Cognitive Bias in Forensic Mental Health Assessment: Evaluator Beliefs About Its Nature and Scope

Patricia A. Zapf
John Jay College of Criminal Justice

Jeff Kukucka
Towson University

Saul M. Kassin
John Jay College of Criminal Justice

Itiel E. Dror
University College London

Decision-making of mental health professionals is influenced by irrelevant information (e.g., Murrie, Boccaccini, Guarnera, & Rufino, 2013). However, the extent to which mental health evaluators acknowledge the existence of bias, recognize it, and understand the need to guard against it, is unknown. To formally assess beliefs about the scope and nature of cognitive bias, we surveyed 1,099 mental health professionals who conduct forensic evaluations for the courts or other tribunals (and compared these results with a companion survey of 403 forensic examiners, reported in Kukucka, Kassin, Zapf, & Dror, 2017). Most evaluators expressed concern over cognitive bias *but* held an incorrect view that mere willpower can reduce bias. Evidence was also found for a *bias blind spot* (Pronin, Lin, & Ross, 2002), with more evaluators acknowledging bias in their peers' judgments than in their own. Evaluators who had received training about bias were more likely to acknowledge cognitive bias as a cause for concern, whereas evaluators with more experience were less likely to acknowledge cognitive bias as a cause for concern in forensic evaluation as well as in their own judgments. Training efforts should highlight the bias blind spot and the fallibility of introspection or conscious effort as a means of reducing bias. In addition, policies and procedural guidance should be developed in regard to best cognitive practices in forensic evaluations.

Keywords: bias blind spot, cognitive bias, forensic evaluation, forensic mental health assessment, expert decision-making

A long history of research and commentary in psychology has focused on issues related to statistical versus clinical (human judgment) predictions of outcomes (Meehl, 1954; Monahan, 1997). In recent years, research on forensic mental health evaluation has shown wide variability in forensic decision-making among professionals (Mossman, 2013; Murrie, Boccaccini, Johnson, & Janke, 2008; Murrie et al., 2009). For example, Murrie and Warren (2005) examined rates of insanity opinions for 59 evaluators who had completed a total of 4,498 forensic evaluation reports. Although most evaluators determined that 5%–25% of the defendants they evaluated met criteria for legal insanity, some evaluators showed very high (>50%) or low (0%) rates of insanity opinions. Similarly, Murrie, Boccaccini, Zapf, Warren, and Henderson (2008) examined the rates of incompetency opinions across

60 clinicians from two U.S. states who had completed over 7,000 forensic evaluation reports. Although aggregate rates of incompetency across the United States are around 25% (Pirelli, Gottdiener, & Zapf, 2011), this research found wide variability in clinician rates of incompetency opinions, ranging from 0% to 62% (Murrie, Boccaccini, Zapf et al., 2008). This variability in evaluators' rates of opining incompetence or insanity can be attributed to multiple possible causes, including referral stream, case difficulty or ambiguity, and other case-relevant variables. To isolate the potential causes for variability in evaluator opinions better, it is important to examine agreement between evaluators for the same case.

Field reliability studies examining consistency between evaluators for the same defendant have found only moderate levels of agreement. In one such study, Gowensmith Murrie, and Boccaccini (2012) examined 216 competence evaluation cases in Hawaii, where the system requires three independent evaluations for each defendant whose competence to stand trial has been questioned. In 29% of these cases, the three evaluators did not agree on the defendant's competency status. Similarly, an analysis of 165 insanity cases in Hawaii—which likewise require three independent evaluations—found disagreement among evaluators in 45% of cases (Gowensmith, Murrie, & Boccaccini, 2013). These discrepancies in evaluator agreement for the same defendant raise the question of what accounts for this variability in forensic evaluation.

Distinguishing between variability that results from a lack of reliability and variability that is due to biasing effects has been examined and established within the forensic science domains. Dror's (2016) Hierarchy of Expert Performance (HEP) uses eight levels to quantify expert performance by examining and distinguishing between reliability and "biasability," with regard to observations and conclusions, between- and within-experts. Whereas reliability refers to the consistency of expert performance based on relevant information *without bias,* biasability refers to decision-making that is affected by irrelevant contextual information. The overall variability is a function of both reliability and biasability (Dror, 2016; Dror & Rosenthal, 2008). Reliability and biasability can be examined within- and between-experts (inter- and intravariability), and with respect to observations (e.g., measures and other data collected during the evaluation) and the final conclusions that experts reach (e.g., the determination of competency or insanity).

It is important to untangle reliability and biasability, a task that has been achieved in various research studies within the forensic sciences (a full summary is presented within the HEP hierarchy; Dror, 2016). Do the same issues of bias shown within the "hard" forensic sciences, such as DNA analysis (see Dror & Hampikian, 2011) and fingerprint identification (both between- and within-experts and with respect to both observations and conclusions, see Dror & Charlton, 2006; Dror, Charlton, & Péron, 2006; Dror et al., 2011), also apply to forensic mental health evaluation? Forensic evaluations may well be as susceptible, if not moreso, to the biasing effect of exposure to irrelevant contextual information. Forensic psychology has yet to fully examine the issue of biasability, as distinct from reliability, but research has already demonstrated its existence (see Dror & Murrie, 2017 for a review of the relevant forensic assessment research on expert performance regarding reliability and biasability).

Murrie, Boccaccini, Johnson et al. (2008) examined interrater agreement on Psychopathy Checklist–Revised (PCL-R; Hare, 2003) scores in 23 real-world cases where both defense and prosecution experts administered the PCL-R to the same defendant. Surprisingly, prosecution experts assigned higher PCL-R scores than defense experts for the same defendant, and the differences between defense and prosecution PCL-R scores were greater than the reported standard error of measurement for the instrument (Murrie, Boccaccini, Johnson et al., 2008). This demonstrates that variability was not solely attributable to a lack of reliability, but was also a function of a bias in favor of the side for which the experts worked.

Murrie et al. (2009) later found this same pattern of bias using actuarial risk assessment tools—namely, the STATIC-99 (Hanson & Thornton, 1999) and the Minnesota Sex Offender Screening Tool–Revised (Epperson et al., 1998). These tools have demonstrated strong psychometric properties, such as interrater agreement, that should increase reliability and reduce subjectivity in forensic evaluation. However, when scores on these instruments were compared between opposing experts in sexually violent predator cases, prosecution experts assigned higher scores (i.e., scores indicative of higher risk) than defense experts for the same defendant. Along these same lines, Chevalier, Boccaccini, Murrie, and Varela (2015) recently found that defense and prosecution experts also differed in their norm selection and reporting, providing additional evidence of an adversarial bias in score reporting and

biased interpretation practices among forensic mental health professionals.

The biasing effects within forensic psychology have also been demonstrated in an experimental study where 108 forensic evaluators were led to believe that they were retained by either the prosecution or the defense and then asked to consider the same case information to score the PCL-R and the STATIC-99R. Evaluators who believed they were retained by the prosecution assigned higher scores on these instruments (indicative of higher levels of risk) than those who believed they were working for the defense; in essence, demonstrating an allegiance effect (Murrie, Boccaccini, Guarnera, & Rufino, 2013). In sum, mounting evidence suggests that there is at least some risk for bias to impact forensic evaluations, either with respect to the outcome (i.e., ultimate opinion) or with respect to the interpretation of specific instruments or measures used in the evaluation (see Zapf & Dror, 2017 for a discussion of the various ways in which bias can arise and impact decision making in forensic evaluation).

These findings raise question as to whether forensic evaluators understand the risk of bias and take steps to minimize it. To begin addressing this question, Neal and Brodsky (2016) conducted a qualitative study of 20 board-certified forensic psychologists regarding their experiences, awareness, and efforts to correct bias in forensic evaluations. They found that bias awareness falls on a continuum, ranging from complete dismissal of the idea of bias in one's own work to a belief that bias is inevitable. In addition, participants reported more concern about bias in their colleagues' work than in their own, with some reporting that "they take cases that might pose significant challenges for others because they see themselves as able to control their biases when others might be unable to do so" (p. 69). These data are consistent with the phenomenon known as the "bias blind spot," wherein people perceive themselves as less vulnerable to bias than others (Pronin & Kugler, 2007; Pronin et al., 2002). Even when individuals acknowledged the existence of bias in the strategies they used to reach their conclusions, they still believed that they were able to overcome these biases and reach objective conclusions (Hansen, Gerbasi, Todorov, Kruse, & Pronin, 2014).

The qualitative data collected by Neal and Brodsky (2016) also revealed 25 strategies that professionals used to mitigate the impact of bias in their work. The authors then surveyed 351 forensic psychologists to determine the perceived usefulness of each of these 25 strategies. Respondents reported nearly all of the strategies (22 of 25) to be "useful" to "very useful," including those—such as introspection—for which empirical evidence suggests the contrary (Nisbett & Wilson, 1977; Pronin & Kugler, 2007). In addition, other bias reducing strategies—such as "blinding" or exposure control (Gilbert, 1993)—were not mentioned by any of the forensic psychologists. Importantly, Neal and Brodsky did not ask respondents about the risk of bias in their own work or the work of others, and so no data exist regarding the distribution of bias awareness along the continuum described in their qualitative study.

The present study was designed to assess the opinions of an international sample of forensic evaluators on a range of bias-related issues, including the extent to which evaluators are aware of biases in their own work and the degree to which they believe bias impacts the work of their peers. This survey reveals the attitudes and beliefs about bias among forensic mental health

evaluators and provides the necessary, foundational information that will assist in determining whether and what policies might be needed to tackle the issue of cognitive bias. The results of a companion survey of 403 forensic examiners are reported elsewhere (see Kukucka et al., 2017): here we present the survey of forensic evaluators and then compare these results to those obtained from forensic science examiners in the discussion.

Method

This study extends that of Neal and Brodsky (2016) by surveying a large international sample of forensic evaluators to determine the extent to which bias in forensic evaluation is acknowledged in one's own evaluations as well as the evaluations of one's peers. In addition, we were interested in whether experience or training on cognitive biases were related to evaluators' opinions regarding the impact of bias in forensic evaluation.

Participants

Participants were 1,099 mental health professionals who were either currently or previously involved in conducting forensic mental health evaluations for the courts and who represented 39 countries, with the United States being most heavily represented (74.3%), followed by Canada (11.2%), the United Kingdom (2.3%), Australia (1.7%), Sweden (1.5%), and New Zealand (1.1%). All other countries were represented by fewer than 10 (<1%) of respondents (Argentina, Austria, Belgium, Brazil, China, Czech Republic, Denmark, Estonia, Finland, France, Germany, Guatemala, India, Ireland, Italy, Lithuania, Malta, Mauritius, Mexico, Netherlands, Nigeria, Norway, Portugal, Puerto Rico, Russia, Saudi Arabia, Singapore, South Africa, South Korea, Spain, Switzerland, Turkey).

The majority of respondents (86.53%) were actively engaged in conducting forensic mental health evaluations, and held a doctoral-level degree (74.43%), with 13.10% of the sample achieving a Masters-level degree, and 4.09% with a medical degree. Almost one quarter (23.93%) had obtained board certification in the United States through either the American Board of Professional Psychology (ABPP) or a medical certification board, or a certification credentialing board in the respondent's home country of practice. Just over half of respondents were female (52.23%), the average age was 51 years ($Mean = 51.24$; $SD = 13.56$), and the average number of years of experience as a forensic evaluator was 16.87 ($SD = 11.28$).

Participants reported engaging primarily in criminal forensic evaluations (62.6%; e.g., competency, insanity, risk), with 5.9% of respondents primarily conducting civil forensic evaluations (e.g., personal injury, disability), 10.7% primarily conducting family court evaluations (e.g., custody, parental fitness), 4.5% in other types of evaluations (mainly fitness for duty evaluations and neuropsychological evaluations), and 16.2% reporting that they conduct forensic evaluations in multiple domains (e.g., both criminal and civil).

Most respondents worked either independently (31.5%) or for large organizations consisting of more than 21 employees (42.6%), with 13.0% working in a small (2–5 people) practice, 6.6% working in a practice or for an organization of 6–10 people, and 6.4% working in a practice or for an organization with 11–20 people. Overall, 34.9% of respondents reported performing evaluations for both the prosecution and the defense, 11.7% reported working primarily or solely for the prosecution, 16.5% reported working primarily or solely for the defense, and 36.9% reported conducting evaluations primarily for other types of tribunals or purposes (e.g., review boards, treatability evaluations). Respondents ($n = 1,066$) estimated that they had completed an average of 922 forensic evaluations ($Mean = 921.85$; $SD = 2,040.51$; $Mdn = 300$; $Modes = 100$ and $1,000$; Range = 1–30,000; IQR = 70–1,000) and had testified an average of 91.52 times ($SD = 396.34$; $Mdn = 15$; $Mode = 0$; Range = 0–9,000; IQR = 1–56.25).

Procedure

Participants were recruited from the member directories of various professional organizations, including the International Association of Forensic Mental Health Services and Divisions 12 (Society of Clinical Psychology), 41 (American Psychology-Law Society), and 42 (Psychologists in Independent Practice) of the American Psychological Association. Professionals whose member profiles indicated that they practiced or had interest in forensic psychology or forensic mental health were selected for recruitment. A total of 3,301 e-mail addresses were collected, 3,122 of which were valid. A total of three recruitment e-mails were sent to the participant pool on days one, two, and seven before the start of the work day to maximize response rate, as suggested by Trouteaud (2004) and Cook, Heath, and Thompson (2000). E-mails were personalized with the recipient's first name and contained an invitation to participate along with a link to the survey. Responses were received from 1,099 professionals, indicating a 35.2% response rate.

After giving electronic consent, participants provided basic demographic (i.e., age, gender, location, education level) and professional background information (i.e., types of evaluations conducted, board certification, years of experience). Participants were also asked to estimate the accuracy rate of forensic evaluations in general as well as the accuracy rate of their own evaluations. Next, participants were provided with the following brief explanation of the issue of cognitive bias:

> In recent years, there has been some debate over whether forensic evaluators / examiners are subconsciously influenced by prior beliefs and expectations formed on the basis of contextual information (e.g., a detective's opinion, evidence from other forensic domains, a suspect's criminal history, a confession, an eyewitness, information presented by the retaining party) that is irrelevant to the forensic cases/ samples they are evaluating. This phenomenon has been referred to as *cognitive bias*.

On a new page, evaluators were asked a series of 13 questions assessing their attitudes and beliefs about the scope (three items) and nature (10 items) of cognitive bias in forensic mental health evaluation. Participants were also asked whether they had received any training on cognitive bias, and if so, to describe it. Overall, 40.9% of our sample reported having received some form of training on cognitive bias.

All procedures were approved by the Institutional Review Board of John Jay College of Criminal Justice, The City University of New York.

Measures

Estimated accuracy. Evaluators were asked to estimate the overall accuracy rate (0% to 100%) of forensic evaluations in their domain as well as the accuracy rate of their own judgments. Responses indicating a range (e.g., "95–100%") were recoded to the midpoint of the range. Inexact responses (e.g., "more than 80%") were coded conservatively (i.e., 80%), and non-numeric (e.g., "I don't know") responses were excluded. Analyses of these two open-ended items are based on the responses of $n = 905$ and 912 evaluators, respectively.

Scope of bias. Evaluators' opinions regarding the scope of cognitive bias in forensic mental health evaluation were assessed using three parallel items (see Table 1). First, evaluators were asked whether they believed cognitive bias to be a cause for concern in the forensic sciences *as a whole*. Second, they were asked whether they believed cognitive bias to be a cause for concern *in their specific domain* of forensic evaluation. Third, they were asked whether they believed *their own judgments* to be influenced by cognitive bias. Response options for these three items included "Yes," "No," and "I don't know."

Nature of bias. To assess beliefs about the nature of bias in forensic evaluation, respondents were asked to rate how strongly they agreed or disagreed with each of 10 statements (see Table 2), on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*), with a rating of 4 indicating neither agreement nor disagreement.

Results

Estimated Accuracy

Evaluators estimated the accuracy rate of forensic evaluations, in general, to be 73.86% (*Mdn* = 75, Range = 0–100) and the accuracy rate of their own evaluations to be 81.85% (*Mdn* = 85, Range = 0–100). Interestingly, 28 evaluators (2.55% of the total sample) reported their own judgments to be 100% accurate. In addition, 12 evaluators (1.09% of the total sample) reported their own judgments to be 0% accurate. Given that we were unable to determine the reason why these evaluators indicated their own judgments were never accurate, we eliminated these 12 participants from subsequent analyses. After eliminating these 12 participants, the estimated accuracy rate for forensic evaluations, in general, was 74.69% (*Mdn* = 75, Range = 20–100) and the accuracy rate of their own evaluations was 82.94% (*Mdn* = 85, Range = 10–100). Those who responded to both items ($n = 869$) estimated their own accuracy ($M = 82.73$, $SD = 10.94$) to be higher than the overall accuracy of forensic evaluations in general ($M = 74.86$, $SD = 13.10$), $t(868) = -18.16$, $p < .001$, $d = 0.65$, 95% CI [0.56, 0.75].

Scope of Bias

Most evaluators (86%) believed cognitive bias to be a cause for concern in the forensic sciences as a whole, with fewer (79%) believing bias to be a cause for concern in forensic evaluation, and even fewer (52%) believing their own judgments to be influenced by bias (see Table 1). When responses of "I don't know" were excluded, a Cochran's Q test confirmed that evaluators were decreasingly likely to answer "yes" to these three questions, $Q(2) = 182.61$, $p < .001$, $\hat{\eta}_Q^2 = .12$ (Serlin, Carr, & Marascuilo, 1982). Interestingly, only 6.15% of evaluators who believed cognitive bias to be a cause for concern in the forensic sciences as a whole also believed that bias was *not* a concern in forensic evaluation, whereas 15.06% of those who believed cognitive bias to be a concern in forensic evaluation also denied that their own judgments were influenced by bias.

Nature of Bias

Frequency of agreement ratings for each of the 10 statements regarding the nature of cognitive bias are shown in Table 2 and the average agreement rating for each statement as well as the results of a one-sample t test comparing the mean against the scale midpoint (indicative of neither agreement nor disagreement) are shown in Table 3.

Are evaluators vulnerable to bias? Overall, 93.78% of our respondents agreed (either slightly, moderately, or strongly) that evaluators' expectations can influence their *analysis* of a case, whereas only 3.43% disagreed (either slightly, moderately, or strongly) with this statement (Item #1; $d = 1.86$). Similarly, most respondents believed that evaluators' expectations can affect their *ultimate opinion* (Item #2; 91.28% agreed vs. 4.55% disagreed; $d = 1.59$) and that evaluators sometimes know what conclusion they are expected to reach (Item #8; 83.12% agreed vs. 7.61% disagreed; $d = 1.13$). Fewer evaluators, however, agreed that this knowledge affects their conclusions (Item #9, 69.25% agreed vs. 17.89% disagreed; $d = 0.56$).

How should bias be addressed? Most respondents agreed (87.16%) that evaluators who consciously try to set aside their expectations are less likely to be influenced by them (Item #4; $d = 1.31$; 7.26% disagreed). With respect to whether respondents believed that evaluators should be shielded from irrelevant contextual information, no clear consensus was found: 39.91% believed that evaluators should *not* be shielded from irrelevant contextual information; 33.58% that evaluators *should* be shielded from this information, and 26.51% neither agreed nor disagreed (Item #7; $d = -0.08$).

Table 1
Forensic Mental Health Professionals' Beliefs About Scope of Cognitive Bias (%)

Question	Yes	No	Don't know	N
In your opinion, is cognitive bias a cause for concern in forensic evaluations *as a whole*?	85.54	8.29	6.17	1,086
In your opinion, is cognitive bias a cause for concern in *your specific domain* of forensic evaluation?	78.74	14.05	7.21	1,082
In your opinion, are *your own judgments* influenced by cognitive bias?	52.17	24.42	23.40	1,081

Table 2
Frequencies (n, %) of Beliefs About the Nature of Bias

Item	N	Strongly disagree	Disagree	Slightly disagree	Neutral	Slightly agree	Agree	Strongly agree
1. An evaluator's/examiner's prior beliefs and expectations can affect how s/he goes about analyzing a forensic case/sample.	1,078	4 (.37%)	18 (1.67%)	15 (1.39%)	30 (2.78%)	212 (19.67%)	**597 (55.38%)**	202 (18.74%)
2. An evaluator's/examiner's prior beliefs and expectations can affect his/her ultimate opinion about a forensic case/sample.	1,078	6 (.56%)	19 (1.76%)	24 (2.23%)	45 (4.17%)	277 (25.70%)	**533 (49.44%)**	174 (16.14%)
3. An experienced evaluator/examiner is less likely than a new evaluator/examiner to be influenced by prior beliefs/expectations.	1,073	58 (5.41%)	190 (17.71%)	150 (13.98%)	167 (15.56%)	189 (17.61%)	**229 (21.34%)**	90 (8.39%)
4. An evaluator/examiner who makes a conscious effort to set aside his or her prior beliefs and expectations is less likely to be influenced by them.	1,075	10 (.93%)	29 (2.70%)	39 (3.63%)	60 (5.58%)	252 (23.44%)	**495 (46.05%)**	190 (17.67%)
5. Having access to irrelevant contextual information can help forensic evaluators/examiners make more accurate judgments.	1,074	102 (9.50%)	**302 (28.12%)**	118 (10.99%)	299 (27.84%)	104 (9.68%)	119 (11.08%)	30 (2.79%)
6. Having access to irrelevant contextual information makes a forensic evaluator's/examiner's job more interesting.	1,076	68 (6.32%)	201 (18.68%)	58 (5.39%)	**392 (36.43%)**	199 (18.49%)	134 (12.45%)	24 (2.23%)
7. To the extent possible, evaluators/examiners in my domain should be shielded from irrelevant contextual information.	1,075	46 (4.28%)	213 (19.81%)	170 (15.81%)	**285 (26.51%)**	164 (15.26%)	158 (14.70%)	39 (3.63%)
8. Evaluators/Examiners sometimes know what conclusion they are expected to find.	1,078	16 (1.48%)	46 (4.27%)	20 (1.86%)	100 (9.28%)	223 (20.69%)	**491 (45.55%)**	182 (16.89%)
9. When evaluators/examiners know what they are expected to find, it affects the conclusions they reach.	1,073	18 (1.68%)	95 (8.85%)	79 (7.36%)	138 (12.86%)	**394 (36.72%)**	264 (24.60%)	85 (7.92%)
10. Cognitive bias is generally less of a problem in my specialty domain than in other domains of forensic evaluation.	1,076	132 (12.27%)	**338 (31.41%)**	128 (11.90%)	271 (25.19%)	89 (8.27%)	92 (8.55%)	26 (2.42%)

Note. Modal responses are shown in bold.

Comparison of Bias-Trained and -Untrained Evaluators

Evaluators who had never received training on cognitive bias were less likely than those who had to view bias as a cause for concern in the forensic sciences as a whole (82.15% vs. 90.11%), $\chi^2(2, N = 1,078) = 19.82$, $p < .001$, Cramér's $V = .14$, 95% CI [.08, .20], in forensic evaluation (74.84% vs. 84.88%), $\chi^2(2, N = 1,075) = 17.69$, $p < .001$, Cramér's $V = .13$, 95% CI [.08, .19], and in their own judgments (48.57% vs. 57.88%), $\chi^2(2, N = 1,074) = 11.34$, $p = .003$, Cramér's $V = .10$, 95% CI [.04, .16].

Table 4 compares bias-trained and -untrained evaluators with respect to their agreement about the nature of bias. Compared with those who had received training, bias-untrained evaluators less strongly believed that evaluators' prior beliefs and expectations can affect how they analyze a forensic case (Item #1; $d = 0.33$) or their ultimate opinion about a forensic case (Item #2; $d = 0.38$), and were less likely to acknowledge cognitive bias as a problem in forensic evaluation compared with other domains of forensic science (Item #10; $d = 0.22$).

Effects of Experience

A series of multinomial logistic regressions was used to test years of experience as a predictor of beliefs about the scope of bias. Years of experience did not predict one's belief that cognitive bias is a cause for concern in the forensic sciences as a whole, Wald's chi-squares < 0.45, p values $> .50$. However, years of experience did predict whether one saw bias as a concern *in one's own domain* of forensic evaluation, such that more experienced evaluators were less likely to answer "I don't know" as opposed to either Yes, Wald's chi-square(1) = 5.70, $p = .017$, $OR = 1.03$, 95% CI [1.01, 1.05], or No, Wald's chi-square(1) = 10.21, $p = .001$, $OR = 1.04$, 95% CI [1.02, 1.07], and were marginally more likely to answer No as opposed to Yes, Wald's chi-square(1) = 3.70, $p = .054$, $OR = 1.01$, 95% CI [1.00, 1.03].

Years of experience also predicted whether evaluators believed that *their own judgments* were affected by cognitive bias, such that more experienced evaluators were more likely to answer No as opposed to either Yes, Wald's chi-square(1) = 9.81, $p = .002$, $OR = 1.02$, 95% CI [1.01, 1.03], or "I don't know," Wald's

ZAPF, KUKUCKA, KASSIN, AND DROR

Table 3
Means and One-Sample T-Tests for Beliefs About the Nature of Bias

Item	M (SD)	t	p	d [95% CI]
1. An evaluator's/examiner's prior beliefs and expectations can affect how s/he goes about analyzing a forensic case/sample.	5.81 (.97)	61.01	<.001	1.86 [1.76, 1.96]
2. An evaluator's/examiner's prior beliefs and expectations can affect his/her ultimate opinion about a forensic case/sample.	5.66 (1.04)	52.09	<.001	1.59 [1.50, 1.68]
3. An experienced evaluator/examiner is less likely than a new evaluator/examiner to be influenced by prior beliefs/expectations.	4.20 (1.76)	3.70	<.001	.11 [.05, .17]
4. An evaluator/examiner who makes a conscious effort to set aside his or her prior beliefs and expectations is less likely to be influenced by them.	5.57 (1.20)	42.84	<.001	1.31 [1.22, 1.39]
5. Having access to irrelevant contextual information can help forensic evaluators/examiners make more accurate judgments.	3.45 (1.61)	−11.33	<.001	−.35 [−.41, −.28]
6. Having access to irrelevant contextual information makes a forensic evaluator's/examiner's job more interesting.	3.88 (1.50)	−2.55	.011	−.08 [−.14, −.02]
7. To the extent possible, evaluators/examiners in my domain should be shielded from irrelevant contextual information.	3.87 (1.55)	−2.70	.007	−.08 [−.14, −.02]
8. Evaluators/Examiners sometimes know what conclusion they are expected to find.	5.48 (1.30)	37.23	<.001	1.13 [1.06, 1.21]
9. When evaluators/examiners know what they are expected to find, it affects the conclusions they reach.	4.80 (1.42)	18.38	<.001	.56 [.50, .63]
10. Cognitive bias is generally less of a problem in my specialty domain than in other domains of forensic evaluation.	3.21 (1.58)	−16.36	<.001	−.50 [−.56, −.44]

chi-square(1) = 4.80, p = .029, OR = 1.02, 95% CI [1.00, 1.03], but were equally likely to answer "Yes" or "I don't know," Wald's chi-square(1) = 0.27, p = .602, OR = 1.00, 95% CI [0.99, 1.02]. These data indicate that for every additional year of experience, the odds of an evaluator answering "No" as opposed to "Yes" or "I

don't know" increased by 2%; thus, a five-year increase in experience was associated with a 10% increase in the odds that evaluators would see themselves as unaffected by bias.

Years of experience did correlate with agreement ratings such that evaluators with more experience were more likely to agree

Table 4
Effect of Bias Training on Beliefs About the Nature of Bias

Item	Untrained M (SD) (n = 629)	Trained M (SD) (n = 434)	t	p	d	95% CI
1. An evaluator's/examiner's prior beliefs and expectations can affect how s/he goes about analyzing a forensic case/sample.	5.68 (.98)	6.00 (.93)	5.37	<.001	.33	.28, .39
2. An evaluator's/examiner's prior beliefs and expectations can affect his/her ultimate opinion about a forensic case/sample.	5.52 (1.07)	5.84 (.99)	5.09	<.001	.31	.25, .37
3. An experienced evaluator/examiner is less likely than a new evaluator/examiner to be influenced by prior beliefs and expectations.	4.31 (1.68)	4.02 (1.82)	2.69	.007	.17	.06, .27
4. An evaluator/examiner who makes a conscious effort to set aside his or her prior beliefs and expectations is less likely to be influenced by them.	5.57 (1.20)	5.54 (1.23)	.41	.684	.03	−.05, .10
5. Having access to irrelevant contextual information can help forensic evaluators/examiners make more accurate judgments.	3.54 (1.57)	3.33 (1.65)	2.11	.035	.13	.04, .23
6. Having access to irrelevant contextual information makes a forensic evaluator's/examiner's job more interesting.	3.93 (1.48)	3.82 (1.53)	1.20	.231	.07	−.02, .16
7. To the extent possible, evaluators/examiners in my domain should be shielded from irrelevant contextual information.	3.92 (1.53)	3.80 (1.55)	1.32	.188	.08	−.01, .17
8. Evaluators/Examiners sometimes know what conclusion they are expected to find.	5.47 (1.29)	5.48 (1.33)	.08	.938	.01	−.07, .09
9. When evaluators/examiners know what they are expected to find, it affects the conclusions they reach.	4.72 (1.39)	4.90 (1.44)	2.00	.046	.13	.04, .21
10. Cognitive bias is generally less of a problem in my specialty domain than in other domains of forensic evaluation.	3.35 (1.53)	3.01 (1.61)	3.57	<.001	.22	.12, .31

Note. A two-group multivariate analysis of variance (MANOVA) using these ten items as DVs showed a significant multivariate difference, Wilks' Λ = .96, $F(10, 1,052)$ = 5.01, $p < .001$, η_p^2 = .05.

that experienced evaluators are less likely than new evaluators to be influenced by prior beliefs and expectations, $r = .24$, $p < .001$, that evaluators who make a conscious effort to set aside their beliefs and expectations are less likely to be influenced by them, $r = .10$, $p = .001$, and that cognitive bias is less of a problem in forensic evaluation than in other forensic domains, $r = .10$, $p = .001$. In addition, evaluators with more years of experience were less likely to agree that when evaluators know what they are expected to find it affects the conclusions they reach, $r = -.12$, $p < .001$, and less likely to agree that forensic evaluation should be shielded from irrelevant contextual information, $r = -.08$, $p = .008$. No relation was found between bias training and experience ($r_{pb} = .054$, $n = 1064$, $p = .077$).

Discussion

Consistent with recent research demonstrating that forensic evaluators are influenced by irrelevant contextual information (e.g., Murrie et al., 2013), many evaluators acknowledge the impact of cognitive bias on the forensic sciences in general (86%), forensic evaluation specifically (79%), and in their own forensic evaluations (52%). In terms of the pattern of responses, most evaluators recognized bias as a general cause for concern, but far fewer saw themselves as vulnerable. This pattern is consistent with research on the *bias blind spot*—the inherent tendency to recognize biases in others while denying the existence of those same biases in oneself (e.g., Pronin & Kugler, 2007; Pronin et al., 2002). For forensic evaluators, the presence of a bias blind spot might impact the perceived necessity of taking measures to minimize bias in forensic evaluation or the selection of measures to use for this purpose.

Many evaluators showed a limited understanding of how to effectively mitigate bias. Overall, 87% believed that evaluators who consciously try to set aside their preexisting beliefs and expectations are less affected by them. This appears to suggest that many evaluators see bias as an ethical problem that can be overcome by mere willpower (see Dror, Kassin, & Kukucka, 2013 for a discussion of this within the forensic sciences). Decades of research overwhelmingly suggest that cognitive bias operates automatically (Klayman & Ha, 1997) and without awareness (Nisbett & Wilson, 1977), and cannot be eliminated through willpower alone (Wilson & Brekke, 1994). Training efforts to educate evaluators about cognitive bias should underscore the fact that bias is innate and universal, and thus can affect even well intentioned and competent forensic evaluators.

Our sample of forensic evaluators was divided over whether they should be blinded to irrelevant contextual information (34% agree, 40% disagree, 26% neutral). This lack of clear agreement regarding whether blinding or context management is an effective strategy for use in forensic evaluation is not surprising given that Neal and Brodsky's (2016) data reveal that none of the board-certified forensic psychologists surveyed mentioned context management, exposure control, or blinding procedures as a means of reducing the impact of bias. This is a strategy that has been highlighted in forensic science (e.g., Kassin, Dror, & Kukucka, 2013) and may be more difficult to implement in forensic evaluation, where the relevance of various pieces of information can be initially ambiguous. While we know that irrelevant contextual information can bias evaluators (Murrie et al., 2013), we do not yet

have a clear understanding of how best to implement strategies to reduce the impact of irrelevant contextual information in forensic evaluation. Neal and Saks (2016) addressed the issue of context management in forensic evaluation, suggesting that the approach be adapted from the forensic sciences with a need for the development of specific procedures to: limit evaluator's exposure to biasing contextual information prior to the initial evaluation, distinguish domain-specific from domain-irrelevant information, ensure that evaluators facing the same decision task in the same case be exposed to the same information, and to ensure independence of observations. More research on this specific bias-reduction strategy is necessary to guide the development of related policies and procedures.

One general strategy that has been used in both forensic science and forensic evaluation is training on bias to increase understanding and awareness of its potential impact. While we cannot conclude that bias training *produced* the observed differences between bias-trained and—untrained evaluators in terms of attitudes and beliefs about bias, our data demonstrate that evaluators with training in bias hold attitudes and beliefs suggestive of an increased awareness and understanding of the potential impact of bias. While it is encouraging that bias-trained evaluators held more enlightened beliefs, it remains to be seen whether mere knowledge translates into improved performance.

Our data also revealed that more experienced evaluators were less likely to acknowledge cognitive bias as a cause for concern both in forensic evaluation and with respect to their own judgments. Without more information it is difficult to know whether this reflects a generational perspective (e.g., those who have been active in the profession longer hold outdated beliefs) or whether experience is related to reduced vulnerability to bias, or whether some other factor(s) is/are at play. Our data do not indicate a relation between bias training and years of experience so these findings are not a result of more experienced evaluators having lower rates of bias training. Interestingly, some literature on ethical transgressions appears to indicate that these typically occur when clinicians are more than a decade postlicensure, as opposed to newly licensed (see Grenyer & Lewis, 2012), so it is possible that this reduced capacity to see one's self as vulnerable to bias may be related to a more general trend to be somewhat less careful midcareer. More research is necessary to tease apart generational and training variables from experience and other potential factors that could account for this perceived reduction in vulnerability to bias on the part of more experienced evaluators.

Comparisons With Forensic Science

This survey of forensic evaluators was a companion survey to that sent out to forensic examiners within various domains of forensic science (Kukucka et al., 2017). Together, these two surveys allow for direct comparisons between forensic mental heath professionals and forensic scientists regarding beliefs about the accuracy, nature, and scope of cognitive bias.

Estimated accuracy rates. Results of these companion surveys indicate that forensic scientists are more optimistic in their estimates regarding the accuracy of decisions, both within their own domain (94% estimated accuracy in judgments within forensic science v. 75% estimated accuracy in judgments within forensic evaluation) as well as with respect to their own judgments

(96% estimated accuracy of own judgments by forensic examiners v. 83% estimated accuracy by forensic evaluators). In both samples, however, we see a similar pattern that supports the presence of a bias blind spot, with more professionals within each domain willing to acknowledge bias in the judgments of others more readily than bias within their own judgments.

Scope of bias.   Regarding the scope of bias, the same pattern of results was found across both forensic scientists and forensic evaluators, with forensic evaluators being more willing to acknowledge cognitive bias as a cause for concern in forensic sciences as a whole (86% of evaluators v. 71% of examiners), within their own domain (79% of evaluators v. 52% of examiners), and in their own judgments (52% of evaluators v. 26% of examiners). With respect to the forensic evaluator data, it is interesting to note that although the "Yes" responses about concern dropped from 79% (for forensic evaluation) to 52% (for their own judgments), the "I don't know" responses increased from 7% (for forensic evaluation) to 23% (for their own judgments). Thus, an alternative interpretation is that "I don't know" admits the possibility of being vulnerable to bias so collapsing "Yes" and "I don't know" responses shows that 85.95% of respondents endorse vulnerability to bias in forensic evaluation and 75.57% endorse vulnerability in their own judgments. In either case, evaluators appear more willing to endorse vulnerability to bias in the forensic sciences as a whole and in forensic evaluation in particular than in their own judgments, a pattern that was similar across forensic examiners.

Nature of bias.   Vulnerability to bias (Items 1, 2, 8, 9) showed the same pattern of results across both forensic examiners and forensic evaluators, albeit with forensic evaluators more willing to acknowledge vulnerability to bias as compared with forensic examiners. How bias should be addressed (Items 4 and 7) also demonstrated a similar pattern across forensic evaluators and examiners with both groups believing that those who try to set aside their preexisting beliefs and expectations are less likely to be influenced by them, although more evaluators (87%) believed this to be true than examiners (69%). A similar pattern of results was also found for whether examiners/evaluators should be shielded from irrelevant contextual information, although with more examiners (49%) than evaluators (34%) agreed with this bias-reducing strategy.

In terms of bias reducing strategies, more support was found for those strategies promoted within and relevant to the specific domain. That is, evaluators show less support for context management strategies—which have been promoted in the forensic sciences and that might be more difficult to implement within forensic evaluation—than forensic evaluators.

Training.   With respect to training on cognitive bias, more forensic scientists/examiners (58%) reported receiving training on cognitive bias than forensic evaluators (41%); however, the same pattern of results was found between examiners and evaluators with respect to the impact of training such that those who had received training on cognitive bias were more likely to acknowledge it as a cause for concern in the forensic sciences as a whole, in their specific domain, and in their own judgments.

Experience.   For forensic examiners, experience was not related to beliefs about the nature and scope of cognitive bias whereas, for forensic evaluators, years of experience as a forensic evaluator appears to be related to beliefs about the increased vulnerability of new or inexperienced evaluators, as well as a general pattern of being less willing to acknowledge bias in forensic evaluation and more willing to believe that evaluators are able to control the impact of bias by making a conscious effort. Given the current survey format, we were unable to tease apart the effects of generational beliefs and/or training effects on beliefs about the impact of cognitive bias in forensic evaluation. More research is needed to tease apart the effects of professional experience on beliefs about, and performance regarding, the impact of cognitive bias on forensic decision-making.

Limitations

A few limitations of this research are worth noting. We utilized a survey methodology that relied on self-report so we were unable to ascertain the validity of the responses or obtain more detailed information to elucidate the reasoning behind respondents' answers to the questions. Related to this, we were unable to ensure that all respondents would interpret the questions in the same way. For example, one reviewer pointed out that with respect to question four about the nature of bias (i.e., An evaluator who makes a conscious effort to set aside his or her prior beliefs and expectations is less likely to be influenced by them), respondents could indicate this to be true but still not believe that this conscious effort would eliminate bias, only that it would result in a reduction of the potential influence. Another pointed out that ambiguity regarding the word "irrelevant" and what that might mean in relation to a particular case could have led to different interpretations by various respondents. In addition, our methodology did not allow us to examine casual influences or anything more than mere associations between variables such as training or experience and beliefs about bias.

Conclusions

Research demonstrating bias in the forensic sciences has had a significant impact in terms of policy implementation and procedures to attempt to minimize the impact of bias. For example, the National Commission on Forensic Science, the National Institute of Standards and Technology, and the Department of Justice have each issued procedures and policies to ensure that forensic scientists be exposed only to relevant information. In the United States, the President's Council of Advisors on Science and Technology (PCAST) *Report on Forensic Science in Criminal Courts* and, in the United Kingdom, the Forensic Science Regulator's *Guidance on Cognitive Bias Effects Relevant to Forensic Science Examinations* each issued policies and procedures for attempting to minimize the impact of cognitive bias in forensic science investigations. Similar changes are happening in forensic science laboratories throughout the United States, where issues of cognitive bias are being acknowledged and policies/procedures are being implemented. Forensic psychology has yet to really take on and acknowledge the impact of irrelevant contextual information on bias or in terms of coming out with policies and procedures to deal with this issue. As reflected in our data, the policies for forensic science do not translate easily to forensic psychology, and one needs to develop specific policies that fit the characteristics of forensic psychology and forensic mental health evaluation. Guidance regarding best cognitive practice in forensic psychology is necessary, and this study has provided a few areas for further investigation.

Cognitive bias is an issue relevant to all domains of forensic science, including forensic evaluation. Our results reveal that cognitive bias appears to be a cause for concern in forensic evaluation. Training models emphasize the necessity and importance of context, and evaluators are trained to consider the impact of many different aspects of context on the particular issue being evaluated. This reliance on context in forensic evaluation might result in forensic evaluators being more willing to acknowledge the potential biasing impact of context, but at the same time, being also more susceptible to bias. What appears clear is that not all evaluators are receiving training on biases that can result from human factors or contextual factors in forensic evaluation. In this sample, only 41% had received training on bias in forensic evaluation, suggesting the need for a systematic means of ensuring that all forensic evaluators receive training on this issue. Implementing policy or procedure at the state licensing level or in other credentialing or certification processes is one means of ensuring that all forensic evaluators receive training on this important issue. As Guarnera, Murrie, and Boccaccini (2017) recommended, "states without standards for the training and certification of forensic experts should adopt them, and states with weak standards (e.g., mere workshop attendance) should strengthen them" (p. 149).

Evidence for a bias blind spot in forensic evaluators was found. Future research is needed to investigate ways in which this bias blind spot might be reduced or minimized. Neal and Brodsky's (2016) survey of forensic psychologists revealed that *all* evaluators endorsed introspection as an effective means of reducing bias, despite research evidence to the contrary. Pronin and Kugler (2007) found that educating individuals about the fallibility of introspection resulted in a reduced reliance on introspection as a means of minimizing bias. Training on bias should explicitly address the bias blind spot and the fallibility of introspection as a bias-reducing strategy.

More research on specific mechanisms to reduce or minimize the effects of cognitive bias in forensic evaluation is required. Techniques such as exposure control, emphasized in the forensic sciences, may be feasible for some aspects of forensic evaluation but not others; however, more research is needed to determine the specific conditions under which these strategies can be effective in forensic evaluation. The use of checklists, alternate hypothesis testing, considering the opposite, and other strategies have been proposed for use in forensic evaluation to reduce the impact of bias (e.g., Borum, Otto, & Golding, 1993; Robb, 2006; Wills, 2008), but more research is needed to determine the specific conditions under which these strategies can be most effective. Cross-domain research, drawing on bias reduction strategies used in the forensic and clinical/medical sciences and their application to forensic evaluation, is necessary to develop the ways in which bias in forensic evaluation can be reduced. As Lockhart and Satya-Murti (2017) recently concluded, "it is time to shift focus to the study of errors within specific domains, and how best to communicate uncertainty in order to improve decision-making on the part of both the expert and the trier-of-fact" (p.1).

What is clear is that forensic evaluators appear to be aware of the issue of bias in general, but diminishing rates of perceived susceptibility to bias in one's own judgments and the perception of higher rates of bias in the judgments of others as compared with oneself, underscore that we may not be the most objective evaluators of our own decisions. As with the forensic sciences, implementing procedures and strategies to minimize the impact of bias in forensic evaluation can serve to proactively mitigate against the intrusion of irrelevant information in forensic decision making. This is especially important given the courts' heavy reliance on evaluators' opinions (see Zapf, Hubbard, Cooper, Wheeles, & Ronan, 2004), the fact that judges and juries have little choice but to trust the expert's self-assessment of bias (see Kassin et al., 2013), and the potential for biased opinions and conclusions to cross-contaminate other evidence or testimony (see Dror, Morgan, Rando, & Nakhaeizadeh, 2017). More research is necessary to determine the specific strategies to be used and the various recommended means of implementing those strategies across forensic evaluations, but the time appears to be ripe for further discussion and development of policies and guidelines to acknowledge and attempt to reduce the potential impact of bias in forensic evaluation.

References

Borum, R., Otto, R., & Golding, S. (1993). Improved clinical judgment and decision making in forensic evaluation. *Journal of Psychiatry & Law, 21,* 35–76. http://dx.doi.org/10.1177/009318539302100104

Chevalier, C. S., Boccaccini, M. T., Murrie, D. C., & Varela, J. G. (2015). Static-99R reporting practices in sexually violent predator cases: Does norm selection reflect adversarial allegiance? *Law and Human Behavior, 39,* 209–218. http://dx.doi.org/10.1037/lhb0000114

Cook, C., Heath, F., & Thompson, R. L. (2000). A meta-analysis of response rates in web- or internet-based surveys. *Educational and Psychological Measurement, 60,* 821–836. http://dx.doi.org/10.1177/00131640021970934

Dror, I. E. (2016). A hierarchy of expert performance. *Journal of Applied Research in Memory & Cognition, 5,* 121–127. http://dx.doi.org/10.1016/j.jarmac.2016.03.001

Dror, I. E., Champod, C., Langenburg, G., Charlton, D., Hunt, H., & Rosenthal, R. (2011). Cognitive issues in fingerprint analysis: Inter- and intra-expert consistency and the effect of a "target" comparison. *Forensic Science International, 208,* 10–17. http://dx.doi.org/10.1016/j.forsciint.2010.10.013

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

Dror, I. E., Charlton, D., & Péron, A. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International, 156,* 74–78. http://dx.doi.org/10.1016/j.forsciint.2005.10.017

Dror, I. E., & Hampikian, G. (2011). Subjectivity and bias in forensic DNA mixture interpretation. *Science & Justice, 51,* 204–208. http://dx.doi.org/10.1016/j.scijus.2011.08.004

Dror, I. E., Kassin, S. M., & Kukucka, J. (2013). New application of psychology to law: Improving forensic evidence and expert witness contributions. *Journal of Applied Research in Memory & Cognition, 2,* 78–81. http://dx.doi.org/10.1016/j.jarmac.2013.02.003

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

Dror, I. E., & Murrie, D. C. (2017). A hierarchy of expert performance applied to forensic psychological assessments. *Psychology, Public Policy, and Law*. Advance online publication. http://dx.doi.org/10.1037/law0000140

Dror, I., & Rosenthal, R. (2008). Meta-analytically quantifying the reliability and biasability of forensic experts. *Journal of Forensic Sciences, 53,* 900–903. http://dx.doi.org/10.1111/j.1556-4029.2008.00762.x

Epperson, D. L., Kaul, J. D., Goldman, R., Hout, S., Hesselton, D., & Alexander, W. (1998). *Minnesota Sex Offender Screening Tool–Revised*. St. Paul: Minnesota Department of Corrections. Retrieved from http://www.psychology.iastate.edu

Gilbert, D. T. (1993). The assent of man: Mental representation and the control of belief. In D. M. Wegner & J. W. Pennebaker (Eds.), *Handbook of mental control* (pp. 57–87). Englewood Cliffs, NJ: Prentice Hall.

Gowensmith, W. N., Murrie, D. C., & Boccaccini, M. T. (2012). Field reliability of competence to stand trial opinions: How often do evaluators agree, and what do judges decide when evaluators disagree? *Law and Human Behavior, 36,* 130–139. http://dx.doi.org/10.1037/h0093958

Gowensmith, W. N., Murrie, D. C., & Boccaccini, M. T. (2013). How reliable are forensic evaluations of legal sanity? *Law and Human Behavior, 37,* 98–106. http://dx.doi.org/10.1037/lhb0000001

Grenyer, B. F. S., & Lewis, K. L. (2012). Prevalence, prediction, and prevention of psychologist misconduct. *Australian Psychologist, 47,* 68–76. http://dx.doi.org/10.1111/j.1742-9544.2010.00019.x

Guarnera, L. A., Murrie, D. C., & Boccaccini, M. T. (2017). Why do forensic experts disagree? Sources of unreliability and bias in forensic psychology evaluations. *Translational Issues in Psychological Science, 3,* 143–152. http://dx.doi.org/10.1037/tps0000114

Hansen, K., Gerbasi, M., Todorov, A., Kruse, E., & Pronin, E. (2014). People claim objectivity after knowingly using biased strategies. *Personality and Social Psychology Bulletin, 40,* 691–699. http://dx.doi.org/10.1177/0146167214523476

Hanson, R. K., & Thornton, D. (1999). *Static-99: Improving actuarial risk assessments for sex offenders* (User Report 99–02). Ottawa, Ontario, Canada: Department of the Solicitor General of Canada.

Hare, R. D. (2003). *The Hare Psychopathy Checklist–Revised* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory & Cognition, 2,* 42–52. http://dx.doi.org/10.1016/j.jarmac.2013.01.001

Klayman, J., & Ha, Y. W. (1997). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Bulletin, 94,* 211–228.

Kukucka, J., Kassin, S. M., Zapf, P. A., & Dror, I. E. (2017). Cognitive bias and blindness: A global survey of forensic science examiners. *Journal of Applied Research in Memory & Cognition.* Advance online publication. http://dx.doi.org/10.1016/j.jarmac.2017.09.001

Lockhart, J. J., & Satya-Murti, S. (2017). Diagnosing crime and diagnosing disease: Bias reduction strategies in the forensic and clinical sciences. *Journal of Forensic Sciences.* Advance online publication. http://dx.doi.org/10.1111/1556-4029.13453

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and review of the evidence*. Minneapolis, Minnesota: University of Minnesota Press. http://dx.doi.org/10.1037/11281-000

Monahan, J. (1997). Clinical and actuarial predictions of violence. In D. Faigman, D. Kaye, M. Saks, & J. Sanders (Eds.), *Modern scientific evidence: The law and science of expert testimony* (pp. 300–318). St. Paul, MN: West.

Mossman, D. (2013). When forensic examiners disagree: Bias, or just inaccuracy? *Psychology, Public Policy, and Law, 19,* 40–55. http://dx.doi.org/10.1037/a0029242

Murrie, D. C., Boccaccini, M. T., Guarnera, L. A., & Rufino, K. A. (2013). Are forensic experts biased by the side that retained them? *Psychological Science, 24,* 1889–1897. http://dx.doi.org/10.1177/0956797613481812

Murrie, D. C., Boccaccini, M. T., Johnson, J. T., & Janke, C. (2008). Does interrater (dis)agreement on Psychopathy Checklist scores in sexually violent predator trials suggest partisan allegiance in forensic evaluations? *Law and Human Behavior, 32,* 352–362. http://dx.doi.org/10.1007/s10979-007-9097-5

Murrie, D. C., Boccaccini, M. T., Turner, D., Meeks, M., Woods, C., & Tussey, C. (2009). Rater (dis)agreement on risk assessment measures in sexually violent predator proceedings: Evidence of adversarial allegiance in forensic evaluation? *Psychology, Public Policy, and Law, 15,* 19–53. http://dx.doi.org/10.1037/a0014897

Murrie, D. C., Boccaccini, M. T., Zapf, P. A., Warren, J. I., & Henderson, C. E. (2008). Clinician variation in findings of competence to stand trial. *Psychology, Public Policy, and Law, 14,* 177–193. http://dx.doi.org/10.1037/a0013578

Murrie, D. C., & Warren, J. I. (2005). Clinician variation in rates of legal sanity opinions: Implications for self-monitoring. *Professional Psychology: Research and Practice, 36,* 519–524. http://dx.doi.org/10.1037/0735-7028.36.5.519

Neal, T. M. S., & Brodsky, S. L. (2016). Forensic psychologists' perceptions of bias and potential correction strategies in forensic mental health evaluations. *Psychology, Public Policy, and Law, 22,* 58–76. http://dx.doi.org/10.1037/law0000077

Neal, T. M. S., & Saks, M. J. (2016, March). *The science of science offers solutions to cognitive bias in forensic psychology*. Paper presented at the annual meeting of the American Psychology-Law Society, Atlanta, GA.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259. http://dx.doi.org/10.1037/0033-295X.84.3.231

Pirelli, G., Gottdiener, W. H., & Zapf, P. A. (2011). A meta-analytic review of competency to stand trial research. *Psychology, Public Policy, and Law, 17,* 1–53. http://dx.doi.org/10.1037/a0021713

Pronin, E., & Kugler, M. B. (2007). Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology, 43,* 565–578. http://dx.doi.org/10.1016/j.jesp.2006.05.011

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin, 28,* 369–381. http://dx.doi.org/10.1177/0146167202286008

Robb, A. (2006). Strategies to address clinical bias in the child custody evaluation process. *Journal of Child Custody: Research, Issues, and Practices, 3,* 45–69. http://dx.doi.org/10.1300/J190v03n02_03

Serlin, R. C., Carr, J., & Marascuilo, L. A. (1982). A measure of association for selected nonparametric procedures. *Psychological Bulletin, 92,* 786–790. http://dx.doi.org/10.1037/0033-2909.92.3.786

Trouteaud, A. R. (2004). How you ask counts: A test of internet-related components of response rates to a web-based survey. *Social Science Computer Review, 22,* 385–392. http://dx.doi.org/10.1177/0894439304265650

Wills, C. D. (2008). The CHESS method of forensic opinion formulation: Striving to checkmate bias. *Journal of the American Academy of Psychiatry and the Law, 36,* 535–540.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116,* 117–142. http://dx.doi.org/10.1037/0033-2909.116.1.117

Zapf, P. A., & Dror, I. E. (2017). Understanding and mitigating bias in forensic evaluation: Lessons from forensic science. *International Journal of Forensic Mental Health, 16,* 227–238. http://dx.doi.org/10.1080/14999013.2017.1317302

Zapf, P. A., Hubbard, K. L., Cooper, V. G., Wheeles, M. C., & Ronan, K. A. (2004). Have the courts abdicated their responsibility for determination of competency to stand trial to clinicians? *Journal of Forensic Psychology Practice, 4,* 27–44. http://dx.doi.org/10.1300/J158v04n01_02