The Wrongful Conviction Law Review

“Not Scientific” to Whom? Laypeople Misjudge Manner of Death Determinations as Scientific and Definitive

Jeff Kukucka et Oyinlola Famulegun

Volume 5, numéro 1, 2024

Innocence Project 2024 Scholarship Issue

URI : https://id.erudit.org/iderudit/1111972ar
DOI : https://doi.org/10.29173/wclawr109

Aller au sommaire du numéro

Éditeur(s)

University of Alberta Library

ISSN

2563-2574 (numérique)

Découvrir la revue

Citer cet article

Kukucka, J. & Famulegun, O. (2024). “Not Scientific” to Whom? Laypeople Misjudge Manner of Death Determinations as Scientific and Definitive. *The Wrongful Conviction Law Review*, *5*(1), 42–58. https://doi.org/10.29173/wclawr109

Résumé de l'article

When someone dies unexpectedly, a medical examiner may perform an autopsy to determine how they died (i.e., manner of death). Recent studies found that cognitive bias can affect manner of death judgments, such that extraneous non-medical information may cause the same death to be judged as either a homicide or accident, which has significant legal ramifications. In response, leading medical examiners clarified that manner of death is “not scientific” and “often does not fit well in court.” Yet medical examiners often testify in court, and little is known about how fact-finders appraise their judgments. To address this gap, we conducted two experiments in which mock jurors read and evaluated a medical examiner’s testimony at a murder trial (modeled after the real-world case of Melissa Lucio), while varying the expert’s opinion (i.e., homicide or accident) and the defendant’s attributes (i.e., an affluent white or underprivileged Latina woman). Overall, participants rated the medical examiner’s testimony as highly scientific, credible, and convincing, and it strongly affected their verdicts and belief in the defendant’s guilt, irrespective of the defendant’s attributes. Moreover, participants unexpectedly rated the expert as even more credible if they ruled the death a homicide rather than an accident. Our data thus reveal a worrisome disconnect between how medical examiners characterize their judgments (i.e., as nonscientific and tentative) and how jurors appraise those judgments (i.e., as highly scientific and practically dispositive). We discuss ways to remedy this disconnect, including reforming death investigation practices to curtail bias and encourage standardization and transparency.

"Not Scientific" to Whom? Laypeople Misjudge Manner of Death Determinations as Scientific and Definitive

Jeff Kukucka, Oyinlola Famulegun
Department of Psychology, Towson University,
Towson, MD
U.S.A.

When someone dies unexpectedly, a medical examiner may perform an autopsy to determine how they died (i.e., manner of death). Recent studies found that cognitive bias can affect manner of death judgments, such that extraneous non-medical information may cause the same death to be judged as either a homicide or accident, which has significant legal ramifications. In response, leading medical examiners clarified that manner of death is "not scientific" and "often does not fit well in court." Yet medical examiners often testify in court, and little is known about how factfinders appraise their judgments. To address this gap, we conducted two experiments in which mock jurors read and evaluated a medical examiner's testimony at a murder trial (modeled after the real-world case of Melissa Lucio), while varying the expert's opinion (i.e., homicide or accident) and the defendant's attributes (i.e., an affluent white or underprivileged Latina woman). Overall, participants rated the medical examiner's testimony as highly scientific, credible, and convincing, and it strongly affected their verdicts and belief in the defendant's guilt, irrespective of the defendant's attributes. Moreover, participants unexpectedly rated the expert as even more credible if they ruled the death a homicide rather than an accident. Our data thus reveal a worrisome disconnect between how medical examiners characterize their judgments (i.e., as nonscientific and tentative) and how jurors appraise those judgments (i.e., as highly scientific and practically dispositive). We discuss ways to remedy this disconnect, including reforming death investigation practices to curtail bias and encourage standardization and transparency.

I Introduction

Jurors tend to believe that expert judgments of forensic evidence are objective, dispositive, and virtually infallible (e.g., Crozier *et al*., 2020; Martire *et al*., 2019; Koehler, 2017). However, it is now well-established that forensic experts' judgments are susceptible to *cognitive bias*, such that extraneous influences (e.g., ordering of information, knowledge of irrelevant case facts or theories) can elicit different interpretations of the same evidence (see Kukucka & Dror, 2023). In one study, for example, firearms experts changed 28% of their opinions of the same bullets when given different information about where those bullets were found (Kerstholt *et al*., 2010). Cognitive bias has likewise been shown to affect expert judgments of DNA mixtures, fingerprints, toxicology results, bloodstain patterns, skeletal remains, and digital evidence, among others (see Kukucka & Dror, 2023). When cognitive bias prompts conflicting opinions of the same evidence, one of those opinions must be incorrect, which can produce a miscarriage of justice—and indeed, forensic science errors have now been implicated in over 850 wrongful convictions (National Registry of Exonerations, 2023), including 53% of DNA exonerations (Innocence Project, 2023).

Medicolegal judgments are likewise susceptible to cognitive bias, such that extraneous information can lead medical experts to interpret the same injury as either criminal or accidental (see Kukucka & Findley, 2023). In one study, for example, physicians and nurses more often misdiagnosed an accidental pediatric leg fracture as child abuse if led to believe that the child's parents were unmarried and had a strained relationship—information that should have no bearing on their medical judgment (Anderst *et al*., 2016). Similarly, another study found that hospital staff more often interpreted the same radiographs as evidence of child abuse if given prejudicial but medically irrelevant information about the child's parents (e.g., that they were unemployed, unmarried, immigrants, drug users, or welfare recipients; Loos *et al*., 2021).

Dror *et al*. (2021) showed that cognitive bias can also affect postmortem *manner of death* determinations (i.e., opinions as to whether a sudden unexpected death was an accident, homicide, suicide, natural, or indeterminate—which, depending on the jurisdiction, may be rendered by a coroner, forensic pathologist, or medical examiner). In their study, medical experts read a case file about a child who was found unresponsive at home, taken to a hospital, and died soon thereafter, and they were asked to opine as to how the child had died. All of the experts received the same information about the child's injuries and medical history; however, some were told that the child was white and brought to the hospital by her grandmother, while others were told that the child was Black and brought to the hospital by her mother's boyfriend. These extraneous details had a strong effect on manner of death opinions: Despite identical medical information, experts in the latter group were about 12 times more likely to rule the child's death a homicide.

Dror *et al*.'s findings sparked strong and varied reactions from the medical examiner community. Some were supportive, lauding the researchers for "call[ing] attention to the powerful impact that bias can have on [medical examiners'] diagnostic decisions" and "offer[ing] important suggestions on how to mitigate the impact of bias" (Graber, 2021; see also Duflou, 2021; Obenson,

2021). Others were extremely hostile, dismissing Dror *et al*.'s results as invalid and politically motivated, levying baseless allegations of research misconduct (all of which were thoroughly investigated and found to have no merit), demanding that the paper be retracted, and threatening legal action against the authors, journal, and publisher (see Peat, 2021).

Importantly, Dror *et al*.'s critics also took this opportunity to clarify the nature and purpose of manner of death determinations. A commentary by 86 prominent forensic pathologists from 29 U.S. states (including the then-President of the National Association of Medical Examiners) explained that "manner determination is not a 'scientific' determination" but rather "a cultural determination… for the purpose of public health statistics" for which "there is no 'right' answer [or] criteria for correctness" and which "often does not fit well in court" (Peterson *et al*., 2021a). Other medical examiners expressed similar viewpoints, including that "manner determination is a nonscientific determination for statistical purposes [such that] any individual determination is questionable" (Oliver, 2021). Indeed, medical examiners often reach different opinions of the same death even in the absence of biasing extraneous information: In one study, for example, 152 experts determined manner of death for 30 case scenarios, only four of which produced greater than 90% consensus, while 17 showed less than 70% consensus (Hanzlick *et al*., 2015).

Nevertheless, medical examiners *do* present manner determinations in courtrooms, where jurors likely consider them highly scientific just as they do other forensic expert testimony. Take for example the case of Melissa Lucio, who is the first Hispanic woman to be sentenced to death in Texas. In February 2007, Melissa's two-year-old daughter Mariah fell down a flight of stairs and suddenly died two days later. Melissa was later convicted of murdering her daughter based largely on testimony from Dr. Norma Jean Farley, who performed Mariah's autopsy and ruled her death a homicide, opining that her injuries could only have resulted from abuse.

Critically, however, two police officers had attended Mariah's autopsy and told Dr. Farley in advance that Melissa had allegedly confessed to abusing her daughter, which raises concern that Farley's determination was biased by this and/or other extraneous information (e.g., that Melissa was an unmarried, impoverished, former drug user). Accordingly, numerous independent experts have since repudiated Farley's testimony in this case; as one pathologist wrote, "the investigation into Mariah's death appears to have been significantly prejudiced, not evidence based, and without an adequate consideration of alternative issues" (Ophoven, 2022).

Unfortunately, jurors are poor at recognizing and discounting biased forensic expert testimony (Kukucka *et al*., 2020; Thompson & Scurich, 2019)—or worse, jurors may actually favor biased testimony insofar as it reinforces their own biases. Research on *stereotype congruence* has shown that the degree of correspondence between a defendant's traits and the stereotypic traits of a criminal offender affects how jurors process and appraise information (Jones & Kaplan, 2003; McKimmie *et al*., 2013), including forensic science evidence (Smalarz *et al*., 2016). As a result, jurors may value the same evidence differently depending on the defendant's characteristics; in one study, for example, jurors found a confession more compelling if the defendant fit a criminal stereotype, even if the confession was coerced (Smalarz *et al.*, 2018). Similarly, jurors may deem the same expert testimony more credible—regardless of its objective validity—when the defendant's characteristics stereotypically congruent with the expert's opinion.

We therefore designed the current studies with two aims. First, in light of pathologists' assertions that manner of death determinations are not scientific, we aimed to test whether jurors feel similarly—or whether there is a worrisome disconnect between how experts and non-experts regard manner determinations. Second, given that stereotypes and other extraneous information have been shown to affect both expert and juror decisions, we aimed to test whether jurors' evaluations of the same medical expert testimony depend on the stereotypic congruence between the expert's opinion and the defendant's traits—e.g., jurors may consider inculpatory testimony more credible if the defendant matches a criminal stereotype, and vice versa.

To address these questions, we conducted two experiments (one with college students and one with community members) that followed the same procedure and used materials modeled closely after the aforementioned case of Melissa Lucio. In both studies, participants acted as mock jurors in the trial of a mother who was accused of murdering her daughter. First, participants read background information about the case, which described the defendant as either an affluent white woman or an underprivileged Latina woman. Then, participants read the trial testimony of the medical examiner who performed the child's autopsy, which described the child's injuries and ruled the death either an accident or a homicide. Finally, participants rendered a verdict, estimated the likelihood of the defendant's guilt, and rated the degree to which the medical examiner's testimony was scientific, credible, and convincing. Our hypotheses were as follows:

> **H1:** The medical examiner's testimony will affect participants' belief in the defendant's guilt, such that they will more often judge the defendant as guilty if the medical examiner testifies that the child's death was a homicide rather than an accident.
>
> **H2:** Participants' evaluations of the medical examiner's testimony will depend on the stereotypic congruence between the manner of death determination and the defendant's characteristics. If the medical examiner rules the child's death a homicide, jurors will rate their testimony more favorably if the defendant is an underprivileged Latina woman as opposed to an affluent white woman. Conversely, if the medical examiner rules the child's death an accident, jurors will rate their testimony more favorably if the defendant is an affluent white woman as opposed to an underprivileged Latina woman.

II Method

All materials are available on the Open Science Framework (http://osf.io/jekvm).

A. Participants and Design

Study 1 included 319 college students who completed the study online for course credit. Study 2 included 251 community members who completed the study online via Prolific for a $3.00 credit. In both studies, each participant was randomly assigned to one of four cells in a 2 (Stereotype: Innocent vs. Guilty) X 2 (Opinion: Accident vs. Homicide) between-person design. We later excluded data from participants who failed a comprehension test (n = 107 in Study 1; n = 32 in Study 2), leaving samples of N = 212 and 219 for Studies 1 and 2, respectively.

Study 1 participants were mostly female (81.1%; 15.6% male, 3.3% non-binary) and white (46.7%; 27.8% Black, 8.5% Hispanic), with a mean age of 19.1 (*SD* = 2.8; range = 18 – 43). Study 2 participants were also mostly female (53.4%; 40.2% male, 4.6% non-binary) and white (64.8%; 11.0% Black, 9.6% Hispanic), with a mean age of 38.6 (*SD* = 13.1; range = 18 – 75).

B. Procedure

Both studies followed the same procedure. First, participants were asked to imagine being a juror in the murder trial of a woman who is charged with killing her two-year-old daughter, and they read a narrative summary of the child's death and the ensuing investigation (approximately 500 words). By random assignment, this narrative described the defendant as either an affluent, married, white stay-at-home mother (*Innocent Stereotype* condition) or a poor, unmarried, unemployed Hispanic mother with a history of drug use (*Guilty Stereotype* condition), but all other details of the narrative (as explained below) were identical between conditions.

Next, participants read a 12-page transcript of trial testimony from the medical examiner who performed the child's autopsy, in which she described her credentials and the child's injuries. By random assignment, the medical examiner concluded that the child died from either falling down the stairs (*Accident* condition) or being physically abused (*Homicide* condition), though her description of the child's injuries was identical between conditions.

After reading the transcript, participants rated the degree to which they found the medical examiner's testimony to be scientific, credible, and convincing. Then, they rendered a verdict in the case and estimated the likelihood that the defendant had killed her daughter. Lastly, participants completed a comprehension test and a demographic questionnaire.

C. Materials

a. Narrative Summary

By random assignment, each participant first read one of two versions of a narrative that provided background information about the case. The two versions differed in terms of how they described the defendant. In the *Innocent Stereotype* condition, the defendant was Natalie Martin—a 38-year-old woman who was born in the U.S., graduated college, and worked as a museum curator until becoming pregnant with her first child. She is now a stay-at-home mom who owns a house with her husband and two children. In the *Guilty Stereotype* condition, the defendant was Natalia Martinez—a 38-year-old woman who immigrated to the U.S, dropped out of high school, and worked as a museum janitor until entering a drug treatment program. She is now unemployed and rents an apartment with her boyfriend and two children from prior relationships.

All other details of the narrative were identical between conditions and were modeled after the case of Melissa Lucio. In both versions, the defendant was unable to wake her two-year-old daughter Olivia from a nap, so she called 911, but emergency personnel were unable to resuscitate Olivia, and she was pronounced dead at 5:30pm. Police then became suspicious of the defendant because of her body language and "unusually calm demeanor," so they interviewed her at the police station from 9:30pm until 3am, during which they accused her of abusing her daughter.

Although she denied this at first, the defendant eventually "accepted responsibility" for her daughter's death, but she later recanted that statement, claiming that police had coerced her. Nonetheless, the defendant was charged with capital murder and pleaded not guilty.

The narrative then explained that the defense is arguing that Olivia's injuries and death were due to her falling down a "steep flight of wooden stairs" two days earlier. The prosecution is not disputing that Olivia fell down the stairs, but they are arguing that her injuries and death were due to physical abuse. Lastly, the narrative explained that participants would now read and evaluate the trial testimony of the medical examiner who performed Olivia's autopsy.

b. Expert Testimony

Next, participants read a 12-page transcript of the trial testimony of the medical examiner who performed Olivia's autopsy, who we named Dr. Claudia J. Farris. To develop this transcript, we began with the original 55-page transcript of Dr. Norma Jean Farley's testimony at Melissa Lucio's 2008 murder trial, changed the names of the individuals involved and/or discussed, and removed portions that were purely procedural (e.g., discussing exhibit numbers, objections, sidebars) or redundant (e.g., repeated questions). Then, we truncated the remainder of the transcript in a way that maintained its overall content and structure: Dr. Farris explained her qualifications (~1 page), explained the duties of a forensic pathologist in general terms (~1 page), described the injuries that she observed during Olivia's autopsy (including her manner of death determination; ~7 pages), was cross-examined (~2 pages), and was briefly re-directed (~1 page).

Participants who were randomly assigned to the *Homicide* condition read this truncated version of the original transcript, which consisted of near-verbatim statements made by Dr. Farley (now Dr. Farris) that described Olivia's injuries as non-accidental (e.g., "This child was severely abused"), refuted the alternative (e.g., "The pattern of injuries from a fall generally wouldn't look like this"), and ruled her death a homicide (e.g., "The manner was homicide"). Dr. Farris then conceded on cross-examination that it is possible for a child to die from falling downstairs, and then reiterated on re-direct her conclusion that Olivia's death was the result of abuse.

Participants who were randomly assigned to the *Accident* condition read a minimally altered version of this transcript in which we changed Dr. Farris' interpretation of Olivia's injuries, but not her description of those injuries. All told, we changed only 106 of the 3,322 words (3.2%) between the two versions of the transcript. Paralleling the Homicide transcript, the Accident transcript described Olivia's injuries as accidental (e.g., "This child suffered a severe fall"), refuted the alternative (e.g., "The pattern of injuries from abuse generally wouldn't look like this"), and ruled her death an accident (e.g., "The manner was accident"). Dr. Farris then conceded on cross-examination that it is possible for a child to die from being physically abused, and then reiterated on redirect her conclusion that Olivia's death was the result of a severe fall.

D. Measures

a. Ratings of Expert Testimony

Participants rated the degree to which they felt that the medical examiner's testimony was scientific, credible, and convincing, each on a separate scale from 1 (*not at all*) to 5 (*very*).

b. Guilt Judgments

Participants rendered a verdict (i.e., guilty or not guilty) and estimated the likelihood that the defendant had in fact murdered her daughter, using a sliding scale from 0% to 100%.

c. Comprehension Test

Lastly, participants completed a five-item multiple-choice test to ensure that they read and understood the narrative and transcript. Two items pertained to details of our manipulations, including: (a) whether the defendant was previously a museum curator (*Innocent Stereotype* condition) or in a drug treatment program (*Guilty Stereotype* condition) and (b) whether the medical examiner ruled the child's death an accident (*Accident* condition) or a homicide (*Homicide* condition). Prior to analysis, we excluded data from 107 (33.5%) and 32 (12.7%) participants who answered one or both of these items incorrectly in Studies 1 and 2, respectively.

III Results

A. Analytic Plan

Data from both studies are available on the Open Science Framework (http://osf.io/jekvm). For verdicts, we performed a binary logistic regression with Stereotype (Innocent vs. Guilty), Opinion (Accident vs. Homicide), and their interaction as predictors of verdict (i.e., guilty or not guilty). For each numerical measure (i.e., likelihood of guilt and three ratings of the medical examiner's testimony), we performed both frequentist and Bayesian 2 (Stereotype) X 2 (Opinion) ANOVAs. Descriptive statistics for these numerical measures are shown in Table 1, and ANOVA results are shown in Table 2. Bayesian analyses compare the relative strength of the evidence for an effect and the evidence against an effect; Bayes factors (i.e., BF_{10} values; see Table 2) greater than 1 indicate stronger evidence of an effect, whereas BF_{10} values less than 1 indicate stronger evidence of no effect (see Quintana & Williams, 2018).

Table 1. Means (and Standard Errors) for Ratings of Guilt (0-100) and Expert Testimony (1-5)

			Stereotype		Opinion	
		Overall	Innocent	Guilty	Accident	Homicide
Likelihood of Guilt	Study 1	63.43 (1.85)	61.34 (2.22)	64.82 (2.14)	48.16_a (2.20)	78.01_b (2.16)
	Study 2	55.58 (2.39)	56.58 (2.09)	54.84 (2.02)	27.40_a (2.02)	84.02_b (2.04)
Scientific	Study 1	4.33 (.05)	4.39 (.08)	4.29 (.07)	4.42 (.07)	4.26 (.07)
	Study 2	4.65 (.04)	4.71 (.06)	4.60 (.06)	4.66 (.06)	4.66 (.06)

Credible	Study 1	4.33 (.05)	4.40 (.08)	4.27 (.07)	4.20_a (.08)	4.47_b (.07)
	Study 2	4.63 (.04)	4.60 (.06)	4.66 (.06)	4.54_a (.06)	4.72_b (.06)
Convincing	Study 1	4.13 (.06)	4.22 (.09)	4.04 (.08)	3.92_a (.09)	4.34_b (.08)
	Study 2	4.50 (.05)	4.50 (.08)	4.50 (.07)	4.42 (.08)	4.59 (.08)

Note. Means not sharing a common subscript were significantly different at $p < .05$ (see Table 2).

Table 2. Results of 2 (Stereotype) X 2 (Opinion) ANOVAs on Ratings of Guilt and Expert Testimony

		Stereotype		Opinion		Interaction	
		F	BF_{10}	F	BF_{10}	F	BF_{10}
Likelihood of Guilt	Study 1	1.27	0.26	93.90***	4.21×10^{15}	0.04	0.21
	Study 2	0.37	0.16	389.81***	1.19×10^{47}	0.46	0.24
Scientific	Study 1	1.07	0.25	2.14	0.42	0.00	0.22
	Study 2	1.62	0.32	0.00	0.15	0.16	0.22
Credible	Study 1	1.70	0.30	6.98**	4.02	1.78	0.47
	Study 2	0.52	0.21	4.80*	1.58	0.01	0.20
Convincing	Study 1	2.08	0.33	12.48***	49.45	3.60	1.00
	Study 2	0.00	0.15	2.42	0.47	0.00	0.22

Note. *$p < .05$; **$p < .01$; ***$p < .001$

B. Verdicts

Consistent with H1, a 2 (Stereotype: Innocent vs. Guilty) X 2 (Opinion: Accident vs. Homicide) factorial logistic regression on verdicts revealed significant effects of Opinion in both Study 1, Wald $\chi^2(1) = 34.75$, $p < .001$, $OR = 2.65$ [95% CI: 1.92, 3.67], and Study 2, Wald $\chi^2(1) = 88.36$, $p < .001$, $OR = 8.72$ [95% CI: 5.55, 13.70], such that participants more often judged the defendant as guilty when the medical examiner ruled the death a homicide (84.3% in Study 1; 90.8% in Study 2) as opposed to an accident (43.3% in Study 1; 12.7% in Study 2).

There was no effect of Stereotype in either Study 1, $\chi^2(1) = 0.44$, $p = .506$, $OR = 0.90$ [95% CI: 0.65, 1.24], or Study 2, Wald $\chi^2(1) = 1.76$, $p = .184$, $OR = 0.87$ [95% CI: 0.86, 2.13]. Moreover, the Stereotype X Opinion interaction was not significant in either Study 1, Wald $\chi^2(1) = 0.03$, $p = .874$, or Study 2, Wald $\chi^2(1) = 0.87$, $p = .351$, indicating that the medical examiner's testimony had an equivalent effect on verdicts regardless of the defendant's characteristics.

C. Likelihood of Guilt

Further supporting H1, a 2 (Stereotype) X 2 (Opinion) ANOVA on likelihood of guilt estimates (0-100%) revealed strong effects of Opinion in both studies (see Table 2), such that participants were more confident in the defendant's guilt when the medical examiner ruled the death a homicide as opposed to an accident (see Table 1 for descriptive statistics).

D. Ratings of Expert Testimony

Overall (i.e., collapsed across studies and conditions), participants rated the medical examiner's testimony as highly scientific (M = 4.50 out of 5, SE = .03), highly credible (M = 4.48, SE = .04), and highly convincing (M = 4.32, SE = .04). In each case, most participants selected either '5' (60.8% for scientific, 60.6% for credible, 53.4% for convincing) or '4' (29.2% for scientific, 28.5% for credible, 30.2% for convincing), which indicates that a sizeable majority of participants considered the medical examiner's testimony to be highly scientific, credible, and convincing regardless of the manner determination or the defendant's characteristics.

As shown in Table 2, there were no main effects of Stereotype on any of these ratings in either study, such that judgments of the medical examiner's testimony were uniformly positive regardless of the defendant's characteristics. However, we unexpectedly found main effects of Opinion on credibility ratings in both studies and on convincingness ratings in Study 1 (but not Study 2), such that—irrespective of the defendant—participants rated the medical examiner's testimony as more credible if they ruled the death a homicide rather than an accident.

Finally, and contrary to H2, there were no significant Stereotype X Opinion interactions on any ratings of the medical examiner's testimony in either study. Thus, participants rated the medical examiner's testimony as equally (and highly) scientific regardless of the stereotypic congruence (or lack thereof) between the examiner's manner determination and the defendant's characteristics, and participants consistently rated homicide determinations as more credible than accident determinations regardless of the defendant's characteristics.

IV Discussion

Dror *et al.* (2021) demonstrated that extraneous information can affect medical experts' judgments of whether a sudden death was a homicide or an accident—judgments that often hold legal ramifications insofar as they imply whether a crime was committed. In response, leading forensic pathologists clarified that manner of death determinations are "not scientific" (Peterson *et al*., 2021a) and that "any individual determination is questionable" (Oliver, 2021). Despite this, our data show that laypeople find manner determinations very compelling: Across two studies, 90% of mock jurors rated a medical examiner's testimony about manner of death as highly scientific (i.e., a 4 or 5 on a 5-point scale), 89% rated it as highly credible, and 84% rated it as highly convincing. Accordingly, the medical examiner's opinion was practically dispositive on trial outcomes: 88% of mock jurors judged the defendant as guilty of murder if the medical examiner ruled the death a homicide, compared to only 28% if they ruled the same death an

accident. Our data thus reveal a troubling disconnect between how non-experts regard manner determinations and how experts feel that their determinations *should* be regarded.

Contrary to our prediction, participants' evaluations of the medical examiner's testimony were equally favorable regardless of whether the defendant fit a criminal stereotype, which may reflect a ceiling effect. However, we unexpectedly found that jurors rated the testimony as more credible (in both studies) and more convincing (in Study 1 only) when the medical examiner ruled the death a homicide rather than an accident, even though the circumstances of the death were otherwise identical. The presumption of innocence—i.e., that jurors should assume the defendant to be innocent unless and until they are proven guilty—is a cornerstone of the American legal system. However, this finding suggests that jurors find otherwise-identical testimony more persuasive when it is inculpatory rather than exculpatory, which casts doubt on their ability to follow this principle. Other studies have likewise found that jurors initially estimate the probability of a defendant's guilt as around 50% (when it should be 0%; Scurich *et al*., 2016), that merely being charged with a crime invites stronger presumptions of guilt (Scurich & John, 2017), and that a sizeable minority of individuals believe it is worse to wrongly acquit a guilty person than to wrongly convict an innocent person (Garrett & Mitchell, 2022). Thus, our data raise the concerning possibility that jurors not only show unwarranted confidence in manner of death determinations in general, but also inherently trust incriminating determinations more than others.

Legal scholars have argued that expert testimony on manner of death should not be admissible, arguing that it is both unreliable and so dispositive that it usurps the factfinding role of the jury (Findley & Strang, 2022; Simon, 2019). Some medical examiners have likewise explained that manner determinations do not belong in courtrooms, noting that they are meant to inform "public health statistics… not trial results" (Peterson *et al.*, 2018b) and voicing concern over "misuse of manner determination by the courts" (Peterson *et al*., 2018a). Notably, manner determination is an American invention; in most other countries, coroners do not certify manner or testify in court (e.g., Oliver, 2014). For example, Australian death investigators do not classify "manner of death into ascribed, clearly defined categories," nor do they include it on death certificates (Phillips *et al*., 2015). In Italy, "the death certificate is not used in court as evidence of the cause and manner of death, but is used only to prove that a person is dead" (Di Vella & Campobasso, 2015). And in India, "the autopsy physician opines only on the cause of death, [whereas] manner of death is determined by the police" (Sharma & Bajpai, 2015).

Short of eliminating manner determinations altogether, how else might we discourage factfinders from giving undue weight to potentially unreliable manner of death opinions? Above all, medical examiners should make abundantly clear in court—just as they have done in writing—that manner determinations are "nonscientific," there are no "criteria for correctness," and their aim is not to find the "right answer" (Peterson *et al*., 2021a). Furthermore, they should be fully transparent about the rationale for their manner determination—i.e., the extent to which it is based on medical information within the purview of their expertise, as opposed to non-medical information gleaned from police or other sources—and if it is based solely on the latter (as in the infamous case of *Iowa v. Tyler*; see Weedn, 2021), then it should not be admissible.

However, these ostensibly simple solutions may not be as foolproof as they appear. First, even if experts' deficiencies are made clear, the mystique of expertise may overshadow those

deficiencies, as in one study where mock jurors rated forensic experts as highly skilled, competent, and convincing even if they admitted to a staggering error rate (~30%; Crozier *et al*., 2020). Second, it is well-known that people have only limited ability to introspect on their own decision-making processes (e.g., Nisbett & Wilson, 1977) and are often blind to their own biases (e.g., Pronin & Hazel, 2023), such that medical examiners may be unable to faithfully articulate the rationale for their determinations even if they are willing. Third, when experts claim to be immune to bias (Kukucka *et al*., 2017), jurors may believe them; in one study, for example, jurors equally trusted forensic examiners who denied that biasing information had affected their opinion and other examiners who never received that information (Kukucka *et al*., 2020).

When medical examiners neglect to explain the limitations and/or basis of their manner determinations, the onus falls on attorneys to draw out that information. Indeed, FRE Rule 705 states that expert witnesses need not disclose the basis for their opinion unless cross-examination requires it. Unfortunately, it appears that attorneys often fail to detect unreliable manner of death opinions. Despodova *et al*. (2020) had defense attorneys imagine representing a client charged with murder and review a case file, including an autopsy report that ruled the death a homicide. They found that attorneys rated the autopsy as equally probative and reliable regardless of whether it was patently biased (i.e., the medical examiner admitted that his knowledge of the defendant's recanted confession colored his interpretation of the decedent's injuries) or unbiased (i.e., he was unaware of the confession). Moreover, fewer than half of attorneys who read the patently biased autopsy report said that they would raise the issue of bias on cross-examination.

For attorneys who do cross-examine forensic experts, its efficacy remains unclear. Some studies found that cross-examination had little or no effect on jurors' trust in dubious testimony (e.g., Garrett *et al*., 2020; McQuiston-Surrett & Saks, 2009), while others found that it lessened trust indiscriminately (e.g., Scanlon *et al*., 2021)—neither of which is ideal. Cross-examination should ideally sensitize jurors to the quality of forensic testimony so that they discount it only when it is unreliable. Some research suggests this may be possible (e.g., Crozier *et al*., 2020; Thompson & Scurich, 2019), but more work is needed to identify specific effective approaches. While our study included a cross-examination, it was rather benign and constant across conditions, and so we cannot say whether or how it affected jurors' decision-making. Some studies have tested other ways to make jurors more discerning in their reliance on forensic expert testimony—such as judicial instructions (Eastwood & Caldwell, 2015), visual aids (Ribeiro *et al.*, 2023), or educational videos (LaBat *et al*., 2023)—and those methods have also yielded mixed results. Future research should continue to examine how laypeople appraise expert testimony from medical examiners specifically, including how cross-examination might influence those appraisals.

A. Policy Recommendations

Rather than relying on attorneys and jurors to detect, expose, and devalue unreliable manner determinations in the courtroom, it would be more effective to address the problem further upstream—by reforming medicolegal death investigation practices in ways that minimize the risk of unreliable manner determinations in the first place. Dror and Kukucka (2021) have described a procedure called *Linear Sequential Unmasking—Expanded* (LSU-E) for reducing bias in decision-making in both forensic and non-forensic arenas. Under this procedure, decision-makers make thoughtful *a priori* decisions about what information should or should not inform their decision,

prioritize information that is most relevant and objective, and document the information they considered and how their opinion evolved over time. Variants of this procedure have now been endorsed and/or adopted by practitioners in a range of forensic disciplines (e.g., Archer & Wallman, 2016; Dahal *et al*., 2022; Found & Ganas, 2013; Whitehead *et al.*, 2022).

In medicolegal death investigations, LSU-E would advise medical examiners to complete the autopsy and document a tentative manner of death opinion based solely on the observations therein *before* considering any other contextual information that might be relevant to their determination, while also strictly avoiding information that is irrelevant to their decision (see also Simon, 2019, for a proposed typology of information that is always, sometimes, or never relevant to manner determinations). Some pathologists have now endorsed LSU-E in writing, calling it an "important suggestion on how to mitigate the impact of bias" (Graber, 2021; see also Ko & Glusac, 2023). Moreover, Dr. Andrew Baker, tacitly endorsed LSU-E at the murder trial of Derek Chauvin when he testified that he "intentionally chose not to" watch the cell phone video of George Floyd's death prior to performing Mr. Floyd's autopsy because he "did not want to bias [the] exam by going in with any preconceived notions." Although Dr. Baker did not describe it as such, this calculated approach is entirely consistent with what LSU-E prescribes.

Manner determinations should also be as standardized, independent, and transparent as possible. Standardization tends to protect against bias insofar as it lessens room for interpretation and disagreement, which, as noted above, is endemic to manner determinations (Hanzlick *et al*., 2015). In Italy, for example, death certificates include instructions and examples of proper manner determinations in an effort to promote consistency between agencies and examiners (Di Vella & Campobasso, 2015). Medical examiners should also operate independently of law enforcement agencies and prosecutors' offices, so as to minimize the risk of overt pressure to produce a certain determination (Luzi *et al*., 2013) and/or implicit allegiance effects (Murrie *et al*., 2013). Finally, consistent with LSU-E, medical examiners should be transparent about what information informed their manner determination (see also Quigley-McBride *et al*., 2022).

Maryland's Office of the Medical Examiner (OCME) recently agreed to reforms of this nature as part of a settlement pertaining to the death of Anton Black, a Black teenager who died while being restrained by police in 2018. To be exact, the OCME will now (a) have a clear policy for the handling of in-custody deaths that follows NAME standards for homicide determinations (i.e., the "but for" principle), (b) prohibit non-OCME personnel (e.g., law enforcement) from giving input on autopsies, and (c) document the presence of law enforcement during autopsy and any investigative information received therefrom. At the same time, Maryland's Office of the Attorney General is conducting an independent audit of OCME's handling of past deaths involving police restraint for signs of bias and/or inappropriate procedures, with an eye toward improving death investigation practices and creating a model for other states to follow.

In sum, the current studies reveal a problematic disconnect between how experts and non-experts regard manner of death determinations, such that jurors consider them highly scientific and persuasive, whereas practitioners acknowledge their dubiety and question their suitability for courtroom presentation. To bridge this gap, we must better educate factfinders on the limitations of manner determinations and/or reform the processes by which medical examiners make and communicate those determinations. Either will require collaboration between psychologists,

attorneys, and medical examiners who share an ambition to optimize death investigation practices in ways that minimize the risk of miscarriages of justice—and we hope that the current studies will inspire more scholars and practitioners to undertake such efforts.

V References

Anderst, J., Nielsen-Parker, M., Moffatt, M., Frazier, T., & Kennedy, C. Using Simulation to Identify Sources of Medical Diagnostic Error in Child Physical Abuse. (2016) Child Abuse & Neglect, 52, 62-69. Online: doi:10.1016/j.chiabu.2015.12.015

Archer, M. S., & Wallman, J. F. Context Effects in Forensic Entomology and Use of Sequential Unmasking in Casework. (2016) Journal of Forensic Sciences, 61, 1270-1277. Online: doi:10.1111/1556-4029.13139

Crozier, W. E., Kukucka, J., & Garrett, B. L. Juror Appraisals of Forensic Evidence: Effects of Blind Proficiency and Cross-Examination. (2020) Forensic Science International, 315, 110433. Online: doi:10.1016/j.forsciint.2020.110433

Dahal, S., Chaudhary, G. K., & Agrawal, N. K. Operation Makalu air crash: Influence of Cognitive and Human Factors on Decision-Making. (2022) Forensic Sciences Research, 7, 803-807. Online: doi:10.1080/20961790.2022.2095691

Despodova, N. M., Kukucka, J., & Hiley, A. Can Defense Attorneys Detect Forensic Confirmation Bias? Effects on Evidentiary Judgments and Trial Strategies. (2020) Zeitschrift für Psychologie, 228, 216-220. Online: doi:10.1027/2151-2604/a000414

Di Vella, G., & Campobasso, C. P. Death Investigation and Certification in Italy. (2015) Academic Forensic Pathology, 5, 454-461. Online: doi:10.23907/2015.050

Dror, I. E., & Kukucka, J. Linear Sequential Unmasking–Expanded (LSU-E): A General Approach for Improving Decision-Making as Well as Minimizing Bias. (2021) Forensic Science International: Synergy, 3, 100161. Online: doi:10.1016/j.fsisyn.2021.100161

Dror, I. E., Melinek, J., Arden, J. L., Kukucka, J., Hawkins, S., Carter, J., & Atherton, D. Cognitive Bias in Forensic Pathology Decisions. (2021) Journal of Forensic Sciences, 66, 1751-1757. Online: doi:10.1111/1556-4029.14697

Duflou, J. Commentary on Dror *et al*. 'Cognitive Bias in Forensic Pathology Decisions.' (2021) Journal of Forensic Sciences, 66, 2561. Online: doi:10.1111/1556-4029.14836

Eastwood, J., & Caldwell, J. Educating Jurors About Forensic Evidence: Using an Expert Witness and Judicial Instructions to Mitigate the Impact of Invalid Forensic Science Testimony. (2015) Journal of Forensic Sciences, 60, 1523-1528. Online: doi:10.1111/1556-4029.12832

Findley, K. A., & Strang, D. A. Ending Manner-Of-Death Testimony and Other Opinion Determinations of Crime. (2022) Duquesne University Law Review, 60, 302-337.

Found, B., & Ganas, J. The Management of Domain Irrelevant Context Information in Forensic Handwriting Examination Casework. (2013) Science & Justice, 53, 154-158. Online: doi:10.1016/j.scijus.2012.10.004

Garrett, B. L., & Mitchell, G. Error Aversions and Due Process. (2022) Michigan Law Review, 121, 707-751.

Garrett, B. L., Scurich, N., & Crozier, W. E. Mock Jurors' Evaluation of Firearm Examiner Testimony. (2020) Law and Human Behavior, 44, 412-423. Online: doi:10.1037/lhb0000423

Graber, M. L. Commentary on Dror *et al*. 'Cognitive Bias in Forensic Pathology Decisions.' (2021) Journal of Forensic Sciences, 66, 2574. Online: doi:10.1111/15564029.14857

Hanzlick, R. L., Goodin, J., & Haden-Pinneri, K. Mind Your Manners: 20 Years Later. (2015) Academic Forensic Pathology, 5, 380-395. Online: doi:10.23907/2015.042

Innocence Project. (2023). *Misapplication of Forensic Science.* Online: https://innocenceproject.org/misapplication-of-forensic-science/

Jones, C. S., & Kaplan, M. F. The Effects of Racially Stereotypical Crimes on Juror Decision-Making and Information-Processing Strategies. (2003) Basic and Applied Social Psychology, *25*, 1-13. Online: doi:10.1207/S15324834BASP2501_1

Kerstholt, J., Eikelboom, A., Dijkman, T., Stoel, R., Hermsen, R., & van Leuven, B. Does Suggestive Information Cause a Confirmation Bias in Bullet Comparisons? (2010) Forensic Science International, 198, 138-142. Online: doi:10.1016/j.forsciint.2010.02.007

Ko, C. J., & Glusac, E. J. Cognitive Bias in Pathology, As Exemplified in Dermatopathology. (2023) Human Pathology, 140, 267-275. Online: doi:10.1016/j.humpath.2023.03.003

Koehler, J. J. Intuitive Error Rate Estimates for the Forensic Sciences. (2017) Jurimetrics, 57, 153-168.

Kukucka, J., & Dror, I. E. Human Factors in Forensic Science: Psychological Causes of Bias and Error. In D. DeMatteo & K. C. Scherr (Eds.), (2023) The Oxford Handbook of Psychology and Law (pp. 621-642). Oxford University Press.

Kukucka, J., & Findley, K. A. (2023). Cognitive Bias in Medicolegal Judgments. In K. A. Findley, C. Rossant, K. Sasakura, L. Schneps, W. Squier, & K. Wester (Eds.), *Shaken Baby Syndrome: Investigating the Abusive Head Trauma Controversy* (pp. 205-217). Cambridge University Press.

Kukucka, J., Hiley, A., & Kassin, S. M. Forensic Confirmation Bias: Do Jurors Discount Examiners Who Were Exposed to Task-Irrelevant Information? (2020) Journal of Forensic Sciences, 65, 1978-1990. Online: doi:10.1111/1556-4029.14546

Kukucka, J., Kassin, S. M., Zapf, P. A., & Dror, I. E. Cognitive Bias and Blindness: A Global Survey of Forensic Science Examiners. (2017) Journal of Applied Research in Memory and Cognition, 6, 452-459. Online: doi:10.1016/j.jarmac.2017.09.001

LaBat, D. E., Goldfarb, D., Evans, J. R., Compo, N. S., Koolmees, C. J., LaPorte, G., & Lothridge, K. Improving Juror Assessments of Forensic Testimony and Its Effects on Decision-Making and Evidence Evaluation. (2023) Law and Human Behavior, 47, 566–578. Online: doi:10.1037/lhb0000539

Loos, M. L. H., Allema, W. M., Bakx, R., Stoel, R. D., van Rijn, R. R., & Karst, W. A. Paediatric Femur Fractures—The Value of Contextual Information on Judgement in Possible Child Abuse Cases: Are We Bias? (2021) European Journal of Pediatrics, 180, 81-90. Online: doi:10.1007/s00431-020-03704-6

Luzi, S. A., Melinek, J., & Oliver, W. R. Medical Examiners' Independence Is Vital for The Health of The American Legal System. (2013) Academic Forensic Pathology, 3, 84-92. Online: doi:10.23907/2013.012

Martire, K. A., Ballantyne, K. N., Bali, A., Edmond, G., Kemp, R. I., & Found, B. Forensic science evidence: Naïve Estimates of False Positive Error Rates and Reliability. (2019) Forensic Science International, 302, 109877. Online: doi:10.1016/j.forsciint.2019.10987 7

McKimmie, B. M., Masters, J. M., Masser, B. M., Schuller, R. A., & Terry, D. J. Stereotypical and Counterstereotypical Defendants: Who Is He and What Was the Case Against Her? (2013) Psychology, Public Policy, and Law, 19, 343–354. Online: doi:10.1037/a0030505

McQuiston-Surrett, D., & Saks, M. J. The Testimony of Forensic Identification Science: What Expert Witnesses Say And What Factfinders Hear. (2009) Law and Human Behavior, 33, 436-453. Online: doi:10.1007/s10979-008-9169-1

Murrie, D. C., Boccaccini, M. T., Guarnera, L. A., & Rufino, K. A. Are Forensic Experts Biased by The Side That Retained Them? (2013) Psychological Science, 24, 1889-1897. Online: doi:10.1177/0956797613481812

National Registry of Exonerations. (2023). *Exonerations in the United States.* Online: https://www.law.umich.edu/special/exoneration/Pages/Exonerations-in-the-United-States-Map.aspx

Nisbett, R. E., & Wilson, T. D. Telling More Than We Can Know: Verbal Reports on Mental Processes. (1977) Psychological Review, 84, 231-259. Online: doi:10.1037/0033-295X.84.3.231

Obenson, K. Commentary on Dror *et al*. 'Cognitive Bias in Forensic Pathology decisions.' (2021) Journal of Forensic Sciences, 66, 2582-2584. Online: doi:10.1111/1556-4029.14855

Oliver, W. R. Manner Determination in Forensic Pathology. (2014) Academic Forensic Pathology, 4, 480-491. Online: doi:10.23907/2014.062

Oliver, W. R. Commentary on Dror *et al*. 'Cognitive Bias in Forensic Pathology Decisions.' (2021) Journal of Forensic Sciences, 66, 2563-2564. Online: doi:10.1111/1556-4029.14841

Ophoven, J. (2022, March 18). *Declaration of Janice Jean Ophoven, M.D. in State of Texas v. Melissa Lucio*. 138th District Court of Cameron County Texas. Online: https://www.documentcloud.org/documents/21657863-dr-janice-ophoven-declaration

Peat, M. A. JFS editor-in-chief preface. (2021) Journal of Forensic Sciences, 66, 2593-2540. Online: doi:10.1111/1556-4029.14844

Peterson, B. L., Arnall, M., Avedschmidt, S., Beers, D., Bell, M., Burton, S. & Wright, R. Commentary on Dror *et al*. 'Cognitive Bias in Forensic Pathology Decisions.' (2021a) Journal of Forensic Sciences, 66, 2541-2544. Online: doi:10.1111/1556-4029.14843

Peterson, B. L., Gill, J., & Oliver, W. Peterson *et al*. Response to Authors' Response. (2021) Journal of Forensic Sciences, 66, 2549-2552. Online: doi:10.1111/1556-4029.14849

Phillips, B., Little, D., McDougall, J., & Langlois, N. E. The Coronial System and Determining Manner of Death in Australia: An Overview. (2015) Academic Forensic Pathology, 5, 436-442. Online: doi:10.23907/2015.047

Pronin, E., & Hazel, L. Humans' Bias Blind Spot and Its Societal Significance. (2023) Current Directions in Psychological Science, 32, 402-409. Online: doi:10.1177/09637214231178745

Quigley-McBride, A., Dror, I. E., Roy, T., Garrett, B. L., & Kukucka, J. A Practical Tool for Information Management in Forensic Decisions: Using Linear Sequential Unmasking-Expanded (LSU-E) in Casework. (2022) Forensic Science International: Synergy, 4, 100216. Online: doi:10.1016/j.fsisyn.2022.100216

Quintana, D. S., & Williams, D. R. Bayesian Alternatives for Common Null-Hypothesis Significance Tests in Psychiatry: A Non-Technical Guide Using JASP. (2018) BMC Psychiatry, 18, 1-8. Online: doi:10.1186/s12888-018-1761-4

Ribeiro, G., Likwornik, H., & Chin, J. M. Visual Decision Aids: Improving Laypeople's Understanding of Forensic Science Evidence. (2023) Journal of Applied Research in Memory and Cognition, 12, 230-240. Online: doi:10.1037/mac0000026

Scanlon, P., Banyai, B., Hart, E., & Cooper, S. L. Juror Certainty About Expert Firearms Identification Evidence and The Impact of Cross-Examination. (2021) Southern California Interdisciplinary Law Journal, 31, 91-105.

Scurich, N., Nguyen, K. D., & John, R. S. Quantifying the Presumption of Innocence. (2016) Law, Probability and Risk, 15, 71-86. Online: doi:10.1093/lpr/mgv016

Scurich, N., & John, R. S. Jurors' presumption of innocence. (2017) The Journal of Legal Studies, 46, 187-206.

Sharma, S. K., & Bajpai, S. Medicolegal Death Investigation in India: An overview. (2015) Academic Forensic Pathology, 5, 443-446. Online: doi:10.23907/2015.048

Simon, D. Minimizing Error and Bias in Death Investigations. (2018) Seton Hall Law Review, 49, 255-305.

Smalarz, L., Madon, S., Yang, Y., Guyll, M., & Buck, S. The Perfect Match: Do Criminal Stereotypes Bias Forensic Evidence Analysis? (2016) Law and Human Behavior, 40, 420-429. Online: doi:10.1037/lhb0000190

Smalarz, L., Madon, S., & Turosak, A. Defendant Stereotypicality Moderates the Effect of Confession Evidence on Judgments of Guilt. (2018) Law and Human Behavior, 42, 355–368. Online: doi:10.1037/lhb0000286

Thompson, W. C., & Scurich, N. How Cross-Examination on Subjectivity And Bias Affects Jurors' Evaluations Of Forensic Science Evidence. (2019) Journal of Forensic Sciences, 64, 1379-1388. Online: doi:10.1111/1556-4029.14031

Weedn, V. W. Bases of Forensic Pathology Expert Testimony With Emphasis on Iowa v Tyler. (2021) Academic Forensic Pathology, 11, 185-195. Online: doi:10.1177/19253621211060961

Whitehead, F. A., Williams, M. R., & Sigman, M. E. Decision Theory and Linear Sequential Unmasking in Forensic Fire Debris Analysis: A Proposed Workflow. (2022) Forensic Chemistry, 29, 100426. Online: doi:10.1016/j.forc.2022.100426