Journal of Applied Research in Memory and Cognition 6 (2017) 193–202

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Cognitive Bias in the Legal System: Police Officers Evaluate Ambiguous Evidence in a Belief-Consistent Manner

Steve D. Charman[a,*], Melissa Kavetski[a], Dana Hirn Mueller[b]

[a] Florida International University, USA
[b] Medical College of Wisconsin, USA

Students' and forensic examiners' beliefs in a suspect's guilt can bias their evaluations of subsequent evidence. The current study examines whether experienced police officers also exhibit similar effects. Police officers ($n = 89$) and undergraduate students ($n = 227$) read a fictional criminal case and received incriminating, exonerating, or neutral initial evidence concerning a suspect before providing their initial beliefs in that suspect's guilt. Participants then evaluated the incriminating/exonerating value of four pieces of ambiguous evidence (an alibi, a facial composite, a handwriting sample, informant testimony), and subsequently provided their final beliefs in the suspect's guilt. Structural equation modeling indicated that (a) police officers' initial beliefs of guilt significantly predicted their evaluations of three types of ambiguous evidence, (b) these biased evaluations significantly predicted the officers' final beliefs of guilt, demonstrating a bias snowball effect, and (c) the pattern of effects were the same for police officers as for students.

General Audience Summary

The current study examines whether police officers' evaluations of different pieces of evidence are related to their beliefs in the suspect's guilt. In other words, will police officers who believe a suspect is guilty evaluate evidence more harshly than police officers who believe the same suspect is innocent? And will these biased evaluations further inflate their beliefs in the suspect's guilt, resulting in a bias snowball effect? Police officers ($n = 89$) read about a crime that contained either incriminating, exonerating, or neutral evidence against a suspect, provided their beliefs that the suspect was guilty, and then evaluated four pieces of evidence: the suspect's alibi, a comparison of handwriting samples, a comparison between a facial composite and the suspect, and an informant's testimony. Results indicated that the more likely the police officers believed the suspect to be guilty, the more harshly they tended to evaluate the evidence, which then further inflated their beliefs of guilt. These results suggest that one reason innocent people get convicted is because police officers form a belief in the innocent person's guilt, leading them to interpret further ambiguous evidence as being more incriminating than it actually is, further inflating their belief in the innocent person's guilt.

Keywords: Cognitive bias, Decision-making, Confirmation bias, Context effects

Traditionally, legal psychologists have approached the question of why innocent people become wrongfully convicted by examining the unique, independent contribution of a given piece of evidence (e.g., mistaken eyewitness identification, false confessions, poor forensic judgments). Recently, however, a more sophisticated approach has begun to appreciate the interdependencies between pieces of evidence, examining how one piece of evidence may influence the evaluation of a subsequent piece of evidence via the evaluator's beliefs about the guilt of the suspect. For example, forensic experts who believe a suspect is

Author Note

* Correspondence concerning this article should be addressed to Steve D. Charman, Department of Psychology, DM 256, Florida International University, Miami, FL 33199, USA. Contact: charmans@fiu.edu

guilty (based on their knowledge of other evidence) can come to interpret subsequent ambiguous evidence as being more incriminating than they otherwise would have. This process can then result in a self-sustaining positive feedback loop: Once evaluators obtain evidence pointing toward a suspect's guilt, they may then interpret subsequent ambiguous evidence as being particularly incriminating, further bolstering their belief in the suspect's guilt, further biasing the evaluation of additional evidence. In this fashion, bias can compound on itself, resulting in what Dror (2012) calls "a bias snowball effect."

Research on this type of contextual bias among legal decision-makers has recently come to the forefront of legal psychologists' awareness, partly as a result of a National Academy of Sciences (2009) publication that excoriated many forensic science disciplines for their unscientific basis and susceptibility to contextual bias, as well as real-world cases in which contextual biases have been officially implicated (e.g., the Brandon Mayfield "Madrid bomber" case; U.S. Department of Justice, 2006). Consequently, researchers have provided data showing that biasing effects exist when making decisions regarding the incriminating or exonerating value of evidence (see Kassin, Dror, & Kukucka, 2013, for an overview). For instance, beliefs in a suspect's guilt have been shown to bias fingerprint examiners' determinations about whether two fingerprints match (Dror, Charlton, & Péron, 2006), similarity judgments between a composite of a perpetrator and a suspect (Charman, Gregory, & Carlucci, 2009), and eyewitness identification decisions (such that knowledge of a lineup member's confession can lead witnesses to identify the confessor; Hasel & Kassin, 2009). Even ostensibly more objective types of evidence, such as DNA—thought to be a gold standard of forensics—can be biased by initial beliefs (Dror & Hampikian, 2011). Furthermore, a real-world analysis of wrongful convictions suggests that false confessions may taint forensic examiners' evaluations of forensic evidence (Kassin, Bogart, & Kerner, 2012).

Researchers have developed numerous recommendations to minimize these effects (see Kassin et al., 2013). For example, keeping forensic experts (e.g., fingerprint examiners) blind to the existence of other evidence against a suspect should obviate the formation of beliefs regarding the suspect's guilt, reducing any potential bias on subsequent evidence evaluation. Similarly, Dror et al. (2015) recommend a sequential unveiling of information, keeping an evaluator unaware of extraneous case information for as long as possible. However, although these suggestions should reduce contextual bias among forensic scientists, it is impossible to keep certain legal decision-makers blind to this information. Police officers and detectives, for instance, must routinely make judgments about the likely guilt of a suspect based on various pieces of accumulating ambiguous and imperfect evidence, and are unable to do so in a contextual vacuum. Unfortunately, virtually all research on contextual biases within a forensic context has focused on either student-participants, forensic examiners, or police trainees (e.g., Ask, Rebelius, & Granhag, 2008); little research has focused specifically on experienced police officer decision-making (for exceptions using Swedish criminal investigators, see Ask & Granhag, 2007; Ask, Granhag, & Rebelius, 2011).

Focusing on the decision-making of police officers expands the scope of judgments that may be biased by beliefs of guilt. Whereas forensic examiners' judgments are limited to forensic testing (e.g., fingerprints, DNA, firearms, hair analysis), police officers' judgments are much broader. For instance, throughout the investigative process, police officers may evaluate a suspect's alibi, an eyewitness's identification (or non-identification), an informant's claims, the similarity between a suspect and a composite of the perpetrator, and many other pieces of evidence, the evaluation of which will likely determine the amount of resources spent pursuing a given suspect. And notably, unlike forensic examiners who may produce official reports with their conclusions, police officers' judgments often leave no trail as to their occurrence, making their biased evaluations particularly difficult to diagnose after the fact. Consequently, it is crucial to determine whether, and to what extent, contextual biases influence decision-making among police officers.

These types of context effects are consistent with a variety of theoretical approaches to decision-making, such as confirmation bias (Kassin et al., 2013; Nickerson, 1998), asymmetric skepticism (Lord, Ross, & Lepper, 1979; Marksteiner, Ask, Reinhard, & Granhag, 2011), and cognitive coherence (Holyoak & Simon, 1999). Although these approaches are all related, we adopt a cognitive coherence theoretical perspective for the purposes of the current manuscript, which provides a particularly flexible framework in which to analyze situations in which people must evaluate multiple pieces of evidence to form a conclusion. Specifically, a defining feature of cognitive coherence models is that decision-making occurs in a bi-directional fashion: Not only does the evaluation of evidence affect an evaluator's emerging conclusion, but that emerging conclusion feeds back to influence the evaluation of evidence (e.g., Simon, Pham, Le, & Holyoak, 2001; Simon, Snow, & Read, 2004). Thus, the evaluation of various pieces of evidence tends to cohere with the emerging conclusion (and with each other).

Cognitive coherence models thus emphasize the dynamically evolving nature of decision-making. But whereas the extant research on context effects on legal decision-making has shown that beliefs of guilt can affect the evaluation of evidence—at least among student-participants and forensic examiners—we are aware of no research that has examined the next step: whether these context effects can result in a bias snowball effect whereby biased evaluations of evidence further predict the extent to which evaluators update their beliefs in a suspect's guilt. Evidence for such a process would highlight a difficulty innocent suspects face in trying to escape suspicion: Once they are believed to be guilty, further ambiguous evidence will tend to be evaluated as overly incriminating, further increasing belief in the suspect's guilt. Thus, the current study has two main purposes: To examine (a) whether police officers exhibit contextual biases in their evaluations of various pieces of forensically relevant evidence; and (b) whether any observed biases in evidence evaluation further influence police officers' updated beliefs in the suspect's guilt. Furthermore, to provide a sense of whether the findings of studies using student samples can

Table 1
Mean Initial Belief of Guilt Ratings, Evidence Ratings, and Final Belief of Guilt Ratings (SDs in Parentheses) Among Police Officers as a Function of Type of Evidence and Valence of Evidence

Valence of initial evidence	Initial belief in suspect's guilt	Evidence evaluation				Final belief in suspect's guilt
		Alibi strength	Composite similarity	Handwriting similarity	Informant implication	
DNA evidence						
Exonerating ($n = 12$)	12.2 (21.9)	4.3 (.9)	64.6 (24.2)	43.5 (35.5)	34.7 (26.8)	39.0 (27.2)
Neutral ($n = 15$)	31.7 (23.2)	3.5 (1.1)	59.3 (27.4)	47.3 (29.3)	47.5 (21.4)	53.7 (23.8)
Incriminating ($n = 18$)	79.4 (26.8)	3.6 (1.1)	74.3 (14.9)	58.7 (33.1)	58.2 (20.2)	78.4 (14.9)
Eyewitness evidence						
Exonerating ($n = 16$)	31.6 (21.4)	3.4 (1.0)	67.7 (26.2)	65.4 (26.3)	46.7 (25.6)	55.2 (25.3)
Neutral ($n = 12$)	29.9 (19.5)	3.6 (1.0)	58.6 (27.2)	48.1 (31.5)	50.5 (28.2)	55.3 (21.8)
Incriminating ($n = 16$)	52.3 (24.0)	3.6 (1.4)	61.5 (30.3)	63.2 (27.6)	49.9 (26.9)	62.0 (23.1)
Total						
Exonerating ($n = 28$)	23.3 (23.4)	3.8 (1.0)	66.3 (24.9)	56.3 (31.7)	41.7 (26.3)	48.3 (26.8)
Neutral ($n = 27$)	30.9 (21.3)	3.5 (1.0)	59.0 (26.8)	47.7 (29.7)	48.8 (24.0)	54.4 (22.5)
Incriminating ($n = 34$)	66.6 (28.6)	3.6 (1.2)	68.3 (24.0)	60.8 (30.2)	54.3 (23.6)	70.7 (20.7)

Note: Alibi strength is measured on a 1–7 scale; all other variables are measured on a 0–100 scale.

generalize to police officers, we also included a comparison student sample.

We hypothesized that (a) police officers' initial beliefs in a suspect's guilt would predict their beliefs about the incriminating/exonerating value of subsequent ambiguous evidence; (b) that this would result in a bias snowball effect such that the biased evaluations of evidence would predict the extent to which police officers update their beliefs in the suspect's guilt; and (c) that these context effects would exist among police officers as well as among students.

Method

Participants

Police and sheriff's departments across the US were contacted via email and asked to participate in this online study.[1] Each department received two additional reminders to participate. Data collection was stopped when no new data were collected for one month. Of those contacted, 89 law enforcement officers from 24 states participated (21% female; mean age = 45 years, SD = 10; 76% White, 4% Hispanic, 1% Black, 1% Asian, 1% Native American, 16% no reply). These officers had an average of 20 years ($SD = 10$) experience, and held various ranks (29% chief of police, 24% detective/investigator, 13% patrol officer, 9% lieutenant, 9% police officer, 6% sergeant, 5% captain, 5% other). See Table 1 for distribution of police officers to condition.

As a comparison sample, 227 students from a large Southeastern university also participated (82% female; mean age = 22 years, $SD = 5$; 70% Hispanic, 8% White, 8% Black, 4% Asian, 2% Haitian, 8% other). Student-participants were recruited via announcements made in large undergraduate psychology classes and were given a deadline by which they needed to have participated. All data collected during that time period were used in analyses. See Table 2 for distribution of students to condition. Both samples engaged in the same procedures.

Materials and Procedure

All data were collected via an online Qualtrics survey. Participants first provided basic demographic information regarding age, gender, and ethnicity; law enforcement participants also provided information regarding rank, years of experience, and location by state. Participants were then given a short synopsis of a crime and subsequent apprehension of a suspect, which they were led to believe was taken from an actual case (see Appendix A). To provide variability in participants' initial beliefs in the suspect's guilt, each participant was randomly assigned to one of six fully crossed conditions, in which they read about (a) DNA or eyewitness evidence, which was either (b) incriminating (a DNA match or an eyewitness identification of the suspect), exonerating (a DNA non-match or eyewitness non-identification of the suspect), or neutral (inconclusive DNA results or a witness who stated that he did not get a good view of the crime and refused to view a lineup). After reading about this evidence, participants were asked to rate the likelihood that the suspect was guilty on a Likert-type scale from 0 (*extremely unlikely*) to 100 (*extremely likely*). This measure served as the initial likelihood of guilt rating that was used to predict participants' subsequent ratings of the ambiguous evidence.

Next, participants were provided with the four pieces of forensically relevant evidence (each created to be ambiguous as to its incriminating/exonerating value) in a randomized order. They rated each piece of evidence before moving on to the next piece.

[1] Response rate is impossible to calculate, since recruitment depended on the person who was contacted—generally the police chief—passing along the invitation to participate to the various officers in that department. It is unknown how often this information was passed along, and thus unknown how many actual detectives/police officers received the invitation.

Table 2
Mean Initial Belief of Guilt Ratings, Evidence Ratings, and Final Belief of Guilt Ratings (SDs in Parentheses) Among Students as a Function of Type of Evidence and Valence of Evidence

Valence of initial evidence	Initial belief in suspect's guilt	Evidence evaluation				Final belief in suspect's guilt
		Alibi strength	Composite similarity	Handwriting similarity	Informant implication	
DNA evidence						
Exonerating ($n = 36$)	19.3 (24.4)	3.5 (1.3)	67.2 (20.2)	58.3 (30.8)	41.9 (24.0)	61.1 (19.2)
Neutral ($n = 39$)	41.9 (17.8)	3.8 (1.1)	57.5 (24.3)	64.6 (24.7)	42.6 (22.9)	55.5 (16.1)
Incriminating ($n = 39$)	78.0 (22.7)	3.3 (1.1)	67.7 (24.2)	59.4 (25.8)	51.4 (25.3)	68.7 (20.3)
Eyewitness evidence						
Exonerating ($n = 36$)	29.8 (20.5)	3.5 (1.3)	52.9 (31.3)	59.3 (32.3)	55.9 (24.3)	58.4 (24.2)
Neutral ($n = 39$)	39.0 (19.4)	3.6 (1.2)	60.6 (26.6)	56.6 (23.6)	42.8 (26.7)	57.2 (21.7)
Incriminating ($n = 38$)	57.4 (21.4)	3.4 (1.2)	63.7 (26.6)	54.8 (29.8)	55.6 (28.0)	61.4 (19.6)
Total						
Exonerating ($n = 72$)	24.5 (23.0)	3.5 (1.3)	60.2 (27.0)	58.8 (31.3)	48.5 (24.9)	59.8 (21.7)
Neutral ($n = 78$)	40.4 (18.6)	3.7 (1.2)	59.0 (25.4)	60.7 (24.3)	42.7 (24.7)	56.4 (19.0)
Incriminating ($n = 77$)	67.8 (24.2)	3.4 (1.2)	65.7 (25.4)	57.1 (27.7)	53.5 (26.6)	65.1 (20.2)

Note: Alibi strength is measured on a 1–7 scale; all other variables are measured on a 0–100 scale.

Alibi evidence. Participants were provided with a written statement that included the suspect's alibi and relevant facts about the case (Appendix B). This statement contained some information that was considered incriminating and some information that was considered exonerating. For instance, participants were told that the suspect's boss recalled seeing and having a conversation with him when he came into work that night, but they were also told that the time cards were misplaced and there was no way of knowing when the suspect left work. Participants rated the strength of the alibi evidence on a Likert-type scale from 1 (*very weak*) to 7 (*very strong*).

Facial composite evidence. Participants were told that a witness had provided a facial composite of the perpetrator, which they were shown alongside a photograph of the ostensible suspect (top of Figure 1). For generalizability purposes, participants were randomly assigned to receive one of two different composite–photograph pairs. Participants rated the similarity of the facial composite to the suspect's photograph on a Likert-type scale from 0 (*not at all similar*) to 100 (*extremely similar*).

Handwriting evidence. Participants were told that a handwritten note from the perpetrator was left at the crime scene, which they were shown alongside a handwriting sample from the suspect (bottom of Figure 1). Again for generalizability purposes, participants were randomly assigned to receive one of two handwriting sample pairs. Participants rated the similarity of the perpetrator's handwriting left on the note to the suspect's handwriting sample on a scale from 0 (*not at all similar*) to 100 (*extremely similar*).

Informant evidence. Participants were provided with a statement describing testimony given by an informant who claimed to have evidence that the suspect was guilty (Appendix C). This statement was deliberately created to be ambiguous regarding its validity in that some of the information implied the informant was reliable (e.g., the informant included several details about the crime scene that had not been made known to the public) and some of the information implied the informant was unreliable (e.g., the informant had testified in three previous cases as an informant in exchange for leniency on his unrelated charges). Participants rated the extent to which the statement by the informant implicated the suspect on a Likert-type scale from 0 (*strong implication of innocence*) to 100 (*strong implication of guilt*).

After rating each piece of evidence, participants provided a final rating of the likelihood that the suspect was guilty. Participants were then debriefed.

Results

Separate analyses were conducted for the police officer sample and the student sample. For each sample, we first report analyses examining the relationships between experimental condition and initial beliefs of guilt. Subsequent analyses use structural equation modeling to examine a specific model of decision-making in which initial beliefs of guilt predict the latent variable of evidence evaluation tendencies, which subsequently predicts final beliefs of guilt.

Police Officer Sample

Police officers' initial beliefs of guilt ratings, evidence evaluation ratings, and final beliefs of guilt ratings as a function of condition are displayed in Table 1.

Initial beliefs of guilt. A 3 (Valence of Evidence: exonerating, incriminating, neutral) × 2 (Evidence Type: DNA, eyewitness) ANOVA conducted on initial likelihood of guilt scores indicated no significant effect of evidence type, $F(1, 83) = .41$, $p = .52$, $\eta_p^2 = .005$ but did reveal a significant main effect of valence of evidence, $F(2, 83) = 31.2$, $p < .001$, $\eta_p^2 = .429$. Pairwise comparisons indicated that, as expected, incriminating information resulted in higher estimates of guilt ($M = 66.6$, $SD = 28.6$) than neutral evidence ($M = 30.9$, $SD = 21.3$), $t(59) = 5.40$, $p < .001$, $d = 1.41$. Exonerating information did not result in significantly lower estimates of guilt ($M = 23.3$, $SD = 23.4$) than neutral evidence, $t(53) = 1.27$, $p = .21$,

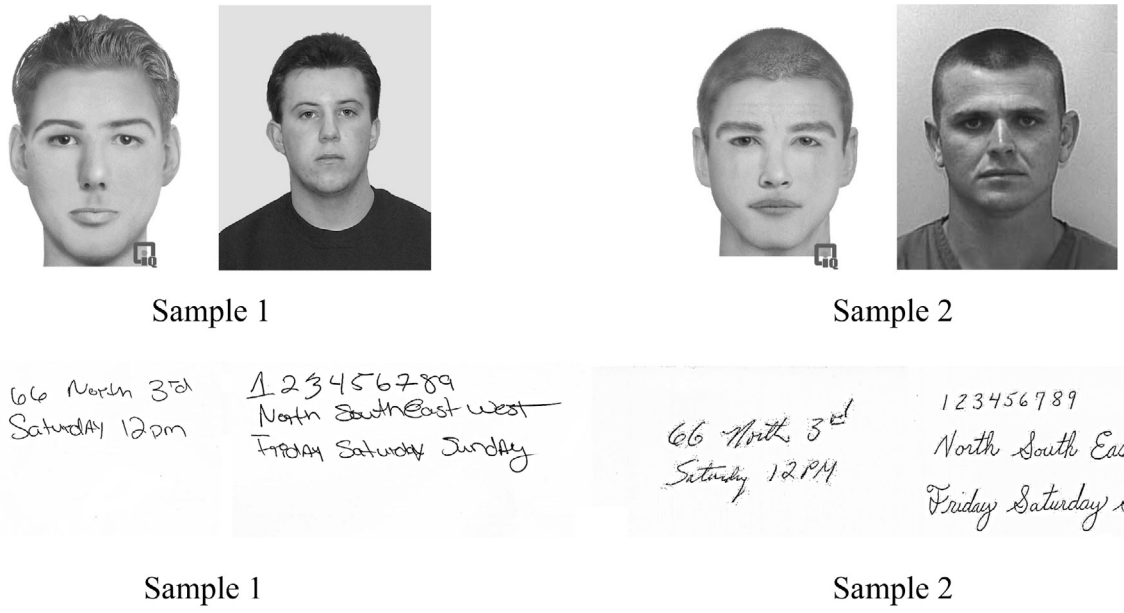

Figure 1. Facial composite–suspect pairs and perpetrator–suspect handwriting samples. Each participant was randomly assigned to receive one of the composite–suspect pair samples, and one of the perpetrator–suspect handwriting samples.

$d = .35$. However, these effects were qualified by a significant Valence of Evidence × Evidence Type interaction, $F(2, 83) = 7.71$, $p = .001$, $\eta_p^2 = .157$. Follow-up analyses indicated that this interaction was the result of incriminating DNA evidence having a greater impact on initial guilt estimates than incriminating eyewitness evidence, $F(1, 57) = 4.19$, $p = .045$, $\eta_p^2 = .068$. Although there was a trend for exonerating DNA evidence to produce a greater decrease in guilt beliefs than exonerating eyewitness evidence, the effect was not significant, $F(1, 51) = 3.25$, $p = .08$, $\eta_p^2 = .060$.

Structural equation modeling. Structural equation modeling was conducted to test whether participants' initial beliefs in the suspect's guilt predicted their evaluation of subsequent evidence, and in turn whether this evaluation predicted the extent to which they updated their beliefs in the suspect's guilt (see Figure 2).

The model fit was acceptable (comparable fit index [CFI] = .92, root mean square error of approximation [RMSEA] = .15, standardized root mean square residual [SRMR] = .07). Fully standardized beta weights were calculated throughout. The total effect, which captures the overall relationship between initial beliefs in the suspect's guilt and final beliefs in the suspect's guilt, was significant, $\beta = .598$, $SE = .070$, 95% CI [.459, .735], $p < .001$. This effect was, however, entirely mediated by police officers' evaluations of evidence. The more strongly the police officers initially believed in the suspect's guilt, the more incriminating they evaluated the subsequent evidence, $\beta = .629$, $SE = .094$, 95% CI [.445, .813], $p < .001$. Specifically, evaluations of three of the four pieces of evidence significantly loaded on the underlying latent variable of evidence evaluation: composite similarity ($\beta = .637$, $SE = .075$, 95% CI [.490, .784], $p < .001$), handwriting similarity ($\beta = .604$, $SE = .075$, 95% CI [.457, .751], $p < .001$), and implication by informant ($\beta = .672$, $SE = .070$, 95% CI [.535, .809], $p < .001$). Perceived alibi strength did not significantly load on the underlying latent variable of evidence evaluation ($\beta = .099$, $SE = .114$, 95% CI [−.124, .322], $p = .39$).

Furthermore, the more incriminating the police officers evaluated the evidence, the stronger were their final beliefs in the suspect's guilt, $\beta = 1.022$, $SE = .129$, 95% CI [.769, 1.275], $p < .001$. There was no direct effect of police officers' initial beliefs in the suspect's guilt on their final beliefs in the suspect's guilt, $\beta = -.044$, $SE = .142$, 95% CI [−.322, .234], $p = .76$, indicating that the relationship between initial beliefs and final beliefs was fully mediated via police officers' evaluations of evidence.

Student Sample

Students' initial beliefs of guilt ratings, evidence evaluation ratings, and final beliefs of guilt ratings as a function of condition are displayed in Table 2.

Initial beliefs of guilt. A 3 (Valence of Evidence: exonerating, incriminating, neutral) × 2 (Evidence Type: DNA, eyewitness) ANOVA conducted on initial likelihood of guilt scores indicated no significant effect of evidence type, $F(1, 221) = 2.41$, $p = .12$, $\eta_p^2 = .011$, but did reveal a significant main effect of valence of evidence, $F(2, 221) = 80.2$, $p < .001$, $\eta_p^2 = .421$. Pairwise comparisons indicated that, as expected, incriminating information resulted in higher estimates of guilt ($M = 67.8$, $SD = 24.2$) than neutral evidence ($M = 40.4$, $SD = 18.6$), $t(153) = 7.91$, $p < .001$, $d = 1.28$. Furthermore, exonerating information resulted in lower estimates of guilt ($M = 24.5$, $SD = 23.0$) than neutral evidence, $t(148) = 4.68$, $p < .001$, $d = .77$. However, these effects were qualified by a significant Valence of Evidence × Evidence Type interaction, $F(2, 221) = 10.12$, $p < .001$, $\eta_p^2 = .084$. Follow-up analyses indicated that this interaction was the result of incriminating DNA evidence having a greater impact on initial guilt estimates than

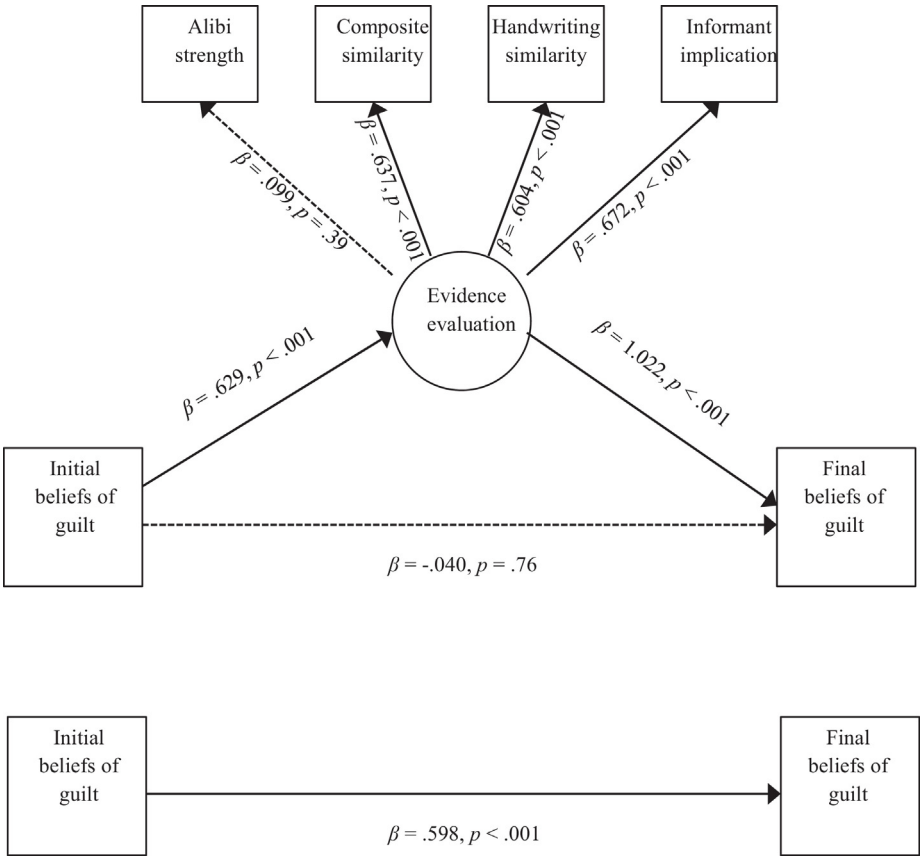

Figure 2. Mediated path diagram (top) and total effect (bottom) of the effects of initial beliefs in a suspect's guilt on final beliefs in the suspect's guilt among police officers. Dotted lines represent non-significant pathways.

incriminating eyewitness evidence, $F(1, 151) = 7.23$, $p = .008$, $\eta_p^2 = .046$. Similarly, exonerating DNA evidence produced a significantly greater decrease in guilt beliefs than exonerating eyewitness evidence, $F(1, 146) = 3.93$, $p = .049$, $\eta_p^2 = .026$.

Structural equation modeling. A similar structural equation model was constructed for students as for police officers (see Figure 3). Model fit was strong (CFI = 1.0, RMSEA = .01, SRMR = .03). Fully standardized beta weights were calculated throughout. Results were similar for students as for police officers. The total effect, which captures the overall relationship between initial beliefs in the suspect's guilt and final beliefs in the suspect's guilt, was significant, $\beta = .457$, $SE = .054$, 95% CI [.351, .563], $p = .001$. This effect was, however, entirely mediated by students' evaluations of evidence. The more strongly students initially believed in the suspect's guilt, the more incriminating they evaluated the subsequent evidence, $\beta = .380$, $SE = .115$, 95% CI [.155, .605], $p = .001$. Specifically, evaluations of three of the four pieces of evidence significantly loaded on the underlying latent variable of evidence evaluation: composite similarity ($\beta = .409$, $SE = .082$, 95% CI [.248, .570], $p < .001$), handwriting similarity ($\beta = .432$, $SE = .076$, 95% CI [.283, .581], $p < .001$), and implication by informant ($\beta = .305$, $SE = .072$, 95% CI [.164, .446], $p < .001$). Alibi strength did not significantly load on the underlying latent variable of evidence evaluation ($\beta = .052$, $SE = .073$, 95% CI [−.091, .195], $p = .47$).

The more incriminating the students evaluated the evidence, the stronger were their final beliefs in the suspect's guilt, $\beta = .898$, $SE = .152$, 95% CI [.600, 1.196], $p < .001$. There was no direct effect of students' initial beliefs in the suspect's guilt on their final beliefs in the suspect's guilt, $\beta = .116$, $SE = .136$, 95% CI [−.111, .343], $p = .39$, indicating that the relationship between initial beliefs and final beliefs was fully mediated via students' evaluations of evidence.

Discussion

Police officers' evaluations of evidence were related to their initial beliefs in a suspect's guilt: The more likely they were to believe the suspect was guilty, the more incriminating they perceived subsequent ambiguous evidence to be. These effects occurred for three pieces of evidence (facial composite similarity, handwriting similarity, and extent to which the informant's testimony implicated the suspect) that have received very little attention from legal psychology researchers. In fact, we are aware of only one study that examined whether preexisting beliefs of guilt affect perceptions of similarity between a facial composite of a perpetrator and a photograph of a suspect (Charman et al., 2009), only one study that examined whether preexisting beliefs affect handwriting sample comparisons (Kukucka & Kassin, 2014), and no studies that have examined the effects of preexisting beliefs on informant believability. Further, we are aware of no studies that have examined

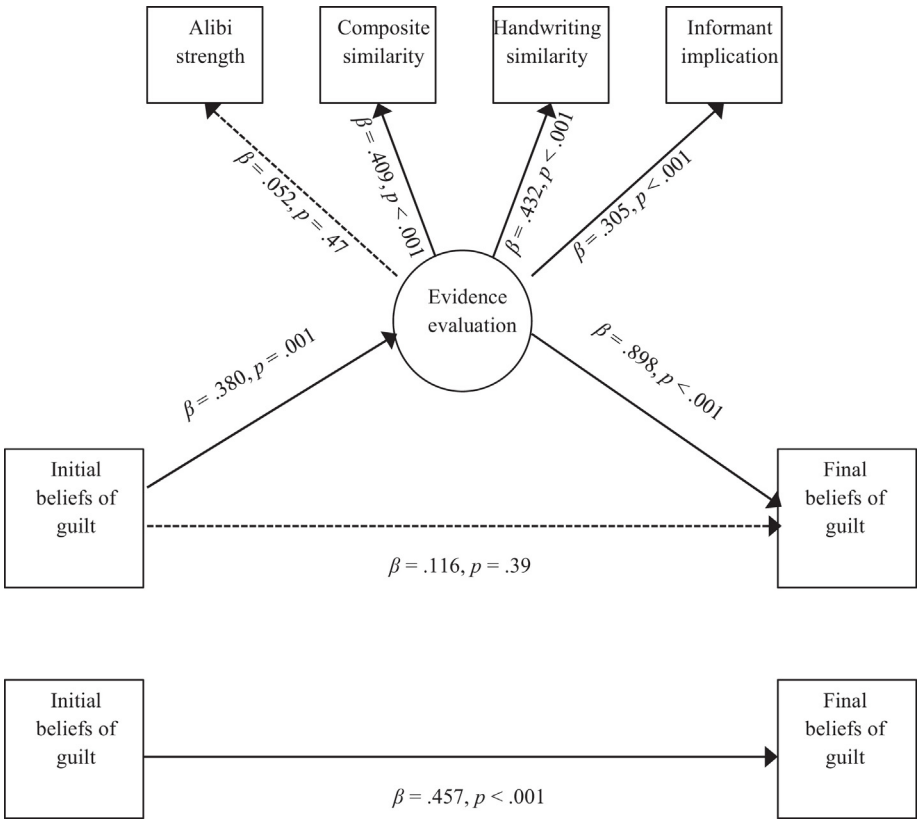

Figure 3. Mediated path diagram (top) and total effect (bottom) of the effects of initial beliefs in a suspect's guilt on final beliefs in the suspect's guilt among students. Dotted lines represent non-significant pathways.

these processes specifically among police officers. The current results thus expand upon the types of judgments that have been shown to be biased by evaluators' beliefs in a suspect's guilt.

Notably, these general context effects occurred for the three pieces of potentially incriminating evidence but not for the single piece of potentially exonerating evidence (alibi strength). Although speculative, it is possible that context effects are larger among evaluations of evidence that imply guilt: It may be easier for a belief to nudge the evaluation of a potentially incriminating piece of evidence than it is to nudge the evaluation of a potentially exonerating piece of evidence.

Data are consistent with the existence of a bias snowball effect (Dror, 2012): Initial beliefs in a suspect's guilt led police officers to interpret ambiguous evidence as indicative of guilt, which led them to more strongly believe in the suspect's guilt. In other words, beliefs of guilt were self-reinforcing, a tendency that should make it particularly difficult for innocent people to exonerate themselves once they become suspects. This is an important point, as it demonstrates not just the immediate negative effects of contextual bias, but also its later downstream consequences: Contextual biases are dynamic and can accumulate. This finding suggests that even a relatively weak initial belief in a suspect's guilt can grow over time, as that belief biases evaluation of evidence to appear more incriminating, further inflating beliefs of guilt.

An inspection of the structural equation modeling results indicates a high degree of similarity between students and police officers: In both samples, the relationship between initial beliefs of guilt and evidence evaluation was moderate-to-strong, the relationship between evidence evaluation and final beliefs of guilt was strong, and the direct effect of initial beliefs of guilt on final beliefs of guilt dropped to statistical non-significance when evidence evaluation was included as a mediator in the model. No evidence was found that police officers are immune to this bias, suggesting that context effect data from past (and future) studies using student samples should tend to generalize to police officers as well. That said, these results should be taken with some caution, as our police officer sample was, of course, comprised of only those officers willing to participate in a research study. Whether these results generalize to other police officers is unknown. However, we might expect that officers willing to participate in research and to have their responses scrutinized would be particularly mindful about their decision-making; if anything, we may therefore expect even stronger effects among police officers when evaluating evidence in the real world.

In fact, data support the assertion that these biases occur among police officers and detectives in the real world. Kassin et al.'s (2012) analysis of real-world false convictions indicated that cases involving false confessions were more likely to be associated with multiple other pieces of incriminating evidence, and that in those cases, confessions were likely to have been obtained first. Kassin et al. interpreted these data as indicating that confession evidence can corrupt other evidence. Our finding that initial incriminating evidence can taint the evaluation of subsequent evidence is consistent with this interpretation. However, we broaden the scope of this effect by (a) showing that corruptive

influences are not limited to confessions, such that eyewitness evidence and DNA evidence also predicted subsequent evidence evaluation (via their effects on initial guilt beliefs), and (b) showing that multiple disparate forms of evidence (i.e., facial composite similarity, handwriting similarity, and extent to which an informant's testimony implicates a suspect) are affected by these corruptive influences.

Practical Application

To avoid context effects, Kassin et al. (2013) suggested keeping evaluators blind to extraneous case information when making their forensic judgments. Along similar lines, Dror et al. (2015) proposed a Linear Sequential Unmasking procedure, in which the evaluator (often a forensic scientist) be kept unaware as much as possible of task-irrelevant information (such as case specifics). We agree with these suggestions, and the current data provide additional evidence for the importance of limiting extraneous information as much as possible. However, current results show these context effects to be particularly insidious for two reasons. First, we showed that initial beliefs of guilt can affect evidence evaluation among police officers and detectives—people who cannot reasonably be kept blind to extraneous case information. Second, we showed that initial beliefs predict judgments that are inherently subjective in nature (e.g., judgments of similarity), and which leave no record of their occurrence. The direct biasing influence of those beliefs on judgments is effectively invisible.

We thus recommend that researchers begin to shift their focus from simply demonstrating these biases to investigating strategies that may mitigate them, as few real-world procedures have been implemented within the legal system to reduce the likelihood of contextual biases (Stoel, Dror, & Miller, 2014). The challenge will be to develop procedures to mitigate these biases even among evaluators, such as police officers and detectives, who, out of necessity, possess contextual information. Unfortunately, efforts to reduce contextual biases have largely failed (e.g., Simon et al., 2001, although see Simon, Krawczyk, Bleicher, & Holyoak, 2008, for evidence that increased time delay between judgments may mitigate these effects). However, some research suggests that activating a thoroughness goal—as opposed to an efficiency goal—may result in deeper processing of information, which may reduce context effects (Ask et al., 2011). Training police officers to adopt this goal may warrant further research.

As the National Academy of Sciences (2009) points out, it is crucial to understand the process of forensic confirmation bias, including the extent of the bias and the conditions under which it occurs. The current study shows what had been suspected but not yet empirically shown: Initial beliefs of guilt can bias police officers' evaluations of a variety of evidence against a suspect, further inflating their beliefs of guilt.

Conflict of Interest Statement

The authors of this manuscript declare no conflicts of interest.

Author Contributions

S.D. Charman developed the idea for the study. All authors contributed to the study design. Data collection was performed by M. Kavetski and D. Hirn. S.D. Charman analyzed the data and drafted the manuscript, with input from M. Kavetski and D. Hirn. All authors approved the final version of the manuscript for submission.

Appendix A. Case Synopsis and Initial Evidence

All participants read the initial case synopsis; they were then randomly assigned to read one of the following six excerpts regarding the eyewitness/DNA evidence.

During the early morning hours, a 911 call was made by a young man. The man was visiting his mother-in-law who lived alone since her husband's passing. When his mother-in-law didn't answer the door, the man used his spare key to enter the home. Upon entering, he discovered the woman's body; she had been stabbed to death. Crime scene investigators determined the woman was killed the night before. Samples of blood and fingerprints were recovered from several rooms throughout the house. Investigators also spoke with several neighbors, two of whom claimed to have seen a suspicious man leaving the home the previous evening. Descriptions of the 'suspicious man' were taken from the witnesses and a composite sketch was constructed. The murder weapon was not found inside the home.

Police had no suspects on the first two days of investigation. On the third day, police received an anonymous tip that a man who lives in apartment 2D across the street was seen snooping around the woman's home a few days prior to her murder. (We will refer to this man as John Doe.) Since they had no other leads, police officers arrived at John Doe's apartment complex and asked him to come to the station to answer a few questions. John Doe agreed.

[Incriminating DNA condition]

Upon arriving at the station, officers took a DNA sample from John Doe. After analyzing John's DNA sample and the blood samples taken from the crime scene, an expert DNA analyst concluded that although most of the blood found at the scene of the crime was the victim's, some of it came from a second person. Based on the DNA results, the analyst was 99.9% sure that it was John Doe's blood.

[Exonerating DNA condition]

Upon arriving at the station, officers took a DNA sample from John Doe. After analyzing John's DNA sample and the blood samples taken from the crime scene, an expert DNA analyst concluded that although most of the blood found at the scene of the crime was the victim's, some of it came from a second person. Based on the DNA results, the analyst was 99.9% sure that it was NOT John Doe's blood.

[Neutral DNA condition]

Upon arriving at the station, officers took a DNA sample from John Doe. After analyzing John's DNA sample and the blood samples taken from the crime scene, an expert DNA analyst concluded that although most of the blood found at the scene of the crime was the victim's, some of it came from a second person. Due to the limited amount of suitable DNA recovered

from the crime scene, the DNA results of this second person were interpreted as inconclusive, meaning that John Doe's DNA sample and the samples of blood from the crime scene could not be compared.

[Incriminating Eyewitness condition]

After apprehending the suspect (John Doe), police returned to the eyewitnesses to see whether they thought they would be able to identify the man they saw leaving their neighbor's house the evening before her body was found. They both were certain they would be able to identify the man if he was in a lineup. Thus, police officers constructed a six-person photographic lineup, which included a photo of the suspect, and asked the neighbors (separately) whether the man he saw leaving the victim's home was present in the lineup.

The first eyewitness told police officers he was 85% certain that the man he saw leaving his neighbor's house was in the lineup.

The second eyewitness almost immediately pointed out John Doe and told officers he was certain this was the man he previously saw. When officers asked the eyewitness how confident he was in this decision, he responded that he was 90% certain.

[Exonerating Eyewitness condition]

After apprehending the suspect (John Doe), police returned to the eyewitnesses to see whether they thought they would be able to identify the man they saw leaving their neighbor's house the evening before her body was found. They both were certain they would be able to identify the man if he was in a lineup. Thus, police officers constructed a six-person photographic lineup, which included a photo of the suspect, and asked the neighbors (separately) whether the man he saw leaving the victim's home was present in the lineup.

The first eyewitness told police officers he was 85% certain that the man he saw leaving his neighbor's house was NOT in the lineup.

The second eyewitness almost immediately told officers he was certain the man he previously saw was NOT in the lineup. When officers asked the eyewitness how confident he was in this decision, he responded that he was 90% certain.

[Neutral Eyewitness condition]

After apprehending the suspect (John Doe), police returned to the eyewitnesses to see whether they thought they would be able to identify the man they saw leaving their neighbor's house the evening before her body was found. Both neighbors felt that they would not be able to identify the man they saw leaving the victim's house if he was in a lineup so the police chose to not administer a lineup to the witnesses.

Appendix B. Suspect's Alibi

Police interviewed John Doe on his whereabouts the night the woman was killed in her home. Initially, he reported having been with his girlfriend that night. However, when officers questioned the girlfriend, she said she had been out of town at that time. When questioned again, John Doe told officers he was mistaken and that he was thinking of the wrong night. Instead, he claimed to have gone into work at a local warehouse to clock in some overtime and that he was there most of the night. Investigators spoke with John Doe's boss, who informed them that he recalled seeing and having a conversation with John that night when he came into work. However, the time cards were misplaced and since the boss left early that night, there was no way of knowing what time John Doe clocked out. To further investigate this matter, police officers requested the security cameras from the warehouse. The footage showed a red car leaving the warehouse approximately half an hour before the time of the crime. John Doe drives a red car. However, the videos were very grainy and didn't allow a very good view of the parking lot where the employees enter and leave. Because of the graininess of the video and the darkness outside, the license plate of the car leaving the warehouse could not be seen and officers could not even be certain of the make and model of the car. Furthermore, two of John's co-workers claimed to police that they believed John was at work all night.

Appendix C. Informant Testimony

During the investigation, J.G., a former co-worked of John Doe's, came forward and offered information to police. He informed officers that he had information regarding a case where a woman was stabbed to death in her home. He fist told officers that he was currently facing robbery and assault charges and asked if he would be granted leniency in return for the information he could provide in the murder case. Police couldn't guarantee J.G. leniency in his unrelated charges. Thus, they interviewed him several times before he provided any information to the police. According to J.G., John Doe told him several details about the crime but that he never actually confessed. During the final time that J.G. was interviewed, he was asked to provide a written statement of what he knew regarding the crime. In his statement, J.G. included several details about the crime scene that had not been made available to the public. For instance, he discussed the fact that the attacker broke into the house through the window on the back door. When asked why he was providing police with this information, J.G. stated that he felt bad for the victim and her family. Upon investigation, officers found that J.G. has been arrested multiple times and has testified as an informant in three previous cases in exchange for leniency on the unrelated charges that were being brought against him.

References

Ask, K., & Granhag, P. A. (2007). Hot cognition in investigative judgments: The differential influence of anger and sadness. *Law and Human Behavior*, *31*, 537–551. http://dx.doi.org/10.1007/s10979-006-9075-3

Ask, K., Granhag, P. A., & Rebelius, A. (2011). Investigators under influence: How social norms activate goal-directed processing of criminal evidence. *Applied Cognitive Psychology*, *25*, 548–553. http://dx.doi.org/10.1002/acp.1724

Ask, K., Rebelius, A., & Granhag, P. A. (2008). The elasticity of criminal evidence: A moderator of investigator bias. *Applied Cognitive Psychology*, *22*, 1245–1259. http://dx.doi.org/10.1002/acp.1432

Charman, S. D., Gregory, A. H., & Carlucci, M. (2009). Exploring the diagnostic utility of facial composites: Beliefs of guilt can bias perceived similarity between composite and

suspect. *Journal of Experimental Psychology: Applied*, *15*, 76–90. http://dx.doi.org/10.1037/a0014682

Dror, I. E. (2012). Cognitive bias in forensic science. In *The 2012 yearbook of science & technology*. pp. 43–45. New York: McGraw-Hill.

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: Journal of the Forensic Science Society*, *51*(4), 204–208. http://dx.doi.org/10.1016/j.scijus.2011.08.004

Dror, I. E., Thompson, W. C., Meissner, C. A., Kornfield, I., Krane, D., Saks, M., & Risinger, M. (2015). Context management toolbox: A linear sequential unmasking (LSU) approach for minimizing cognitive bias in forensic decision making. *Journal of Forensic Sciences*, *60*, 1111–1112.

Hasel, L. E., & Kassin, S. M. (2009). On the presumption of evidentiary independence: Can confessions corrupt eyewitness identifications? *Psychological Science*, *20*, 122–126. http://dx.doi.org/10.1111/j.1467-9280.2008.02262.x

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General*, *128*, 3–31.

Kassin, S. M., Bogart, D., & Kerner, J. (2012). Confessions that corrupt: Evidence from the DNA exoneration case files. *Psychological Science*, *23*, 41–45. http://dx.doi.org/10.1177/0956797611422918

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

Kukucka, J., & Kassin, S. M. (2014). Do confessions taint perceptions of handwriting evidence? An empirical test of the forensic confirmation bias. *Law and Human Behavior*, *38*, 256–270. http://dx.doi.org/10.1037/lhb0000066

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098–2109.

Marksteiner, T., Ask, K., Reinhard, M. A., & Granhag, P. A. (2011). Asymmetrical scepticism towards criminal evidence: The role of goal-and belief-consistency. *Applied Cognitive Psychology*, *25*, 541–547. http://dx.doi.org/10.1002/acp.1719

National Academy of Sciences. (2009). *Strengthening forensic science in the United States: A path forward*. Washington, DC: National Academies Press.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220. http://dx.doi.org/10.1037/1089-2680.2.2.175

Simon, D., Krawczyk, D. C., Bleicher, A., & Holyoak, K. J. (2008). The transience of constructed preferences. *Journal of Behavioral Decision Making*, *21*, 1–14. http://dx.doi.org/10.1002/bdm.575

Simon, D., Pham, L. B., Le, Q. A., & Holyoak, K. J. (2001). The emergence of coherence over the course of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1250–1260.

Simon, D., Snow, C. J., & Read, S. J. (2004). The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology*, *86*, 814–837. http://dx.doi.org/10.1037/0022-3514.86.6.814

Stoel, R. D., Dror, I. E., & Miller, L. S. (2014). Bias among forensic document examiners: Still a need for procedural changes. *Australian Journal of Forensic Sciences*, *46*(1), 91–97. http://dx.doi.org/10.1080/00450618.2013.797026

U.S. Department of Justice, Office of the Inspector General. (2006). *A review of the FBI's handling of the Brandon Mayfield case*. https://oig.justice.gov/special/s0601/final.pdf

Received 17 August 2016;
received in revised form 26 January 2017;
accepted 2 February 2017
Available online 24 May 2017