Opponents, such as Butt (2013), Curley et al. (2020), and Elaad (2013), by contrast, have expressed concerns that the restriction to task-relevant information might limit the reliability or completeness of forensic evaluations.[3] Nonetheless, the arguments in favor of the principle have generally been regarded as convincing. As a result, the

[1] The "DNA wars" of the late 1980s and early 1990s are a well-documented example of legal and scientific debates over whether DNA profiling satisfies admissibility standards such as *Daubert* and *Frye* (Saks & Koehler, 2005).

[2] See Haack (2015) for a discussion of the relationship between the two problems.

[3] Interestingly, it appears that both sides of the debate are, to some extent, correct. In certain cases, the use of task-irrelevant information can indeed improve the reliability of forensic analyses. However, at the same time, such use can make the outcome of a forensic procedure dependent on contextual factors, such as the results of prior procedures, thereby reducing the epistemic independence and transparency of the analysis. As a consequence, the findings become less useful for the fact-finder, who seeks evidence that can be independently evaluated. This tension between improving reliability through broader informational input and maintaining the objectivity and utility of forensic results constitutes what has been termed the *criminalist's paradox* (see, e.g., Thompson, 2011 or Sikorski & Gebharter, 2025).

relevance principle is now widely viewed as the gold standard in forensic science for determining what information forensic experts should consider when evaluating evidence that may assist the trier of fact in reaching a well-informed judgment (Bird et al., 2024; Koehler et al., 2023; Kunkler & Roy, 2023). It is intended to distinguish between relevant information that can safely be used by experts from relevant information that cannot, where relevant information – following orthodoxy in law (cf. National Court Rules Committee, 2024, rule 401) – is understood as information that has a positive or negative probabilistic impact on the outcome of the evaluation. In a nutshell, the relevance principle says that forensic experts should consider all and only the relevant information that would make a probabilistic difference for the evidence if the hypothesis they have been tasked to evaluate were true (or false).

It has been argued that the relevance principle is not only relevant for forensic science, but for any science in which experts are tasked to evaluate hypotheses on whose basis a decision maker may act (Sikorski, 2022; Sikorski & Gebharter, 2025). For example, a radiologist might analyze an X-ray image on the basis of which a medical doctor may decide whether a patient should undergo a certain medical treatment. Or a team of scientists monitoring seismic activity to gather evidence to be considered by policy makers in order to decide whether the region needs to be evacuated because of an imminent threat of a severe earthquake. To keep things simple, we content ourselves with applications to cases from forensic science in this paper. However, our findings can easily be generalized for any scientific discipline that relies on expert evaluations.

The relevance principle is attractive for several reasons. It can be justified on the basis of a Bayesian network analysis (Thompson, 2016) and helps to avoid shifts in the decision thresholds of forensic scientists that would undermine the probative value of their conclusions (Thompson, 2023). It also allegedly helps to avoid the double-counting fallacy that might influence the trier of fact's decision in unfavorable ways (Sikorski & Gebharter, 2025). In this paper, we focus on the latter. In particular, we put the relevance principle to the test by looking at its performance from a causal perspective. We consider elementary ways of how a suspect's guilt to be evaluated by the trier of fact, a source level hypothesis – e.g., the source of the fingerprint collected at the crime scene and one collected from a suspect are the same – to be evaluated by a forensic expert, the evidence to be analyzed by that expert, and a piece of additional information – e.g., possible contamination, the result of another forensic procedure, etc. – can be causally related to each other and then explore whether the relevance principle allows us to derive the intended results. We will have a look at the principle's performance in these structures under two different assumption: the trier of fact being ignorant about the causal structure and treating the expert's assessment of the hypothesis as well as the additional information as independent evidence for assessing the suspect's guilt vs. the trier of fact being a causal reasoner, meaning that they will evaluate the guilt hypothesis based on the causal structure underlying the case at hand. It will turn out that the probabilistic version of the relevance principle encounters problems in some causal settings in each of these two cases. Finally, we propose alternative versions of the relevance principle that will return the intuitively correct results in each of the causal structures and under each of the two assumptions about how the trier of fact evaluates the suspect's guilt. We would like to emphasize

that our critique is not aimed at the very intuitions underlying the relevance principle, but rather at the purely probabilistic version intended to capture these ideas. We are in favor of the relevance principle, but believe that it cannot be adequately formalized in purely probabilistic terms.

The paper is structured as follows: In Section 2, we provide more details on the relevance principle as well as its motivation and role in forensic science. In Section 3, we introduce the formal tools we will use in later sections to model causal structure: causally interpreted Bayesian networks (Pearl, 2000; Spirtes et al., 2000). In Section 4, we present six elementary causal settings and provide an example for each of them. In Section 5, we look at the relevance principle's performance in these causal structures under the assumption that the trier of fact is ignorant about the causal structure and in Section 6 under the assumption that the trier of fact is a causal reasoner. We highlight problems for the probabilistic version of the relevance principle under each assumption and propose alternative versions that return better results in each causal setting. We wrap up and conclude in Section 7.

2 The relevance principle

Here is a toy example: Assume there is a legal case. The main question of interest to the trier of fact is whether the suspect committed the murder and is guilty. Several traces were secured at the crime scene and are now being analyzed by forensic experts in search of evidence. Assume you are a fingerprint analyst tasked with finding out whether the fingerprints on the murder weapon are the suspect's fingerprints, and your friend is a DNA analyst tasked with finding out whether the DNA found on the murder weapon is the suspect's DNA. You both have to come up with a result: match, no match, or the analysis was inconclusive. However, realistically, you will not be 100% certain about your result, but rather have a certain degree of confidence that the prints on the murder weapon are the suspect's prints, on which you will base your conclusion.

Assume that your friend told you about the result they obtained from the DNA analysis. Now the question is whether you should take this information into account for your own evaluation of the fingerprint samples you received. On the one hand, it would increase the accuracy of your result because DNA analysis is generally considered to be much more reliable than fingerprint analysis.[4] Because of this, some – though a minority – of forensic scientists have argued that you should consider this information (see, e.g., Curley et al., 2020; Elaad, 2013). On the other side, if you decide to base your conclusion partly on the DNA results, it can lead to at least two related problems: bias and double-counting.

In the methodology of forensic science, bias is typically understood as any influence of task-irrelevant information on the way forensic procedures are conducted (see, e.g., National Commission on Forensic Science, 2016). As

[4] For illustration, we chose an example in which one of the two methods used is clearly more reliable than the other. However, neither the analysis nor the causal approach we develop in this paper depends on how the involved methods compare to each other in terms of reliability.

discussed in the introduction, bias became a central concern in the methodology of forensic science following the publication of the National Research Council's (2009) report. Empirical studies suggest that contextual information cannot only influence the outcomes of forensic procedures, but also divert them from the correct result. While the underlying mechanisms behind such biases have not been empirically established in detail, a plausible hypothesis was proposed by Thompson (2016):

> "This information is likely to influence the expert toward identifying the print as the suspect's for at least two reasons. First, it creates a strong motive to identify the suspect. The motive in question does not arise from malice or partisanship; it arises from one of the most basic and common desires of mankind – a desire that may be particularly strong among those who make their livings as experts – it is the desire to reach the correct conclusion, to be right. The eyewitness evidence helps assure the fingerprint examiner that the suspect really is the perpetrator and hence that identifying him as the source of the crime-scene print is the right decision." (Thompson, 2016, p. 144.)

Additionally, Dror et al. (2017) argued that interference between different biases or biases at different stages of a forensic procedure can create snowball-and-cascade-like effects that amplify the overall negative effect. The biasing potential of contextual information and the emergence of bias due to snowball-and-cascade-like effects were demonstrated in a recent simulation study (Cuéllar et al., 2022).

A related concern is the so-called double-counting fallacy. One commits this fallacy if one considers a piece of evidence more than once for confirming or disconfirming a hypothesis. Let us further elaborate by coming back to our toy example: You have learned from your friend that they found that the DNA samples do not match. Your own result (not considering this information) is that the fingerprint samples do match, but your degree of certainty is just slightly above your acceptance threshold. So you would report "match" to the trier of fact on the basis of your own analysis only. However, since you know that DNA analysis is generally much more reliable, you would update your degree of certainty downwards if taking the negative result of your friend into account for your own analysis. Thus, you would tell the trier of fact "no match" in that case. Let us assume that you decided for the latter option, resulting in you reporting a "no match" result. This decision might be fatal. To illustrate why, let us assume further that the suspect was in fact guilty, that they committed the crime, but the trier of fact ruled "not guilty" based on the evidence it received. Had you not taken your friend's result into account, the trier of fact would not have received two results (yours and your friend's) pointing towards the same conclusion (not guilty) and it might have requested further evidence or another round of assessments before forming an opinion about whether the suspect is guilty. As a result, the trier of fact might have come up with the verdict "guilty" which would, in our example, have been the correct decision. Summarizing, the case is such that an unjust court decision results from you having your own verdict about whether the fingerprints match being influenced by the DNA result of your friend.

The problem described above is that part of the DNA result has been counted twice by the trier of fact when they evaluated whether the suspect was the murderer, because it was also a crucial part of your fingerprint results after your own analysis. It was exactly this double-counting that led to the misguided court decision. Taking the DNA result into account made your own evaluation more reliable, but led the trier of fact astray at the same time. As we have seen, the dominant view among forensic scientists is that the danger from bias and double-counting outweighs the possible gain in accuracy and that therefore you should not take information about the DNA analysis into account in your own analysis (see e.g., Kukucka, 2020; Thompson, 2011). The relevance principle is intended to capture exactly this preference for minimizing the risk of bias and avoiding the double-counting fallacy over and above a possible gain in accuracy. In the following, we use a version of the principle based on (National Commission on Forensic Science, 2016, technical apendix):

> **Relevance principle:** When analyzing evidence E for a hypothesis H an expert should only consider additional information I such that $Pr(E|H, I) \neq Pr(E|H)$ or $Pr(E|\neg H, I) \neq Pr(E|\neg H)$.

In words this means that forensic experts should consider all and only the information that makes a probabilistic difference for the evidence in the light of the hypothesis they are tasked to evaluate. Applied to our example, the principle implies that you should not rely on the DNA result from your friend because whether the DNA sample taken from the murder weapon matches the DNA sample taken from the suspect does not influence the probability of the evidence you are tasked to analyze given the hypothesis (that the suspect is the source of the fingerprints) is true or false. Thus, the relevance principle is a clear step in the right direction. As we will see later on, however, the validity of its probabilistic formulation is limited to specific causal settings and assumptions of how the trier of fact uses evidence when evaluating the suspect's guilt. In some causal settings it fails to prevent double-counting at the level of the trier of fact and, thus, needs to be further improved.

Before we go on and introduce the formal framework we will use to model causal structures, a few remarks are in order to set the stage: Throughout the paper we use G to represent the guilt hypothesis to be evaluated by the trier of fact, H for the source hypothesis – e.g., the suspect being the source of the prints collected from the crime scene – to be evaluated by the expert, E for the evidence the expert is tasked to analyze, typically consisting in the features of analyzed traces, and I for a piece of additional information. We only consider cases in which (i) the hypothesis H to be evaluated by the expert depends on whether the suspect is guilty G, (ii) the evidence E probabilistically depends on the hypothesis H, and (iii) I as well as the expert's evaluation of H are available to the trier of fact for evaluating whether the suspect is guilty. All three assumptions are plausible. (i) is required for the hypothesis H to count as evidence for the suspect's guilt G. (ii) captures that E is evidence for the hypothesis H. Finally, (iii) is required to render the cases we will have a look at possible instances of double-counting by the trier of fact.

3 Causal Bayesian networks

To get a more formal grip on causation, we introduce the relevant basics of the causal Bayesian network machinery in this section. A causal Bayesian network consists of a graph $\mathbf{G} = \langle \mathbf{V}, \mathbf{E} \rangle$ and a probability distribution Pr over $\mathbf{V}$. $\mathbf{V}$ is a set of random variables $X_1, ..., X_n$ that can represent properties, events, hypotheses, pieces of evidence, and pretty much anything one likes them to represent. $\mathbf{E}$ is a set of directed edges ($\longrightarrow$) connecting pairs of variables in $\mathbf{V}$. $X_i \longrightarrow X_j$ stands for X_i being a direct cause of X_j with respect to variable set $\mathbf{V}$. Chains of such edges are called paths. A path of the form $X_i \longrightarrow ... \longrightarrow X_j$ is called a directed path from X_i to X_j and X_i is called an ancestor of X_j while X_j is called a descendant of X_i.[5] A variable X_i's direct ancestors in the graph are called its parents and its direct descendants are called its children. $\mathbf{Par}(X_i)$ is the set of X_i's parents. A variable X_i with two incoming arrows on a path is called a collider on that path. The graphs of causally interpreted Bayesian networks are assumed to be acyclic, meaning that they do not feature causal cycles or paths of the form $X_i \longrightarrow ... \longrightarrow X_j$. Finally, Pr is a probability distribution over $\mathbf{V}$ that can be interpreted to represent the strengths of the causal influences propagated along the causal structure underlying $\mathbf{V}$.

The graphs and probability distributions of causally interpreted Bayesian networks are assumed to satisfy the following factorization:

$$Pr(X_1, ..., X_n) = \prod_{i=1}^{n} \mathbf{Par}(X_i) \tag{1}$$

A directed acyclic graph and a probability distribution factor like this if and only if they satisfy the causal Markov condition (Spirtes et al., 2000, p. 29):

> **Causal Markov condition:** A graph $\mathbf{G} = \langle \mathbf{V}, \mathbf{E} \rangle$ and a probability distribution Pr over $\mathbf{V}$ satisfy the causal Markov condition if and only if every variable $X_i \in \mathbf{V}$ is probabilistically independent of its non-descendants conditional on its parents.

The causal Markov condition is equivalent to the *d*-separation condition (Pearl, 2000, sec. 1.2.3):

> ***d*-separation condition:** A graph $\mathbf{G} = \langle \mathbf{V}, \mathbf{E} \rangle$ and a probability distribution Pr over $\mathbf{V}$ satisfy the *d*-separation condition if and only if for all $X_i, X_j \in \mathbf{V}$ and $\mathbf{Z} \subseteq \mathbf{V} \backslash \{X_i, X_j\}$: If X_i and X_j are *d*-separated by $\mathbf{Z}$, then X_i and X_j are probabilistically independent conditional on $\mathbf{Z}$.

Note that $\mathbf{Z}$ can be the empty set. *d*-separation is defined as follows:

> ***d*-separation:** X_i and X_j are *d*-separated by $\mathbf{Z}$ if and only if X_i and X_j are not *d*-connected given $\mathbf{Z}$.

[5] For technical reasons, every variable is considered to be a descendant of itself.

> ***d*-connection:** X_i and X_j are *d*-connected given $\mathbf{Z}$ if and only if X_i and X_j are connected by a path such that (i) all colliders on this path are in $\mathbf{Z}$ or have a descendant in $\mathbf{Z}$ and (ii) no non-collider on this path is in $\mathbf{Z}$.

The idea is that certain causal connections (i.e., the *d*-connections) can produce and, thus, explain probabilistic dependence, while there can be no probabilistic dependence in the absence of proper causal connection (i.e., *d*-connection). For more details and a detailed explanation for why *d*-connection has exactly those features see, for example, Gebharter (2017, sec. 4.2).

Throughout the paper we assume that all the structures under consideration also satisfy the causal faithfulness condition (Spirtes et al., 2000, p. 31):

> **Causal faithfulness condition:** A graph $\mathbf{G} = \langle \mathbf{V}, \mathbf{E} \rangle$ and a probability distribution *Pr* over $\mathbf{V}$ satisfy the causal faithfulness condition if and only if for all $X_i, X_j \in \mathbf{V}$ and $\mathbf{Z} \subseteq \mathbf{V} \backslash \{X_i, X_j\}$: If X_i and X_j are probabilistically independent conditional on $\mathbf{Z}$, then X_i and X_j are *d*-separated by $\mathbf{Z}$.

The causal faithfulness condition is the converse of the *d*-separation condition. Thus, it states that every probabilistic independence implies that there is no proper causal connection (i.e., a *d*-connection), or equivalently: For every proper causal connection (i.e., a *d*-connection) there is a corresponding probabilistic dependence.

4 Elementary causal structures

Let us now introduce the elementary possibilities of how the suspect's guilt G to be evaluated by the trier of fact, the hypothesis H to be evaluated by the expert, the evidence E for that hypothesis to be analyzed by the expert, and an additional piece of information I can be causally connected. First of all, each of these causal structures will feature the directed causal path $G \longrightarrow H \longrightarrow E$. The hypothesis to be evaluated by the expert is a direct effect of the suspect being guilty G (or not) and the evidence E to be analyzed by the expert is a direct effect of H being true (or false), reflecting the assumptions that (i) H is direct evidence for G and that (ii) E is direct evidence for H. The only question left open then is how the piece of additional information I is connected to $G \longrightarrow H \longrightarrow E$. There are only six elementary possibilities. They are depicted in Fig. 1.

In (a), I is a direct cause of G. A realistic interpretation of (a) might be a case where G stands for the suspect being the murderer, H for whether the suspect is the source of the fingerprints found at the crime scene, E for the fingerprint samples the expert received (one from the suspect and one from the crime scene), and I for whether the suspect had a motive to murder the victim. Under this interpretation, I (motive) would be a direct cause of G (suspect's guilt) and would, due to faithfulness, make a probabilistic difference for G, H, and E. Under this interpretation, I is an important piece of information usually considered by the trier of fact, but it may also be available to the expert tasked with analyzing E and evaluating H.

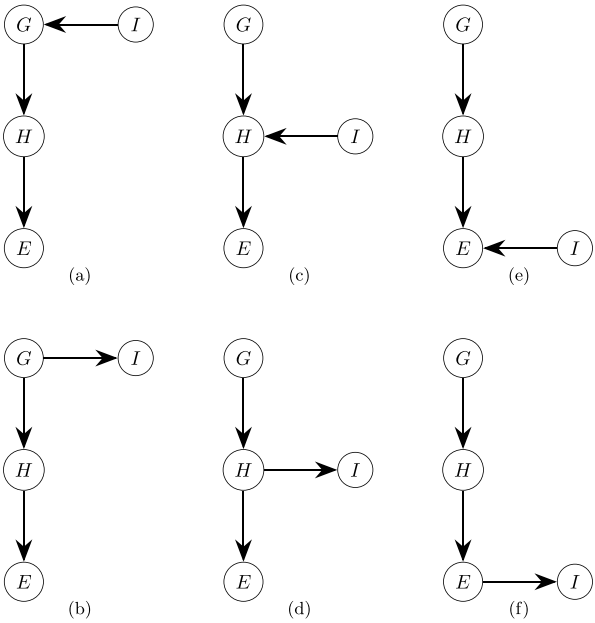

Fig. 1 Elementary possibilities how *G*, *H*, *E*, and *I* can be causally connected to each other

In (b), *I* is a direct effect of *G*. A realistic interpretation would be the one we used for illustrating and motivating the relevance principle in Section 2. *G* would again stand for the suspect's guilt (being the murderer or not being the murderer), *H* for whether the suspect is the source of the prints found at the crime scene, *E* the actual fingerprint samples to be analyzed by the expert, and *I* the result of another expert who was tasked to evaluate whether the DNA traces from the crime scene fit the suspect's DNA. *I* would be a direct effect of *G* being on a par with *H* to be evaluated by the fingerprint expert. Like in (a), *I* would make a probabilistic difference for *G*, *H*, and *E* due to faithfulness. The information *I* is typically made available to the trier of fact since it may be crucial for arriving at a verdict about the suspect's guilt, but also the fingerprint expert may have access to it.

The next possible structure is (c) in which *I* is a direct cause of *H*. Here, *I* would need to be a piece of information not having a probabilistic impact on the suspect's guilt *G*, but only on the hypothesis *H* to be evaluated by the expert and, through *H*, also on the evidence *E* to be analyzed by the expert. Here is a realistic instance of structure (c): *G* does, again, stand for the suspect's guilt, *H* stands for the fingerprints found at the crime scene coming from the suspect, and *E* for the fingerprints to be analyzed by the expert. *I* might then stand for the suspect having been present at the crime scene, but independently of the murder, maybe on another day before the crime scene was investigated. Thus, *I* would cause the prints found at the crime scene being the suspect's prints, but be independent of whether the suspect was the murderer *G*. Again, information *I* might be available to the trier of fact as well as to the expert.

A realistic interpretation of causal structure (d) would be *G* standing for the suspect's guilt, *H* for the suspect being the source of the fingerprints found at the crime scene, and *E* for the fingerprint samples to be analyzed. Finally, *I* could be the result of an analysis of the same fingerprint samples carried out by another expert.

Here I would be a direct effect of H being on a par with H's other direct effect E. Both analyzing E as well as considering I would have an impact on the likelihood of the suspect being the source of the prints H and, furthermore, also on the suspect being guilty G. Such information I is typically made available to the trier of fact, but also other fingerprint experts may have access to it.

Another interesting case is (e). Here I is a direct cause of E. A realistic interpretation would be G being the suspect's guilt, H the suspect having deposited the DNA sample to be analyzed at the crime scene,[6] E the DNA sample to be analyzed, and I information that the sample was contaminated (see Gill, 2014 for examples). In this interpretation, the information about contamination I would be independent of the truth of the hypothesis H. Furthermore, the DNA sample would be an effect of both the Hypothesis H as well as the information about contamination I. Information I concerning contamination is usually taken into consideration by the trier of fact, but may also be available to the DNA expert.

Finally, structure (f) is the less interesting one. Here the additional information I would be a direct effect of the evidence to be analyzed by the expert. It seems, however, hard to find a realistic interpretation of this structure where the expert or the trier of fact would consider an effect of the evidence to be analyzed by the expert. For the sake of completeness, we will nevertheless consider (f).

Before we go on and apply the probabilistic version of the relevance principle to each of these six structures, we would like to say a few more words on what the different agents involved in each of these structures are doing in formal terms. The expert's task is to evaluate the hypothesis H based on evidence E and possibly on additional information I. In more technical terms, this means that the expert may condition on E and I in order to update their credence in H. Depending on their specific thresholds, their credence in H after updating determines whether the expert reports to the trier of fact that H is true, false, or that the outcome of their analysis was inconclusive. The trier of fact, on the other hand, is tasked to evaluate G. They, according to our assumption (iii) from Section 2, update their credence in G after conditioning on the expert's verdict about H as well as on I. Also the trier of fact may have a threshold such that if their credence in G is above that threshold, they rule guilty and not guilty otherwise. The "beyond a reasonable doubt" standard of legal proof can be interpreted as such a threshold. Thus, each agent has a different reasoning goal and partly considers different pieces of evidence when updating and arriving at an opinion.

[6] This interpretation of H is consistent with how forensic testimony is typically understood. For example, "'Support for inclusion' is an examiner's conclusion that there is evidentiary support for the inclusion of a known individual as a possible contributor to the DNA typing results obtained from an evidentiary sample" (U.S. Department of Justice, 2022, p. 2), where the evidentiary sample is understood as a "biological sample recovered from a crime scene or collected from persons or objects associated with a crime" (National Institute of Standards and Technology, 2026).

5 The trier of fact being ignorant about the causal structure

In this section, we apply the probabilistic version of the relevance principle to the six causal structures introduced in Section 4 under the assumption that the trier of fact is agnostic about the causal structure underlying the situation at hand. By this, we mean that they will treat the expert's evaluation of H as well as the additional information I as independent information about the suspect being guilty G. More formally, we assume that the trier of fact will update their credence in G independently of the causal structure by conforming to the probabilistic constraint $Pr^*(G|H, I) \neq Pr^*(G|H)$.[7] For each of the six causal structures for which we provided a realistic interpretation in Section 4, we use that interpretation for further illustration. Structure (f) lacking a realistic interpretation we discuss only theoretically.

In structure (a), where the motive I is a direct cause of the suspect being guilty G, conditioning on H screens E off from I since H d-separates E and I. Thus, I does not impact the probability of E conditional on H, and the relevance principle rules that I should not be used by the expert when evaluating H on the basis of the fingerprint samples E. Intuitively, this is the right result, since if the expert would use I for evaluating H and, at the second step, the trier of fact would evaluate G on the basis of the expert's evaluation of H and I, the trier of fact would count the part of the information I that went into the expert's evaluation of H twice when evaluating G.

(b) is the paradigmatic case in which H is the hypothesis that the fingerprints on the murder weapon are the suspect's prints, and I is the result of the DNA analysis (i.e., whether the DNA traces found on the murder weapon match the sample taken from the suspect). It features the same d-connection relations as (a) and, thus, also in (b) H screens E off from I, meaning that I will not provide any information about E in addition to H. Hence, the relevance principle dictates that the DNA results I should not be used by the expert when evaluating whether the fingerprints match H based on the fingerprint samples E. As we already saw in the earlier discussion of this example in Section 2, this is the intuitively correct result, since if I would be considered in the evaluation of H by the expert and the trier of fact would update their credence in G conditional on the expert's report on H and I independently of the causal structure, the trier of fact would count I twice when forming an opinion about G.

In structure (c), I is a direct cause of H. Also here H d-separates E from I and, thus, I is uninformative about E in the light of H. Accordingly, the probabilistic version of the relevance principle rules that the expert tasked to evaluate H should not consider I in addition to E. And once again, this seems to be the correct result. Had the expert used information I about the suspect's presence at the crime scene on a day different from the day of the murder to further increase the accuracy of their evaluation of H, then, because the trier of fact would have considered the expert's assessment of H and I as independent evidence for G, they would have counted a part of the additional information I twice in their evaluation of the suspect's guilt G.

Next, let us have a look at causal structure (d). Like in structure (c), H d-separates E and I, which implies that I does not provide any information about E conditional

[7] The asterisk in $Pr^*(\cdot)$ indicates that the probability distribution over G, H, and I utilized by the trier of fact may differ from the distribution $Pr(\cdot)$ of the respective causal structure.

on H. Thus, the relevance principle tells us that an expert tasked to analyze H on the basis of E is not allowed to also consider I. This is the intuitively correct verdict based on our interpretation that I represents the results of an analysis of the same samples by another expert. The relevance principle succeeds in excluding double-counting by the trier of fact as intended: Were the expert tasked to assess H on the basis of E allowed to consider I in their evaluation, then part of the additional information I would have gone into the expert's assessment of H and the trier of fact would have counted this part twice in their evaluation of G based on H and I.

Structure (e) is more interesting. Here, we interpreted I as information about contamination of the DNA taken from the crime scene and, thus, as a direct cause of E that is independent of H. To flesh out the example a bit further, assume that the expert after having analyzed E in isolation, came to the belief that the sample was indeed deposited by the suspect at the crime scene H. If they now learn about I and the contamination consists, for example, in accidental partial inclusion of the suspect's DNA in the sample, the expert will most likely decrease their degree of belief in H conditional on E after learning about I. This is reflected by the causal structure: Conditioning on I does not d-separate E from H, meaning that I will have an impact on E's probability distribution conditional on H, i.e., $Pr(E|H, I) \neq Pr(E|H)$. Accordingly, the relevance principle suggests to consider the information about contamination I. Though it seems intuitively correct to consider I to be task-relevant, taking I into account may lead to double-counting by the trier of fact in their evaluation of G since the portion of I relevant for the expert's evaluation of H based on E would be counted twice by the trier of fact. A famous example of a defense lawyer forcefully – and ultimately successfully – challenging the reliability of DNA evidence on the basis of possible contamination occurred in the notorious O. J. Simpson case (see, e.g., Thompson, 1996), which resulted in the possibility of contamination having been considered by the trier of fact. If contamination had also been considered by the forensic scientist at the evidentiary stage, this would have resulted in double-counting.

Finally, structure (f). This is the only case without a clear realistic interpretation. It is, however, interesting from a theoretical perspective. Here, I is a direct effect of E. Since E and H are d-connected given I, conditioning on I will impact E's probability given H. Thus, the relevance principle implies that the expert tasked with evaluating H should consider I in their evaluation. However, since the trier of fact considers both the expert's assessment of H as well as the additional piece of information I, they would count a part of that additional information twice when evaluating the suspect's guilt G. Thus, the relevance principle fails to avoid double-counting in structure (f) for the same reason as in (e).

Summarizing, the relevance principle gave the intuitively correct advice in structures (a)–(e), while it failed in (f). Furthermore, the relevance principle succeeded in avoiding double-counting by the trier of fact in structures (a)–(d), but failed in (e) and (f). Another interesting observation is that under the assumption that the trier of fact ignores the causal structure and considers both the expert's assessment of H and the additional information I as independent evidence for the suspect's guilt G, double-counting happens in all six cases if the relevance principle is not implemented. This shows that the probabilistic formulation of the relevance principle is partially effective in preventing double-counting. What the causal analysis added was the develop-

ment of a systematic way to distinguish between possible settings and, thus, to ensure that no scenario potentially problematic for the relevance principle was overlooked when putting the principle to the test. And, as we will see in the next section, causal structure can play a much bigger role when it comes to deciding which relevant evidence is safe to be taken into consideration by the expert if the trier of fact updates their credence in G in a causal way.

Since, as we observed, double-counting happens anyway if the trier of fact ignores the causal structure and the expert considers I, a formulation of the relevance principle in purely probabilistic terms was not much promising in the first place. What gives rise to double-counting under the assumption that the trier of fact ignores the causal structure is rather that a piece of evidence, viz. I, may be considered twice, first by the expert when evaluating H and then by the trier of fact when evaluating G. Thus, for cases in which the trier of fact is agnostic about the causal structure, a better and simpler version of the relevance principle would go like this:

> **Relevance principle*:** When analyzing evidence E for a hypothesis H an expert should only consider additional information I that is not available to the trier of fact.

Following this version of the principle will allow to avoid double-counting independently of the causal structure. The downside of this version is that in some cases experts may have to ignore additional information that would have increased the reliability and accuracy of their analysis. We will have a closer look at an example in the next section.

6 The trier of fact as a causal reasoner

As we saw in Section 5, the probabilistic version of the relevance principle is not adequate for all basic cases if the trier of fact is agnostic about the causal structure. Let us now shed some light on the alternative. In this section, we assume that the trier of fact is indeed a causal reasoner, meaning that they will update their credence in the suspect being guilty G by updating on the expert's assessment of H and additional information I based on the causal structure underlying the case at hand. In more technical terms, this means that the trier of fact updates their credence in G as $Pr^*(G|H, I)$, where $Pr^*(G|H, I)$ is constrained by the d-connection and -separation relations of the respective causal structure.

We begin again with structure (a). Here, the motive I is a direct cause of the suspect being the murderer G. We already know that the relevance principle dictates that the expert should not consider I when evaluating whether the suspect is the source of the prints found at the crime scene H based on the prints from the crime scene and from the suspect E. Since we assume the trier of fact to be a causal reasoner, they will update their credence in G based on the expert's assessment of H and I and the causal structure. Since both H and I are d-connected to G, the trier of fact would indeed count the motive I twice if the expert had also used I for evaluating H. Thus, the relevance principle provides the correct result if the trier of fact updates in a causal way.

The situation is similar in causal structure (b). Since this structure features the exact same *d*-connection relations as structure (a), the relevance principle forbids the expert to use the DNA result *I* for increasing the reliability of their own fingerprint analysis. This is the correct result, since the trier of fact as a causal reasoner would count part of the DNA information *I* twice when evaluating the suspect's guilt *G*. Since this interpretation of (b) is the paradigmatic case we already discussed in Section 2, this result is not unexpected.

The next structure is (c). As we already know from Section 5, the relevance principle rules that information *I* about the suspect's presence at the crime scene on another day and independently of the murder should not be used by the expert for evaluating *H*. Under the assumption that the trier of fact is a causal reasoner, this is again the right result. Updating their credence in *G* based on the expert's assessment of *H* and *I* would indeed result in double-counting if the expert would consider *I* in their evaluation. The reason is that conditioning on *H* activates the causal path between *G* and *I* or, in other words, that *G* and *I* are *d*-connected given *H*.

Next we have structure (d). Recall that the relevance principle applied to this structure implies that the expert should not use information *I* when evaluating *H* on the basis of *E*. Under the assumption that the trier of fact is ignorant about the causal structure, we found this implication to be the intuitively correct result. But what if the trier of fact is a causal reasoner? In that case, since *H* screens *G* and *I* off each other, the trier of fact would not consider *I* in addition to *H* if presented with both. Thus, they would not count *I* twice even if the expert would consider *I* in their evaluation of *H* based on *E*. This stands in contrast to what the relevance principle suggests for structure (d), viz. that the expert should not consider *I*. But under the assumption that the trier of fact is a causal reasoner, the expert considering *I* would do no harm. It might even be beneficial if the forensic expert assesses *H* based on their own analysis of *E* as well as on the other expert's assessment *I*. After all, the forensic expert may be expected to aggregate results from other fingerprint analyzes more reliably than the trier of fact.

Also structure (e) is interesting. We already know that the relevance principle implies that the expert should take information about contamination *I* into consideration when evaluating whether the DNA traces to be analyzed were deposited by the suspect at the crime scene *H*. Under the assumption that the trier of fact ignores the causal structure and will change their credence in *G* conditional on the expert's assessment of *H* based on *E* when also considering *I*, this was the wrong result because it led to double-counting by the trier of fact. But the situation is different when assuming that the trier of fact updates their credence in *G* based on the causal structure underlying the case at hand. Since *H* *d*-separates *G* from *I*, the trier of fact as a causal reasoner will not consider *I*, even if available, because *H* already contains all the relevant information according to the causal structure. As a consequence, if the trier of fact is a causal reasoner, then the expert should indeed consider the information *I* about contamination in their evaluation because this allows them to make their inference about whether the sample was deposited by the suspect more reliable and accurate. Moreover, the DNA expert has more expertise and experience in evaluating how contamination might influence the likelihood of *H* given the evidence *E* than the trier of fact.

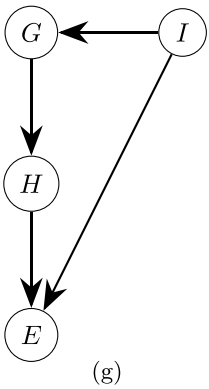

Fig. 2 The suspect being framed I by putting personal belongings at the crime scene

Finally, only structure (f) is left. As we saw in Section 5, the relevance principle implies that the expert should consider I when evaluating H based on E in this scenario. Assuming that the tier of fact is agnostic about the causal structure, this would give rise to double-counting. If the trier of fact is a causal reasoner, on the other hand, double-counting is avoided because H d-separates G from I, meaning that I does not provide any information about G over and above the expert's assessment of H. Thus, if the trier of fact is a causal reasoner, the relevance principle correctly rules that I is safe to use for the expert.

Summarizing, we found that under the assumption that the trier of fact is a causal reasoner, the probabilistic relevance principle gets things right for structures (a), (b), (c), (e), and (f). It does, however, imply that the expert should not consider I in structure (d), though I would increase the expert's accuracy without leading to double-counting by the trier of fact. What led to the underperformance in structure (d) is that the relevance principle is formulated in purely probabilistic terms and that it does not incorporate G. Whether I does have an impact on the conditional probability of E given H is neither relevant for whether considering I does improve the reliability and accuracy of the expert's evaluation of H on the basis of E, nor does it reliably guarantee that double-counting is avoided by the trier of fact. The former we could see in our discussion of structure (d), where the principle rules that the expert should not take I into consideration, though it would increase the reliability of their analysis without running the risk of double-counting by the trier of fact. The latter can be illustrated by a slightly more complex structure. As always, we assume that $G \longrightarrow H \longrightarrow E$. But this time, let I be connected to both G and E as in Fig. 2. A realistic interpretation would be G modeling whether the suspect is guilty, H whether the suspect was at the crime scene at the time of the murder, and E the belongings of the suspect found at the crime scene. Then I might be another person framing the suspect by having put personal belongings of the suspect with the suspect's fingerprints on them at the crime scene. Under this interpretation, I would be directly causally relevant for the suspect being guilty (or not) G as well as for the items E to be analyzed by the expert. In this scenario, the relevance principle would say that I should be used by the expert because it impacts the conditional probability of E given H. However, since neither H d-separates G from I nor I d-separates G from H, the trier of fact as a

causal reasoner would count I twice when updating their credence in G conditional on the expert's evaluation of H and I.

For cases where the trier of fact updates in a causal way, what is more relevant is what we just found: The expert considering I will not lead to double-counting on the trier of fact's side exactly if Hd-separates G and I or Id-separates G and H:

Relevance principle****:** When analyzing evidence E for a hypothesis H an expert should consider only additional information I such that

- H d-separates G and I, or
- I d-separates G and H.

Following this principle will rule that any information I that can increase the expert's analysis' reliability and does not lead to double-counting by the trier of fact can be used by the expert for evaluating H based on evidence E, regardless of how complex the causal structure is. A precondition is, of course, that the trier of fact is a causal reasoner. One may wonder how realistic this assumption is. We acknowledge that it is somewhat revisionary and that one cannot simply assume that the trier of fact updates in a causally adequate way. However, the required competence could easily be achieved, for example, by a crash course on causal reasoning. Another typical problem with causal accounts is that the specific causal structure is often unknown or underdetermined by the data. But in court cases, it seems quite easy to get an idea of the appropriate causal structure, as the simple structures in Fig. 1 show. If I is a motive, for example, it is usually a cause of the suspect being guilty G. If it is evidence for the hypothesis H to be evaluated by a forensic expert, then it is usually an effect of H. If it is evidence for another hypothesis to be considered by the trier of fact, then it is typically a separate effect of G on a par with H, and so on.

7 Conclusion

In this paper, we discussed the relevance principle, which plays a crucial role in the methodology of forensic science, but is potentially relevant for other scientific disciplines as well. In particular, we were focusing on the precise probabilistic version of the principle as put forward by the National Commission on Forensic Science (2016, p. 9). We introduced the basics of the causal Bayesian network machinery and applied it to construct basic causal settings involving a hypothesis about the suspect's guilt G to be evaluated by the trier of fact, a hypothesis H to be assessed by an expert, evidence E for that hypothesis H, and some additional information I that may or may not be used by the expert for evaluating H on the basis of E. We distinguished two possible scenarios: the trier of fact ignoring the specific causal structure underlying the case at hand and updating their credence in G on the basis of the expert's evaluation of H and I as if they were independent – meaning that considering I in addition to H further impacts the trier of fact's credence in G – vs. the trier of fact being a causal reasoner – meaning that they update their credence in G based on the expert's assessment of H, the additional information I, and the causal structure underlying the case at hand.

When assuming that the trier of fact is agnostic about the causal structure, double-counting arises in all of the six basic structures when the relevance principle is not used. The probabilistic version of the relevance principle correctly implies that the expert should ignore I in structures (a)–(d), while it suggests considering I in structures (e) and (f). Thus, the relevance principle does not succeed in avoiding double-counting in (e) and (f). We presented an improved version of the relevance principle simply demanding that the expert should not use any piece of information I also available to the trier of fact, since, as we saw, this is exactly what gave rise to double-counting by the trier of fact. Thus, if the trier of fact is agnostic about the causal structure, no probabilistic relevance principle is required and also a causal principle does not help much to demarcate which additional information I is safe to be used by the expert.

Under the assumption that the trier of fact is a causal reasoner, the situation is somewhat different. While the relevance principle still succeeds in structures (a)–(d), its performance given (e) and (f) improves. While the principle suggests that the expert should use I in both of these structures, the trier of fact, being a causal reasoner, avoids double-counting since they would only consider the part of the information about I that went into the expert's assessment of H when updating their credence in G. Especially the realistic cases (d) and (e) are interesting, since they are somewhat revisionary. While under the assumption that the trier of fact is agnostic about the causal structure, the expert should indeed ignore additional information I to avoid double-counting, they should take this information into account if the trier of fact is a causal reasoner. In both (d) and (e) the responsibility to assess the impact of contamination on H is moved from the trier of fact to the expert. This may even be beneficial, since the expert can be expected to have far more expertise and understanding of the extent of E's confirmatory impact on H and how the analysis might be impacted by other fingerprint results or contamination than the trier of fact. Finally, we proposed an alternative causal version for cases in which the trier of fact updates causally. What is relevant in order to avoid double-counting is not that I must have an impact on E's probability given H, but rather on how H and I are causally connected to G. In particular, if the causal structure guarantees that a competent causal reasoner would not change their credence in G conditional on the expert's assessment of H when considering additional information I (or the other way round), then double-counting on the trier of fact's side will not happen.

What does all of this mean in practice? It shows that the probabilistic version of the relevance principle is not without problems. For each case, there are principles that are guaranteed to avoid double-counting, regardless of the underlying causal structure's complexity. Which principle should one prefer then? The relevance principle* is clearly the safer version, since it would work regardless of whether the trier of fact is a causal reasoner or not, simply by forbidding the expert to use any piece of information I also available to the trier of fact. It requires, however, more coordination between the trier of fact and the experts tasked to assess different pieces and kinds of evidence for the case at hand. Another disadvantage is that in some cases it might be too restrictive because it would forbid the expert to use information I that would increase the reliability and accuracy of their assessment of the hypothesis H under the assumption that the trier of fact is a causal reasoner.

Thus, from a theoretical and formal perspective, it is more likely that the trier of fact arrives at better decisions about *G* if they update based on the causal structure and the expert uses the causal relevance principle** rather than the more restrictive relevance principle* simply forbidding the expert to use information available to the trier of fact. This only requires a basic understanding of causal reasoning that could, for example, be achieved in a crash course. As is known from empirical studies, human beings are much better at causal reasoning than at purely probabilistic reasoning (cf. Hagmayer, 2016). It also requires knowledge of the causal structure by both the trier of fact and the expert. However, as we saw, the causal relations among hypotheses and pieces of information relevant for court decisions and forensic experts' work are usually intuitively clear enough. Thus, we believe that it might be a promising way forward to promote causal reasoning by providing basic training for the trier of fact as well as experts and requiring experts to endorse the causal version of the relevance principle in order to maximize their reliability and accuracy while, at the same time, avoiding double-counting by the trier of fact.

Acknowledgements We would like to thank two anonymous reviewers for helpful comments and suggestions, and especially one reviewer for suggesting to model the suspect's guilt to be assessed by the trier of fact and the hypothesis to be evaluated by the forensic expert as different variables.

Author Contributions Both authors contributed equally to this paper.

Funding Open access funding provided by Università Politecnica delle Marche within the CRUI-CARE Agreement. Funded by the European Union - Next Generation EU, Mission 4, Component 1, CUP I53D23006890001.

Data Availability Statement No data has been used for this article.

Declarations

Conflicts of Interest The authors have no competing interests to declare that are relevant to the content of this article.

Ethical approval n/a

Informed consent n/a

References

Bird, C., Jones, K., & Ballantyne, K.N. (2024). Cognitive bias and contextual information management: Considerations for forensic handwriting examinations. *WIREs Forensic Science*. https://doi.org/10.1002/wfs2.1530

Butt, L. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions-commentary by a forensic examiner. *Journal of Applied Research in Memory and Cognition, 2*, 59–60.

Chin, J. (2014). Psychological science's replicability crisis and what it means for science in the courtroom. *Psychology, Public Policy and Law, 20*, 225–238.

Cooper, G., & Meterko, V. (2019). Cognitive bias research in forensic science: A systematic review. *Forensic Science International, 297*, 35–46.

Cuéllar, M.-F., Mauro, J. A., & Luby, A. (2022). A probabilistic formalisation of contextual bias: From forensic analysis to systemic bias in the criminal justice system. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 185*, S620–S643.

Curley, L.J., Munro, J., Lages, M., Maclean, R., & Murray, J. (2020). Assessing cognitive bias in forensic decisions: A review and outlook. *Journal of Forensic Sciences*, *65,* 354–360.

U.S. Department of Justice. (2022). Uniform language for testimony and reports for forensic autosomal dna examinations using probabilistic genotyping systems [Computer software manual]. Retrieved from https://www.justice.gov/olp/page/file/1095961/dl?inline (Effective December 12, 2022)

Dror, I., Mogan, R., Rando, C., & Nakhaeizadeh, S. (2017). The bias snowball and the bias cascade effects: Two distinct biases that may impact forensic decision making. *Journal of Forensic Sciences*, *62*, 832–833. https://doi.org/10.1111/1556-4029.13496

Elaad, E. (2013). Psychological contamination in forensic decisions. *Journal of Applied Research in Memory and Cognition, 2*, 76–77.

Gebharter, A. (2017). *Causal nets, interventionism, and mechanisms*. Springer.

Gill, P. (2014). *Misleading DNA evidence: Reasons for miscarriages of justice*. Academic Press.

Haack, S. (2015). The expert witness: Lessons from the U.S. experience. *Humana Mente, 28*, 39–70.

Hagmayer, Y. (2016). Causal Bayes nets as psychological theories of causal reasoning: Evidence from psychological research. *Synthese, 193*(4), 1107–1126.

Koehler, J.J., Mnookin, J.L., & Saks, M.J. (2023). The scientific reinvention of forensic science. *Proceedings of the National Academy of Sciences of the United States of America*, *120*, e2301840120.

Kukucka, J. (2020). People who live in ivory towers shouldn't throw stones: A refutation of Curley, et al. *Forensic Science International: Synergy*, *2*,110–113. https://doi.org/10.1016/j.fsisyn.2020.03.001

Kunkler, K.S., & Roy, T. (2023). Reducing the impact of cognitive bias in decision making: Practical actions for forensic science practitioners. *Forensic Science International: Synergy*, *7*, 100341.

National Commission on Forensic Science (2016). Ensuring that forensic analysis is based upon task-relevant information. Retrieved from https://www.justice.gov/ncfs/file/818196/dl?inline

National Court Rules Committee (2024). Federal rules of evidence. Retrieved from https://www.rulesofevidence.org.

National Institute of Standards and Technology (2026). Nist glossary: Evidentiary sample. Retrieved from https://www.nist.gov/glossary-term/42721

National Research Council. (2009). *Strengthening forensic science in the United States: A path forward*. National Academies Press.

Pearl, J. (2000). *Causality* (1st ed.). Cambridge University Press.

Peruzzi, E., & Cevolani, G. (2024). The gatekeeper's dilemma: Daubert standard, economic experts and judicial reasoning. Retrieved from https://ssrn.com/abstract=4763570

Quigley-McBride, A., Dror, I., Roy, T., Garrett, B.L., & Kukucka, J. (2022). A practical tool for information management in forensic decisions: Using linear sequential unmasking-expanded (LSU-E) in casework. *Forensic Science International: Synergy, 4*, 100216.

Saks, M. J., & Koehler, J. J. (2005). The coming paradigm shift in forensic identification science. *Science, 309*(5736), 892–895. https://doi.org/10.1126/science.1111565

Sikorski, M. (2022). Is forensic science in crisis? *Synthese*, *200,* 188. https://doi.org/10.1007/s11229-022-03685-z

Sikorski, M., & Gebharter, A. (2025). The criminalist's paradox as a counterexample to the principle of total evidence. *The British Journal for the Philosophy of Science*. https://doi.org/10.1086/735282

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). MIT Press.

Thompson, W. (1996). DNA evidence in the O.J. Simpson trial. Retrieved from https://api.semanticscholar.org/CorpusID:142766396

Thompson, W. (2011). What role should investigative facts play in the evaluation of scientific evidence. *Australian Journal of Forensic Sciences*, *43*, 123–134.

Thompson, W. (2016). Determining the proper evidentiary basis for an expert opinion. In C. Robertson & A. Kesselheim (Eds.), *Blinding as a solution to bias: Strengthening biomedical science, forensic science, and law* (pp. 133–150). Academic Press.

Thompson, W. (2023). Shifting decision thresholds can undermine the probative value and legal utility of forensic pattern-matching evidence. *Proceedings of the National Academy of Sciences, 120*(41), e2301844120.