






## Law and order effects: on cognitive dissonance and belief perseverance

Enide Maegherman<sup>a,b</sup> , Karl Ask<sup>b</sup> , Robert Horselenberg<sup>a</sup> and Peter J. van Koppen<sup>a,c</sup> 

<sup>a</sup>Department of Criminal Law and Criminology, Faculty of Law, Maastricht University, Maastricht, The Netherlands; <sup>b</sup>Department of Psychology, Gothenburg University, Gothenburg, Sweden; <sup>c</sup>Department of Criminal Law and Criminology, Faculty of Law, VU University Amsterdam, Amsterdam, The Netherlands

Order of evidence presentation affects the evaluation and the integration of evidence in mock criminal cases. In this study, we aimed to determine whether the order in which incriminating and exonerating evidence is presented influences cognitive dissonance and subsequent display of confirmation bias. Law students ( $N=407$ ) were presented with a murder case vignette, followed by incriminating and exonerating evidence in various orders. Contrary to a predicted primacy effect (i.e. early evidence being most influential), a recency effect (i.e. late evidence being most influential) was observed in ratings of likelihood of the suspect's guilt. The cognitive dissonance ratings and conviction rates were not affected by the order of evidence presentation. The effects of evidence presentation order may be limited to specific aspects of legal decisions. However, there is a need to replicate the results using procedures and samples that are more representative of real-life criminal law trials.

**Key words:** cognitive dissonance; confirmation bias; criminal law; evidence; judges; legal decision-making; legal psychology; order effects.

In criminal trials, the main question to be answered by the judge or jury is whether it is proven beyond reasonable doubt that the defendant committed the crime he or she is accused of. The judge or jury must do so based on the evidence that has been presented to them by the prosecution and the defence. That process will usually require the judge or jury to draw inferences from the evidence, and to combine these to reach a final decision on guilt. The reasoning process required to reach a decision on guilt ultimately relies on human cognition, and is therefore likely to be

vulnerable to cognitive biases. Such biases can affect how information is processed, including how it is evaluated and integrated (Charman et al., 2019). One form of bias that has been argued to be particularly relevant to the legal system is confirmation bias (Findley & Scott, 2006). In this study, we aimed to determine whether the order in which evidence is presented influences mock legal decision-makers' cognitive dissonance and the subsequent presence of confirmation bias in the context of criminal proceedings.

---

Correspondence: Enide Maegherman. Department of Criminal Law and Criminology, Faculty of Law, Maastricht University, Box 616, 6200, Maastricht, The Netherlands. Email: [enide.maegherman@maastrichtuniversity.nl](mailto:enide.maegherman@maastrichtuniversity.nl)

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.  
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

***Biased reasoning***

Confirmation bias refers to the human tendency to seek and interpret evidence in ways that favour an existing belief, thereby influencing reasoning and decision-making. That influence can have a negative impact. It has been argued to be one of the most problematic aspects of human reasoning (Nickerson, 1988). The existence of confirmation bias has been demonstrated in various areas of criminal investigations, ranging from interviews to forensic investigation (e.g. Ask & Granhag, 2005; Dror et al., 2006; Kassin et al., 2003). Even domains that were thought to be objective, such as forensic toxicology, have been shown to be impacted by confirmation bias (Hamnett & Jack, 2019). Dror (2020) argued that bias does not merely affect the interpretation of information, but even the obtainment of data and the use of analyses, suggesting a wide-ranging influence.

In order for confirmation bias to influence reasoning, one first needs to hold a belief. Therefore, confirmation bias is closely related to belief perseverance, a basic human tendency to adhere to accepted theories or explanations even when faced with discrediting information (Burke, 2007). Multiple researchers have shown that changing people's initial beliefs by confronting them with contradicting evidence can be difficult to achieve. In a classic study, Lord et al. (1979) presented participants with information either confirming or disconfirming their view on capital punishment. Participants found the evidence confirming their initial belief to be more convincing, and there was a lack of belief adjustment in response to contradicting information. In a more recent study, Green and Donahue (2011) presented participants with a story about a young heroin addict. Participants were then told that the story was inaccurate, due to either accidental error or intentional deception. Although those in the intentional deception condition were motivated to change their belief, participants in both conditions maintained beliefs about items specific to the story (e.g. failure of social programmes

to assist young people). Furthermore, Burke (2007) describes how, in many exoneration cases, prosecutors have maintained that the exonerated individual is guilty despite overwhelming evidence to the contrary. In short, belief perseverance is a well-documented issue, for which effective solutions are currently lacking.

Belief perseverance has also been researched in the applied setting of legal proceedings. Schünemann (1983, as cited in Schünemann & Bandilla, 1989) offered valuable insights by investigating belief perseverance in judges. In his study, judges who had been given more incriminating information prior to trial were more likely to convict the defendant than judges who were given the same case file, but less incriminating prior information. Therefore, judges also appear to be prone to belief perseverance despite the need for impartiality.

In the Netherlands, the case file that judges read before the trial proper is likely to consist mainly of incriminating information (Crombag, 2017). That case file, or a summary of the case file, is likely to form the basis of the judge's initial belief. Although the case file is mostly incriminating, it is also likely to contain some information that is contradictory to the suspect being guilty. Such information will often be emphasised by the defence. The judge would then have to process arguments supporting two different beliefs. Experimental psychological research has shown that the consideration of conflicting beliefs can create an uneasy feeling, known as cognitive dissonance (Festinger, 1957). One way in which people try to escape cognitive dissonance is to adopt, and adhere to, one of the beliefs, while refuting or downplaying the other (Jonas et al., 2008).

Cognitive dissonance theory has been researched in various experimental studies. Harmon-Jones and Harmon-Jones (2007) reviewed the development of Festinger's theory and concluded that cognitive dissonance is mainly due to inconsistency rather than the possibility of aversive consequences due to

the chosen alternative. Legal psychological research indicates that cognitive dissonance may arise in the presence of belief-inconsistent criminal evidence (Ask et al., 2011). Participants who received evidence inconsistent with their initial belief about a suspect's guilt reported experiencing more dissonance than those who received consistent evidence. They also found that those who experienced the strongest dissonance were the most reluctant to adjust their judgement of guilt in line with the contradicting evidence. Moreover, participants who received contradicting witness evidence reported experiencing stronger dissonance than those who received contradicting DNA evidence. According to Ask et al. (2011), that difference can be explained by the fact that witness evidence has higher elasticity (i.e. room for subjective interpretation) than DNA evidence. Thus, participants had little choice but to change their belief in line with the compelling DNA evidence, and, as the conflict between beliefs was resolved, these participants experienced less dissonance. In contrast, the contradicting witness evidence evoked stronger dissonance as it created conflicting beliefs, none of which was clearly more compelling than the other (Ask et al., 2008). In sum, Ask et al. (2011) found support for the fact that the cognitive dissonance experienced is related to the evaluation of criminal evidence. They also found support for a mechanism whereby cognitive dissonance can be reduced—ignoring contradicting evidence.

The preference for supporting information in response to cognitive dissonance is akin to confirmation bias (Jonas et al., 2001). There is a strong theoretical relation between the concepts demonstrated above. Namely, after one forms an initial idea, belief perseverance sets in by trying to adhere to that idea. Cognitive dissonance will subsequently arise in response to information that contradicts that idea. As cognitive dissonance is an uncomfortable feeling, the tendency to prefer evidence that confirms the idea, while paying disproportionately less attention to contradicting evidence, will

develop. In that way, confirmation bias can help reduce cognitive dissonance and maintain belief perseverance. Therefore, manipulating belief perseverance and cognitive dissonance may affect the presence of confirmation bias. Whether cognitive dissonance is affected by the presentation of multiple pieces of evidence, varying in consistency with an existing belief, has not yet been researched.

### **Order effects**

As shown above, cognitive dissonance is associated with a reluctance to adjust perceptions of guilt in line with contradicting evidence. A way to reduce such belief perseverance may be to present evidence to jurors in a way that prevents the formation of a strong prior belief and, subsequently, the development of dissonance. For instance, less cognitive dissonance may be experienced when one is, from the outset, presented with both supporting and contradictory information, rather than first receiving all the evidence in favour of one view, followed by all the evidence in favour of the other. In other words, it may be possible to reduce the development of confirmation bias by changing the order in which evidence is presented.

Effects of the order in which evidence is presented were examined by Pennington and Hastie (1988). Participants perceived evidence as stronger when it was presented in a story order rather than in the order in which the witnesses provided their statements. When the prosecution's evidence was presented as a story, a majority of participants convicted the suspect. Similarly, when the defence's evidence was presented in the story order, a majority of participants acquitted the suspect. The totality of evidence presented to participants in these conditions was identical, suggesting that the difference in verdicts must be due to the different presentation orders. The importance of a coherent story has also been recognised in other influential theories on legal decision-making (e.g. Simon, 2004; Wagenaar et al., 1993).

In previous research, the order in which evidence is presented has been investigated mostly in relation to specific types of evidence (e.g. alibi, DNA or witness evidence). For instance, Price and Dahl (2014), who used alibi and eyewitness evidence, found an influence of strong evidence that was presented last on the evaluation of the suspect's guilt overall, as well as on the evaluation of individual pieces of evidence. The finding of such a *recency effect* is contrary to what would be expected in line with belief perseverance; belief perseverance can be considered a form of *primacy effect*, where the conclusion is based disproportionately on the initial evidence.

Charman et al. (2016) made use of DNA evidence and alibi evidence to test the influence of order of presentation. They concluded that the initial evidence had an influence on the evaluation of subsequent evidence, which can be considered supportive of a primacy effect and belief perseverance. Moreover, they concluded that the piece of evidence that is presented last does not retroactively influence evaluations of the evidence that was presented earlier. However, similarly to the finding by Price and Dahl (2014), Charman et al. (2016) also found that the piece of evidence that was presented last had a greater impact on the overall assessment of guilt. Charman et al. (2016) therefore suggested that *evidence evaluation* is likely to be influenced by an existing belief, whereas the *evidence integration* is likely to be influenced by a recency effect. As these studies on order effects looked only at the presentation of two pieces of evidence, the effect of alternating the presentation of incriminating and exonerating evidence remains unexamined.

### **The current study**

The aims of the current study were two-fold. Firstly, we attempted to replicate the previous finding that participants maintain their initial beliefs when presented with contradicting information (i.e. belief perseverance; e.g. Green & Donahue, 2011). Secondly, we

investigated whether the order in which incriminating and exonerating evidence are presented to participants affects the occurrence of cognitive dissonance and belief perseverance. The types of evidence used in this study included hair, CCTV, shoeprints and witness evidence, thereby also extending previous research that used a limited number of evidence types.

Participants in this study were first presented with a summary description of a homicide case, in which a suspect had been arrested. Their cognitive dissonance was then measured, as well as their initial impression of the suspect's guilt. They then received two pieces of incriminating evidence, exculpatory evidence or a mixture of incriminating and exculpatory evidence. After having rated the likelihood of the suspect being guilty and their experienced cognitive dissonance, they received an additional two pieces of exculpatory evidence, incriminating evidence or a mixture of the two. Depending on experimental condition, the valence of the evidence received in the second set was either similar to or different from the evidence participants received in the first set. Participants then made final ratings of suspect guilt and cognitive dissonance.

First, we predicted that the order in which the evidence is presented would make a difference to participants' final guilt ratings. Specifically, participants who received incriminating evidence in the first set and exonerating evidence in the second set (*Inc/Ex*) would make higher final ratings of guilt than participants who received mixed evidence in both sets (H1). Those in the mixed conditions were not expected to strongly commit to one belief, as they had received both incriminating and exonerating evidence. However, those in the contradictory conditions had only seen either incriminating or exonerating evidence and were therefore expected to commit to guilt or innocence, respectively. Similarly, participants who received exonerating information in the first set and incriminating

evidence in the second set (*Ex/Inc*) would make lower final ratings of guilt than participants who received mixed evidence in both sets (H2). Because all the above participants had received the same pieces of evidence in total, but in different order, the predicted effects can be attributed to order effects.

Second, we expected that participants in the contradictory conditions would show belief perseverance, thus preserving their initial belief when presented with contradictory evidence. Hence, we predicted that participants in the *Inc/Ex* condition would make higher final judgements of guilt than participants in the *Ex/Inc* condition (H3).

Third, we predicted that the final judgment of cognitive dissonance would differ between conditions. Specifically, participants who received mixed evidence in both sets would report less cognitive dissonance than participants who received contradictory evidence (i.e. *Inc/Ex* and *Ex/Inc*; H4), but more cognitive dissonance than participants in the control conditions (who received uniformly incriminating or exonerating evidence across both sets; H5). Finally, we explored whether the amount of cognitive dissonance was associated with changes in guilt ratings between the first and the second set of evidence – that is, whether dissonance was associated with belief perseverance, as reported in previous research (Ask et al., 2011). As no such hypothesis was included in the preregistration, we tested this exploratorily. The preregistration for this study can be found at:<sup>1</sup> [https://osf.io/ywajg/?view\\_only=a94c0cd904ed4c2cbe3bb8d82e5185b6](https://osf.io/ywajg/?view_only=a94c0cd904ed4c2cbe3bb8d82e5185b6)

## Method

### Participants

Law students were recruited through several communication platforms at Maastricht University as well as through social media and other forms of advertisement (e.g. flyers

handed out at university buildings). It was decided to use law students due to their affinity with the context of the study, namely legal proceedings, as well as their familiarity with the decision that participants were required to make. Professional judges are notoriously difficult to use as participants in experimental research. As the focus of the study was decision-making by judges in a European, inquisitorial system, European law students were considered to be the most appropriate sample.

Participants could win a shopping voucher worth €10 for their participation in the study. An *a priori* power analysis, conducted using G\*power 3.1 (Faul et al., 2009), showed that a total of 276 (46 per condition) participants were needed to detect a small effect size ( $f = 0.10$ ) with 80% power at a significance level of .05. Data were collected in both Dutch and English at the same time (see below). Random allocation was used in both language versions, which resulted in some groups receiving more participants than necessary. It was therefore decided, prior to conducting any analyses, to add participants in the remaining conditions to achieve approximately equal group sizes. Across both samples, a total of 474 participants completed the survey. In total, 67 participants were excluded from the initial sample for three reasons: firstly, 51 participants failed to complete the initial cognitive dissonance measure. Secondly, 13 participants were removed because they had an educational background other than law. Finally, three participants completed the study in under 240 seconds. Prior to conducting any analysis, that completion time was deemed insufficient to read through the material properly, and these participants were also excluded. After exclusions, the final sample consisted of 407 participants. Participants' average age was 22.97 years ( $SD = 4.24$ ), and the majority of participants were female (62.7%). The survey was offered in both English and Dutch, and participants chose their preferred language. Most participants completed the survey in English (58.7%). Ethical approval for this

<sup>1</sup>The data from this study are available from the corresponding author upon request.

Table 1. Representation of the order presentation in the conditions.

Condition	Evidence	
	First set	Second set
Mixed evidence		
Mixed 1	Incriminating/exonerating	Incriminating/exonerating
Mixed 2	Exonerating/incriminating	Exonerating/incriminating
Contradictory evidence		
Ex/Inc	Exonerating/exonerating	Incriminating/incriminating
Inc/Ex	Incriminating/incriminating	Exonerating/exonerating
Control conditions		
Ex/Ex	Exonerating/exonerating	Exonerating/exonerating
Inc/Inc	Incriminating/incriminating	Incriminating/incriminating

study was obtained from the Ethical Committee at Maastricht University.

### Design

The study used a 6 (evidence order)  $\times$  3 (time) mixed factorial design, with time as within-participants factor. Participants were randomly assigned to one of the six evidence order conditions specified in Table 1. In two of the conditions, participants received a mix of incriminating and exonerating evidence at both Time 2 and Time 3. Those two conditions differed in that one of them started with incriminating evidence as the first of two pieces of evidence at Time 2 (*Mixed 1*), whereas the other condition started with exonerating evidence (*Mixed 2*). Another two conditions consisted of contradicting evidence between Time 2 and Time 3. In those conditions, participants received either only incriminating (*Inc/Ex*) or only exonerating evidence (*Ex/Inc*) at Time 2, and then the contradicting evidence at Time 3. In the last two conditions, which served as control conditions, participants received only exonerating (*Ex/Ex*) or incriminating evidence (*Inc/Inc*) at both Time 2 and Time 3. All participants were asked to rate the measures outlined below after reading only the case description (Time 1), after receiving the first set of evidence (Time 2) and after receiving the second set of evidence (Time 3).

### Materials

Participants completed an experimental online survey administered using the Qualtrics survey platform. The survey consisted of three components: a case vignette, evidence and a set of measures (Appendix).

### Case description

Participants were given a vignette describing the early stages of a homicide investigation. The case was based on material previously used by Ask et al. (2008) and was adapted for use in the current study. The information in the vignette was intended to be ambiguous regarding any specific individual's guilt. The information in the vignette was limited mainly to what had happened and where, namely that a taxi driver had been shot in his car at a cul-de-sac. The vignette also mentioned that shoeprints had been found around the car, and that a pair of gloves was found near the crime scene. It mentioned that the police had arrested a suspect who had a history of violent behaviour, but no evidence incriminating the suspect for this crime was included in the vignette.

In order to test the hypotheses, the vignette should not be biased towards guilt or innocence. Hence, no strongly incriminating or exculpatory information was included in the vignette. By using a neutral vignette to introduce the crime scenario, it was possible to have a clear distinction between the material

that was necessary for the experimental manipulation. The vignette was pretested several times, and adapted according to the results each time. In the pretest of the final version of the case ( $N=35$ ), the average rating of the likelihood of the suspect being guilty was 47.34 ( $SD=19.56$ ) on a scale of 0 (*very unlikely*) to 100 (*very likely*).

### Evidence

All participants were presented with four pieces of evidence after reading the case vignette. There were four types of evidence: eyewitness identification, hair comparison, shoeprint comparison and CCTV footage. The evidence had been manipulated, such that each type of evidence existed in an incriminating version (e.g. the witness identified the suspect in a line-up) and an exculpatory version (e.g. the witness rejected the suspect in a line-up). The pieces of evidence were pretested extensively to ensure that they were perceived as approximately equally strong. In the final pretests ( $N=78$ ; with the exception of shoe prints which was added later,  $N=35$ ), the average strength ratings for the different pieces of evidence were all within 10 points from each other on a 100-point scale (for the detailed outcomes, see the [Supplemental Materials](#)).

### Measures

Participants were asked to fill out a similar set of measures at three points in time: after reading the case vignette (Time 1), after having received the first set of evidence (Time 2) and after having received the second set of evidence (Time 3). They were asked to rate the extent to which they felt uncomfortable, uneasy and bothered on scales ranging from 1 (*not at all*) to 7 (*very*; Elliot & Devine, 1994). The three ratings at each point in time were averaged to form composite measures of cognitive dissonance (Cronbach's  $\alpha = .87, .90$  and  $.91$  for Time 1, Time 2 and Time 3, respectively). Participants were then asked to

rate the likelihood that the suspect was guilty on a scale of 0 (*very unlikely*) to 100 (*very likely*). As court decisions on guilt in real life are dichotomous, participants were also asked whether they would convict the suspect (convict/acquit), and how confident they were about their decision (0 = *not at all confident*, 100 = *very confident*). At the end of the experiment, participants were shown a list of the evidence they had received at Time 2 and Time 3, and were asked to rate how important each piece of evidence was for their decision regarding guilt (0 = *not at all important*, 100 = *very important*).

### Procedure

Participants were provided with a link to the survey through advertisement posters, university communication or social media. The first page of the survey contained information about the study, and participants were asked to provide informed consent. Those who consented were then directed to the demographics section of the questionnaire. Next, participants were shown the case vignette. After reading the case, they completed the measures for the first time (Time 1). Participants were then presented with the first set of two pieces of evidence. The valence of the presented pieces of evidence (i.e. both incriminating, both exonerating or mixed) depended on the condition to which they had been randomly assigned. An overview of evidence combinations in each condition can be found in [Table 1](#). Participants then completed the measures for the second time (Time 2). Next, they were presented with a second set of evidence, the valence of which was again dependent on participants' condition (see [Table 1](#)). Finally, they filled out the measures for the third time (Time 3), and then proceeded to rating the importance of the individual pieces of evidence. Participants were then debriefed through the presentation of written information about the purpose of the study, and were asked to leave their contact details in case they wanted a chance to win a 10 EUR voucher.

Table 2. Means for likelihood of suspect guilt at different times across conditions.

Condition	Time 1	Time 2	Time 3
Mixed evidence			
Mixed 1	44.95 (21.32)	43.00 (22.59)	42.71 (24.12)
Mixed 2	47.63 (17.41)	54.79 (20.87)	54.55 (23.97)
Contradictory evidence			
Ex/Inc	47.97 (20.88)	26.96 (18.82)	52.50 (22.00)
Inc/Ex	46.38 (21.14)	71.16 (23.08)	42.26 (24.71)
Control conditions			
Ex/Ex	46.28 (21.63)	24.83 (18.48)	13.78 (16.03)
Inc/Inc	51.92 (19.69)	75.22 (16.59)	86.05 (12.10)

Note: Standard deviations in parentheses.

Completion of the survey took on average 20 min 35 s (*SD* = 75 min 30 s).<sup>2</sup>

**Results**

*Preliminary analyses*

Participants could complete the questionnaire in either Dutch or English. Language did not interact with the dependent variables (analyses can be found in the [Supplemental Material](#)). It was therefore decided that language was not a confounding variable, and it was not included as a covariate in subsequent analyses. The data for the English and Dutch groups were combined.

*Likelihood of guilt*

A 3 (time: Time 1 vs. Time 2 vs. Time 3) × 6 (condition: Mixed 1 vs. Mixed 2 vs. Ex/Inc vs. Inc/Ex vs. Ex/Ex vs. Inc/Inc) mixed analysis of variance (ANOVA), with time as repeated measures, was conducted on participants’ ratings of the likelihood of the suspect being guilty. The mean ratings for the different conditions at the different times can be found in [Table 2](#).

There was no significant main effect of time,  $F(1.80, 719.64)^3 = 1.82, p = .167, \eta_p^2 = .005$ , 90% confidence interval, CI [0.00,

0.02]. There was, however, a significant main effect of condition,  $F(5, 401) = 43.32, p < .001, \eta_p^2 = .351$ , 90% CI [0.28, 0.40] and a significant interaction effect between time and condition,  $F(8.97, 719.64) = 73.40, p < .001, \eta_p^2 = .478$ , 90% CI [0.43, 0.51]. The nature of the interaction effect is illustrated in [Figure 1](#). The interaction effect is broken down into focused comparisons below.

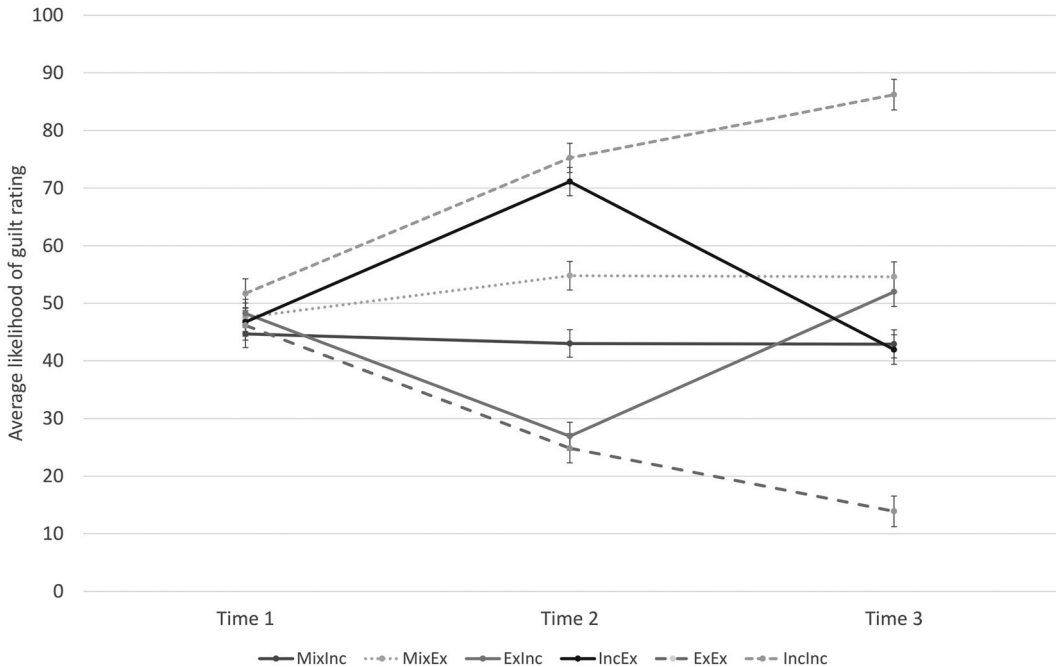
A linear trend analysis was conducted for the control groups (Ex/Ex and Inc/Inc), which had received uniformly incriminating or exonerating evidence in both the first and the second set of evidence. As would be expected, the linear trend was significant for both the Ex/Ex condition,  $F(1, 63) = 118.72, p < .001, \eta_p^2 = .653$ , 90% CI [0.53, 0.73] and the Inc/Inc condition,  $F(1, 63) = 194.27, p < .001, \eta_p^2 = .755$ , 90% CI [0.66, 0.81]. That is, participants made increasingly lower guilt rating in response to accumulating exonerating evidence and increasingly higher guilt ratings in response to accumulating incriminating evidence. This confirmed that the valence of the evidence was interpreted as intended.

Hypotheses 1 through 3 predicted that there would be an effect of the order in which participants viewed the evidence. Specifically, the likelihood of guilt at Time 3 for the condition that had first received incriminating evidence and then exonerating evidence (Inc/Ex)

<sup>2</sup>We suspect the large standard deviation can be explained by participants not closing the tab after they completed the survey.

<sup>3</sup>As the assumption of sphericity was not met, a Huyn–Feldt correction was applied, resulting in adjusted degrees of freedom.





**Figure 1.** Ratings for likelihood of suspect guilt. Error bars represent standard error.

would be higher than that for the conditions that had received mixed evidence at both times (H1). Conversely, the condition that first received exonerating evidence and then incriminating evidence (Ex/Inc) was expected to have a lower rating of likelihood of guilt at Time 3 than the mixed conditions (H2). We had also predicted that participants would show belief perseverance: those who were first presented with exonerating evidence in the first set and incriminating evidence in the second set (Inc/Ex) were expected to have a lower rating of likelihood of guilt at Time 3 than participants who received the evidence in the opposite order (Ex/Inc; H3).

A one-way ANOVA including the four experimental conditions was conducted on the likelihood of guilt at Time 3 to test these hypotheses.<sup>4</sup> There was a significant effect of

condition,  $F(3, 278) = 4.99, p = .002, \eta^2 = .051, 90\% \text{ CI } [0.01, 0.09]$ . A planned contrast showed that the Inc/Ex group did not make significantly higher ratings of likelihood of guilt than the mixed groups combined,  $t(275) = 1.83, p = .069, d = 0.26, 95\% \text{ CI } [-0.02, 0.56]$ , failing to support H1. A second planned contrast analysis showed that the Ex/Inc group did not make significantly lower ratings of likelihood of guilt than the mixed groups combined,  $t(275) = -1.03, p = .305, d = -0.15, 95\% \text{ CI } [0.00, 0.43]$ , failing to support H2. A third planned contrast, comparing the two contradictory conditions, showed that participants in the Ex/Inc condition reported a significantly higher final likelihood of the suspect being guilty than participants in the Inc/Ex condition,  $t(275) = -2.47, p = .014, d = -0.42, 95\% \text{ CI } [0.11, 0.75]$ . However, as we had predicted a difference in the opposite direction, H3 was not supported. Rather than belief perseverance, the latter finding is indicative of a recency effect (i.e. the evidence received last had the largest influence on the final guilt ratings).

<sup>4</sup>This analysis differs from that in the preregistration. We realised that the registered analyses were overly complex to test this specific hypothesis, so we chose to do a between-groups ANOVA at Time 3 only, and using only the experimental conditions.

Table 3. Means for cognitive dissonance at different times across conditions.

Condition	Time 1	Time 2	Time 3
Mixed evidence			
Mixed 1	2.34 (1.38)	2.69 (1.62)	2.65 (1.53)
Mixed 2	2.00 (1.11)	2.05 (1.16)	2.13 (1.23)
Contradictory evidence			
Ex/Inc	2.49 (1.50)	2.18 (1.27)	2.79 (1.61)
Inc/Ex	2.03 (1.11)	1.98 (1.00)	2.41 (1.40)
Control conditions			
Ex/Ex	2.24 (1.13)	1.96 (1.16)	1.85 (1.20)
Inc/Inc	2.09 (1.32)	1.96 (1.22)	1.91 (1.19)

Note: Standard deviations in parentheses.

### Cognitive dissonance

We also predicted that cognitive dissonance would differ between the conditions. Specifically, the contradictory conditions (Ex/Inc and Inc/Ex) were expected to report more cognitive dissonance at Time 3 than the mixed conditions (Mixed 1 and Mixed 2), which in turn were expected to report more cognitive dissonance than the control conditions (Ex/Ex and Inc/Inc). As the homogeneity of variance assumption was violated, a Kruskal–Wallis test was used.<sup>5</sup> The mean ratings for cognitive dissonance for the different conditions at all three times can be found in Table 3. A significant difference between the conditions was found,  $H(5) = 24.15$ ,  $p < .001$ ,  $\epsilon^2 = .081$ , 90% CI [0.070, 0.155].<sup>6</sup> A planned contrast analysis showed that the mixed conditions combined reported significantly more cognitive dissonance than the control conditions combined,  $t(260.78) = -3.25$ ,  $p = .001$ ,  $d = 0.40$ , 95% CI [0.17, 0.64], supporting H5. However, the mixed conditions combined did not differ significantly from the contradictory

conditions combined,  $t(268.42) = 1.18$ ,  $p = .239$ ,  $d = -0.14$ , 95% CI [-0.09, 0.37], failing to support H4.

### Exploratory analysis

#### Dissonance and belief perseverance

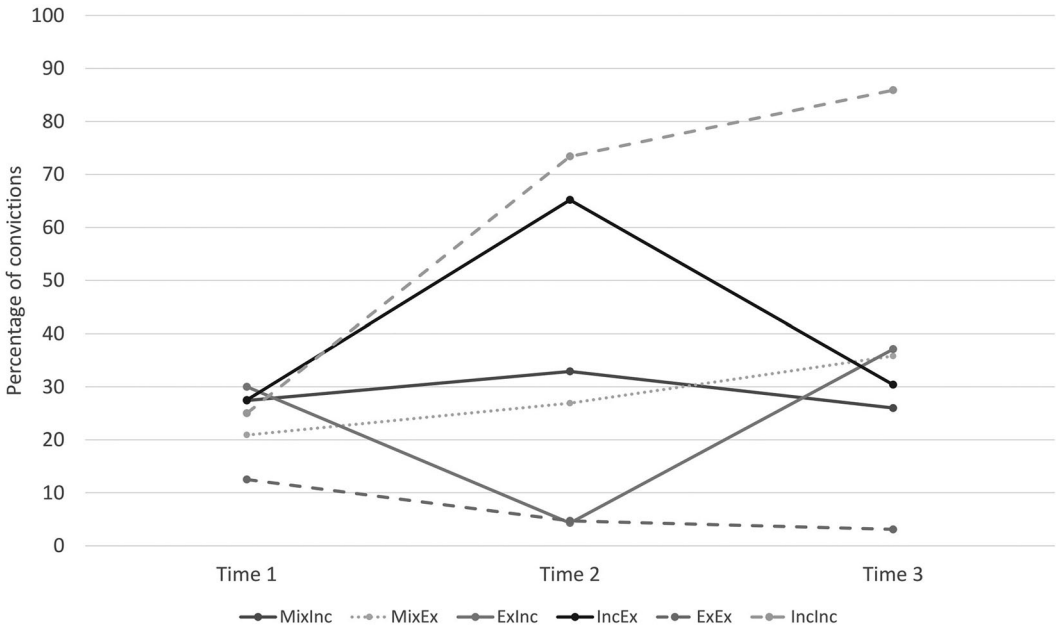
We explored whether the cognitive dissonance experienced by participants at Time 3 was correlated with their change in likelihood of guilt rating from Time 2 to Time 3 – that is, whether dissonance was a predictor of belief perseverance. The analysis is particularly relevant for the contradictory conditions, where belief perseverance would be indicated by a reluctance to change the ratings in response to the second set of (contradictory) evidence. Because the direction of change should logically differ between the conditions (upward in the Ex/Inc condition, downward in the Inc/Ex condition), they were analysed separately. The correlation was not significant in the Ex/Inc condition,  $r = .049$ ,  $p = .687$ , or in the Inc/Ex condition,  $r = -.014$ ,  $p = .909$ . The corresponding correlations in all conditions are reported in the Supplemental Materials.

#### Conviction rates

Participants were asked whether they would convict the suspect. An overview of the conviction rates can be seen in Figure 2. To explore whether the decision to convict after

<sup>5</sup>This test also differed from that in the preregistration for the same reasons as described earlier. We realised a one-way ANOVA of cognitive dissonance at Time 3 only would be a more appropriate test of the hypothesis. Due to the violated assumption of homogeneity of variance, a Kruskal–Wallis test was ultimately used.

<sup>6</sup>The confidence interval was computed using bootstrapping with 10,000 samples.



**Figure 2.** Percentage of convictions in all conditions at different times.

seeing all the evidence differed because of the order in which the evidence was presented, a binary regression analysis was conducted. As the control conditions did not receive the same evidence as the experimental conditions, they were not included in the analysis. The outcome variable was participants' decision on whether or not to convict the suspect at Time 3. Order of evidence presentation was not a significant predictor of the decision on whether to convict the suspect at Time 3,  $\chi^2(3) = 2.58, p = .461$ .

### Importance ratings

To explore whether the order in which the evidence was presented influenced the perceived importance of that evidence, the ratings for the pieces of evidence were analysed (see [Supplemental Material](#) for mean ratings for the evidence). As participants in the control conditions (Conditions 5 and 6) only received either incriminating or exonerating evidence, they were not included in the analysis. A  $2 \times 4$  mixed ANOVA was conducted, using the average importance ratings for both the incriminating and exonerating evidence as a

within-subjects factor, and using experimental condition as a between-subjects factor. There was a main effect of whether the evidence was incriminating or exonerating,  $F(1, 275) = 11.33, p = .001, \eta_p^2 = .040, 90\% \text{ CI } [0.01, 0.08]$ . The average rating of exonerating evidence ( $M = 59.78, SD = 22.11$ ) was higher than the average importance rating for incriminating evidence ( $M = 53.15, SD = 23.31$ ). However, there was no significant effect of condition,  $F(3, 275) = 0.41, p = .749, \eta_p^2 = .004, 90\% \text{ CI } [0.00, 0.01]$ , nor was there a significant interaction effect,  $F(3, 275) = 1.37, p = .252, \eta_p^2 = .015, 90\% \text{ CI } [0.00, 0.04]$ . In sum, the exonerating (vs. incriminating) evidence was considered to be more important, and this did not differ depending on the order in which the evidence was presented.

### Discussion

In the current study, we investigated whether the order in which incriminating and exonerating evidence is presented affects belief perseverance and cognitive dissonance. Contrary to

our expectations, participants did not show belief perseverance in their ratings of the likelihood of the suspect being guilty. Instead, the effect of order on ratings for likelihood of guilt appeared to be a recency effect. There was also no difference between the mixed evidence conditions and the contradictory evidence conditions for cognitive dissonance, again failing to support our hypotheses.

Participants in all conditions seemed to appropriately adjust their ratings of guilt likelihood in response to the evidence they were presented with. For the mixed conditions, likelihood of guilt stayed approximately neutral. For the contradicting conditions, participants' rating of likelihood of guilt increased when they were presented with incriminating evidence and decreased when they were presented with exonerating evidence. In contrast, other researchers have found that participants who expressed a belief in the suspect's guilt subsequently focused more disproportionately on the suspect than those who did not express a belief in guilt (O'Brien, 2009; Rassin et al., 2010). Our findings, however, seem to be in line with the findings by Price and Dahl (2014) and Charman et al. (2016), as the last piece of evidence received by the participants seemed to disproportionately affect their final rating of likelihood of guilt; those who saw incriminating evidence last had a higher final rating of likelihood of guilt than those who saw exonerating evidence last. Participants seemed to be unaware of the recency effect, as they did not rate the final pieces of evidence to be more important than previous pieces. Moreover, the evaluation of evidence did not seem to be affected by the order in which it was presented, as there was no difference in the importance ratings of the evidence between the different conditions.

The cognitive dissonance ratings also did not seem to be affected by the order in which the evidence was presented. Furthermore, we found no significant correlation between the cognitive dissonance reported by participants and the change in their rating of likelihood of

guilt. Favouring information that supports one's belief is a strategy to avoid or reduce cognitive dissonance (Festinger, 1957; Jonas et al., 2008). However, if participants did not experience cognitive dissonance in response to the contradicting information, they would also have no need to adhere to a specific belief in order to reduce cognitive dissonance. The generally low reported levels of cognitive dissonance are therefore in agreement with the lack of belief perseverance that we observed.

### *Recency effect*

Contrary to the theory of belief perseverance, but in line with previous research (Charman et al., 2016; Costabile & Klein, 2005), we observed a recency effect in the ratings of likelihood of guilt. The observed recency effect did not, however, appear to affect participants' decision to convict or acquit the suspect: there was no significant difference between the conditions for final conviction rates. As conviction rates were the main measure used by Costabile and Klein (2005), it is unclear why the recency effect in our study did not affect the conviction rates. Costabile and Klein (2005) suggested that the overall evaluation of information is delayed until a time when one is specifically asked for a decision, at which point the decision is made based on the evidence that is most readily accessible (i.e. the evidence received last). In our study, participants were asked for an overall evaluation of the information, namely whether they would convict the suspect, at three different times. In doing so, they integrated the evidence they had seen until that point. The integration might therefore have taken place in several steps rather than all at once, which could have limited the influence of the evidence that was most easily accessible. It could be argued that, despite the last evidence disproportionately affecting the likelihood of guilt rating, the effect was not strong enough to produce a change in participants' decision to convict or acquit the defendant. Such an interpretation would suggest that

the practical consequences of a recency effect may be limited.

Another explanation can be found in the study by Kerstholt and Jackson (1998). They compared the integration of evidence when participants were asked to give a probability rating of guilt after each witness statement as opposed to after reading all the evidence. They found a recency effect in the former condition. According to their explanation, the recency effect is due to an anchoring-and-adjustment process, whereby new information is evaluated in relation to the general impression of previous evidence, and more weight is attached to the last evidence in adjusting the anchor. Relating their explanation to the findings in the current study, it could be that participants in the contradictory-evidence conditions adjusted their rating towards the impression of the evidence presented at Time 3 and away from the opposite impression created at Time 2. Due to the considerable difference between the earlier impression and the latter impression, however, the latter evidence may not have created a sufficiently strong impression of guilt or innocence to affect conviction rates.

Both of the explanations offered above relate to the difference between the rating of the likelihood of the suspect being guilty and the conviction rates. The discrepancy between our findings and the existing literature (e.g. Charman et al., 2016) suggests that the relation between perceived likelihood of guilt and the absolute decision on guilt may be more complex than previously thought. If researched further, additional data on this relationship could add to the general literature on legal decision-making.

### ***Evidence evaluation***

Based on the changes in rated likelihood of guilt in response to the presentation of evidence, we can conclude that participants were not disproportionately influenced by a prior belief in their evaluation of the evidence. There was also no effect of condition on the importance ratings of the individual pieces of

evidence. That finding is not in line with the finding by Charman et al. (2016), who found that the initial piece of evidence viewed by participants affected their evaluation of the subsequent evidence. It is also not in line with the findings by Price and Dahl (2014), who found the recency effect extended to the evaluation of the individual pieces of evidence.

One unexpected finding was that participants rated the exonerating evidence as more important than the incriminating evidence. Marksteiner et al. (2011) found that only police trainees with a prior belief in the suspect's innocence rated incriminating and exonerating evidence as equally reliable, whereas those with a guilty hypothesis rated incriminating evidence as more reliable than exonerating evidence. Furthermore, the pretests of the evidence used in the current study also showed a slight guilt bias; the incriminating evidence was consistently rated as somewhat stronger than its otherwise identical exonerating equivalent. We therefore expected the incriminating evidence to be more important according to participants. It should be noted that the evaluation of the evidence in earlier research has often been done by asking participants to rate the reliability or strength of the evidence itself, whereas we asked them to rate the importance of the evidence for their decision whether to convict or acquit the defendant. As participants were asked about the importance of the evidence retroactively, the question related more to how they integrated the evidence to reach their decision. That may have limited the effect of either the last piece or first piece of evidence on the evaluation of the individual evidence. By examining the importance of the pieces of evidence that contributed to the overall integration of the evidence, namely the participant's decision on guilt, the influence of a single piece of evidence may have been prevented. That finding serves as a contribution of the current study to the existing literature, and provides a basis for further research of that potential effect. The fact that participants in the main study rated

exonerating evidence as more important might also indicate an innocence bias in the participants who took part. That could be due to the sampled population. When Ask and Granhag (2005) found that professional investigators were less receptive to a suggested alternative scenario than lay participants, one explanation for their finding was that professional investigators are already likely to consider alternative scenarios by default. Similarly to police officers, law students may also spontaneously consider different scenarios, or may even have a bias towards innocence. After all, law students are not only the judges of the future, but also the defence lawyers of the future. That could explain the higher importance ratings for the exonerating evidence, as well as provide an explanation as to why, despite the higher rating of likelihood guilt, participants who saw incriminating evidence last were no more likely to convict the suspect. In order to determine whether the proposed explanation affected the findings, the current study could be conducted with practising lawyers.

### ***Implications and future research***

The fact that we did not observe belief perseverance, and the fact that the recency effect did not affect conviction rates, leads to a tentative optimistic conclusion based on the data; the order in which evidence is presented may not strongly influence the occurrence of biased decision-making. However, further research in a more realistic context is necessary. Furthermore, it must also be remembered that treatment and consideration of cases are far more elaborate in real trials than in the current study, which likely also affects the evaluation of evidence. In addition, real-life case files are also likely to contain mainly incriminating information when a case is brought to trial (Crombag, 2017). As several researchers have previously found that the presentation of guilt-biased information influences the evaluation of later evidence (e.g. Ask et al., 2008, 2011; Greenspan & Scurich, 2016), a replication of

the current study using an incriminating case file may also provide further insights.

The mixed conditions in the current study differed from the majority of previous research. The mixed condition that started with the incriminating evidence can be argued to be more similar to trial proceedings in the Netherlands. Van der Post and Van Toor (2019) summarised the trial proceedings as following an incriminating–exonerating–incriminating–exonerating sequence. The trial starts with the prosecutor presenting the indictment and the questioning of the suspect about the evidence, followed by the defence lawyer’s plea. The prosecutor then responds to the lawyer, after which the lawyer responds again, and the suspect gets to speak last (Van der Post & Van Toor, 2019). According to the results of the current study, that order should not affect the perception of the evidence, as the mixed conditions did not differ from the contradicting conditions in their rating of likelihood of guilt. However, the complexity of a trial can hardly be compared to the procedure of the current study. It would therefore be beneficial to attempt to replicate the current findings in a trial setting with higher ecological validity.

There are also some differences between the procedure used and real-life proceedings which limits the application of our findings. For instance, Kerstholt and Jackson (1998) found that participants who were asked to judge the defendant’s guilt after seeing all the evidence showed a recency effect when background information was provided, but a primacy effect when no background information was provided. In our study, participants were not provided with background information about the suspect. In reality, judges in the Netherlands receive the case file and then have about a week to prepare for the trial (Van der Post & Van Toor, 2019). Such background information would, according to Kerstholt and Jackson (1998), contribute to a recency effect. Furthermore, the time delay between the presentation of evidence and having to make a

decision can also have an impact on possible order effects. More time between the different pieces of information that are being presented has been associated with an increased recency effect, whereas more time between the final presentation and the recall has been associated with a decreased recency effect (Insko, 1964; Stout et al., 2005). In the Netherlands, smaller criminal cases processed by a single judge will be decided immediately at the end of the trial. More serious or complicated cases with multiple judges will usually be discussed by the judges immediately after the trial or at the end of the day (Van der Post & Van Toor, 2019). A decision is then written, and the judgement is usually announced two weeks after the trial. Therefore, the current study was most similar to the trial by the single judge, as participants were asked to make a decision on their own immediately after seeing the evidence. Replicating the study with a longer time delay, and including background information, could therefore provide insight into possible order effects in a greater variation of proceedings.

There are a few limitations to the current study. Perhaps the biggest limitation is the use of law students instead of actual judges. Cognitive dissonance may be particularly likely to arise when the belief in question is considered important. It can be expected that judges are more invested in cases they work on than law students are in an experimental study. Furthermore, the training and experience of judges may cause them to respond differently than law students. That would be in line with the findings by Schmittat and Englich (2016), who found that criminal law experts showed less preference for confirming information in a criminal law case than did experts in other areas of law and laypeople. In addition to the need to determine whether law students and defence lawyers have a preference for alternative scenarios, the question of whether and how the consideration of clearly exonerating evidence differs between the different parties at trial also warrants investigation. Furthermore, based on the current study,

it also seems that findings on order effects among mock jurors (e.g. Costabile & Klein, 2005) may not generalise to populations that have received legal training.

A second possible limitation is the fact that the study was conducted online. While this sped up the data collection process and extended the achievable sample size, we cannot be sure how attentive participants were while taking the survey. We did, however, exclude participants who took the questionnaire very quickly in an attempt to filter out participants who rushed through the survey or who could not have read the material carefully. A third limitation of the study is that only one type of crime was used, namely a murder case. The case we used here has been used in multiple other studies (Ask et al., 2008, 2011; Marksteiner et al., 2011), although it was adapted slightly for the purpose of the current study. Other studies on order effects (e.g. Charman et al., 2016) have also used murder cases. It therefore seems unlikely that the choice of case can account for the unexpected results in the current study. Nevertheless, it may be beneficial for future research to include several types of crimes, although what is needed for the judge to become convinced of the suspect's guilt should not differ based on the type of crime.

A final limitation that should be considered is related to the measurement of cognitive dissonance. We expected that the simultaneous presentation of mixed evidence would not give rise to substantial cognitive dissonance. However, the mixed conditions experienced stronger dissonance than anticipated, resulting in dissonance not too different from that caused by the contradicting evidence in the other conditions. There is also an inherent difficulty to measuring dissonance. As it is an internal feeling, there is a risk that simply measuring dissonance could interfere with participants' experience of dissonance. Nevertheless, the dissonance measure has been used successfully in previous related research (Ask et al., 2011). Elliot and Devine

(1994) created the measure based on Festinger's (1957) description of the state of dissonance and on conceptually related research on affective responses to manipulations. Therefore, we considered this particular instrument to be the most valid available measure of cognitive dissonance.

## Conclusion

While the hypotheses of the current study were not supported, we observed a recency effect in the ratings of likelihood of guilt. The recency effect did not extend to the decision whether or not to convict the suspect, and also did not influence the perceived importance of the evidence. Thus, despite the final rating of likelihood of guilt shifting towards the last piece of evidence that participants saw, participants were not sufficiently influenced by the evidence to cause a difference in conviction rates. The recency effect therefore did not seem to undermine the impact of the evidence presented earlier. It would be meaningful to attempt to replicate the current findings using a more ecologically valid procedure and sample.

## Acknowledgements

This research is supported by a fellowship awarded from the Erasmus Mundus Joint Doctorate Program The House of Legal Psychology (EMJD-LP) with Framework Partnership Agreement (FPA) 2013-0036 and Specific Grant Agreement (SGA) 532473-EM-5-2017-1-NL-ERA MUNDUS-EPJD. We would like to thank Paulien van Rymenant for her help with data collection.

## Supplemental material

Supplemental material is available via the 'Supplementary' tab on the article's online page (<http://dx.doi.org/10.1080/13218719.2020.1855268>).

## Ethical standards

### *Declaration of conflicts of interest*

Enide Maegherman has declared no conflicts of interest

Karl Ask has declared no conflicts of interest

Robert Horselenberg has declared no conflicts of interest

Peter van Koppen has declared no conflicts of interest

### *Ethical approval*

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee at Maastricht University and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

### *Informed consent*

Informed consent was obtained from all individual participants included in the study

## ORCID

Enide Maegherman  <http://orcid.org/0000-0001-9872-0708>

Karl Ask  <http://orcid.org/0000-0001-5093-5902>

Peter J. van Koppen  <http://orcid.org/0000-0002-3127-7649>

## References

- Ask, K., & Granhag, P. A. (2005). Motivational sources of confirmation bias in criminal investigations: The need for cognitive closure. *Journal of Investigative Psychology and Offender Profiling*, 2(1), 43–63. doi:10.1002/jip.19
- Ask, K., Rebelius, A., & Granhag, P. A. (2008). The 'elasticity' of criminal evidence: A moderator of investigator bias. *Applied Cognitive Psychology*, 22(9), 1245–1259. doi:10.1002/acp.1432
- Ask, K., Reinhard, M. A., Marksteiner, T., & Granhag, P. A. (2011). Elasticity in evaluations of criminal evidence: Exploring the role of cognitive dissonance. *Legal and*



- Criminological Psychology*, 16(2), 289–306. doi:10.1348/135532510X510153
- Burke, A. (2007). Neutralizing cognitive bias: An invitation to prosecutors. *New York University Journal of Law Liberty*, 2, 512–530.
- Charman, S. D., Carbone, J., Kekessie, S., & Villalba, D. K. (2016). Evidence evaluation and evidence integration in legal decision-making: Order of evidence presentation as a moderator of context effects. *Applied Cognitive Psychology*, 30(2), 214–225. doi:10.1002/acp.3181
- Charman, S. D., Bradfield Douglass, A., & Mook, A. (2019). Cognitive bias in legal decision making. In N., Brewer, & A., Bradfield Douglass (Eds.), *Psychological Science and the Law* (pp. 30–53). The Guilford Press.
- Costabile, K. A., & Klein, S. B. (2005). Finishing strong: Recency effects in juror judgments. *Basic and Applied Social Psychology*, 27(1), 47–58. doi:10.1207/s15324834basp2701\_5
- Crombag, H. F. M. (2017). Invloeden op rechterlijke beslissingen [Influences on judicial decisions]. In P.J. van Koppen, J.W. de Keijser, R. Horselenberg, & M. Jelicic (Eds.), *Routes van het recht* [Routes of the law] (pp. 385–397). Boom Juridisch.
- Dror, I. E. (2020). Cognitive and human factors in expert decision making: Six fallacies and the eight sources of bias. *Analytical Chemistry*, 92(12), 7998–8004. doi:10.1021/acs.analchem.0c00704
- Dror, I. E., Charlton, D., & Péron, A. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International*, 156(1), 74–78. doi:10.1016/j.forsciint.2005.10.017
- Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology*, 67(3), 382–394. doi:10.1037/0022-3514.67.3.382
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. doi:10.3758/bf03193146
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.
- Findley, K. A., & Scott, M. S. (2006). Multiple dimensions of tunnel vision in criminal cases. *Wisconsin Law Review*, 2, 291–398.
- Green, M. C., & Donahue, J. K. (2011). Persistence of belief change in the face of deception: The effect of factual stories revealed to be false. *Media Psychology*, 14(3), 312–331. doi:10.1080/15213269.2011.598050
- Greenspan, R., & Scurich, N. (2016). The interdependence of perceived confession voluntariness and case evidence. *Law and Human Behavior*, 40(6), 650–659. doi:10.1037/lhb0000200
- Hamnett, H. J., & Jack, R. E. (2019). The use of contextual information in forensic toxicology: An international survey of toxicologists' experiences. *Science & Justice: Journal of the Forensic Science Society*, 59(4), 380–389. doi:10.1016/j.scijus.2019.02.004
- Harmon-Jones, E., & Harmon-Jones, C. (2007). Cognitive dissonance theory after 50 years of development. *Zeitschrift für Sozialpsychologie*, 38(1), 7–16. doi:10.1024/0044-3514.38.1.7
- Insko, C. A. (1964). Primacy versus recency in persuasion as a function of the timing of arguments and measures. *Journal of Abnormal Psychology*, 69(4), 381–391. doi:10.1037/h0042765
- Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology*, 80(4), 557–571. doi:10.1037/0022-3514.80.4.557
- Jonas, E., Traut-Mattausch, E., Frey, D., & Greenberg, J. (2008). The path or the goal? Decision vs. information focus in biased information seeking after preliminary decisions. *Journal of Experimental Social Psychology*, 44(4), 1180–1186. doi:10.1016/j.jesp.2008.02.009
- Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior*, 27(2), 187–203. doi:10.1023/A:1022599230598
- Kerstholt, J. H., & Jackson, J. L. (1998). Judicial decision making: Order of evidence presentation and availability of background information. *Applied Cognitive Psychology*, 12(5), 445–454. doi:10.1002/(SICI)1099-0720(199810)12:5<445::AID-ACP518>3.0.CO;2-8
- 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(11), 2098–2109. doi:10.1037/0022-3514.37.11.2098
- 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(4), 541–547. doi: 10.1002/acp.1719
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220. doi:10.1037/1089-2680.2.2.175
- O'Brien, B. (2009). Prime suspect: An examination of factors that aggravate and counteract confirmation bias in criminal investigations. *Psychology, Public Policy, and Law*, 15(4), 315–334. doi:10.1037/a0017881
- Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 521–533. doi:10.1037/0278-7393.14.3.521
- Price, H. L., & Dahl, L. C. (2014). Order and strength matter for evaluation of alibi and eyewitness evidence. *Applied Cognitive Psychology*, 28(2), 143–150. doi:10.1002/acp.2983
- Rassin, E., Eerland, A., & Kuijpers, I. (2010). Let's find the evidence: An analogue study of confirmation bias in criminal investigations. *Journal of Investigative Psychology and Offender Profiling*, 7(3), 231–246. doi: 10.1002/jip.126
- Schmittat, S. M., & Englich, B. (2016). If you judge, investigate! Responsibility reduces confirmatory information processing in legal experts. *Psychology, Public Policy, and Law*, 22(4), 386–400. doi:10.1037/law0000097
- Schünemann, B., & Bandilla, W. (1989). Perseverance in courtroom decisions. In H. Wegener, F. Lösel, & J. Haisch (Eds.), *Criminal behavior and the justice system: Psychological perspectives* (pp. 181–192). Springer.
- Simon, D. (2004). A third view of the black box: Cognitive coherence in legal decision making. *The University of Chicago Law Review*, 71(2), 511–586.
- Stout, S. C., Amundson, J. C., & Miller, R. R. (2005). Trial order and retention interval in human predictive judgment. *Memory & Cognition*, 33(8), 1368–1376. doi:10.3758/bf03193369
- Van der Post, N., & Van Toor, D. (2019). Volgorde-effecten in het Nederlandse strafproces? Bayesiaanse methode als hulpmiddel tegen de invloed van het primacy en het recency effect op rechterlijke bewijswaardering. *Expertise en Recht*, 11(1), 41–54.
- Wagenaar, W. A., van Koppen, P. J., & Crombag, H. F. M. (1993). *Anchored narratives: The psychology of criminal evidence*. Harvester Wheatsheaf.

## Appendix

### Case vignette

#### *The event*

At 02:21 on a Sunday morning, the emergency central receives a call from a woman (Witness A) who reports having seen a taxi driver probably being shot outside her apartment. A police patrol and an ambulance are sent to the given address. In the taxi, which is parked in a lit up turning space, the police find a dead man in the driver's seat, with a shot wound in the head. Reinforcements are requested and the scene is cordoned off.

Witness A, who called the emergency central, lives on the first floor with her bedroom in the end of the house, facing the turning space. At 02:15 she was woken up by an aggressive male voice coming from the turning space, but she could not hear what was being said. After a minute or so she heard a bang, after which she walked to the window to see what had happened. At the turning space she saw the taxi, and since the interior lighting of the taxi was on she could see the driver hanging over the wheel. Then she called 112. In the outskirts of the residential area, a police officer on the way to the scene noticed a man running. The man was arrested and brought in for questioning.

#### *Investigation results*

According to the taxi company, a customer called in and ordered a taxi to the turning space at the specified time. The customer, however, has a reliable alibi and has no involvement in the crime.

At the crime scene, the police found shoe-prints around the car.

The post-mortem examination showed that the victim had died as a result of a shot to the head. The bullet had entered slightly above

the victim's right eye and had been fired from a very short distance (<1 m). Both the fired bullet and a 9 mm cartridge were found inside the taxi.

The police's canine squad also found a pair of leather gloves and a 9 mm pistol of the make Zastava in a grove near the place where the suspect was arrested.

The results from the National Laboratory of Forensic Science showed that the bullet and the cartridge had been fired from the discovered gun, and that the leather gloves showed traces of powder stain.

A check of the taxi's daily receipts showed that €110 in cash was missing from the car. It also appeared as if a GPS and a mobile phone were missing.

#### *The suspect*

The suspect, who had been seen running in the outskirts of the residential area, was identified using his driver's license.

Records showed that he had been previously convicted of aggravated assault and illegal carrying of a knife. In the first police interview, the suspect claimed that he was running to get away from three men with whom he had been in a fight earlier that night.

When asked to explain which way he had been running, the man answered that he had passed the turning space and that he recalled a taxi being parked there. After that, the suspect refused to answer any further questions.

The suspect lives in the vicinity of the crime scene.

The man is 27 years old, 187 cm tall, and weighs 79 kg. At the time of the arrest he was dressed in dark-green, baggy trousers, a black leather jacket, and a light-gray hooded sweater.

**Additional evidence**

Type of evidence	Incriminating	Exonerating
Eyewitness	The police have found an eyewitness who saw someone leaving the taxi area around the time of the crime. The witness was presented with a line-up at the police station the following day and identified the suspect as the perpetrator. The witness was confident in his decision.	The police have found an eyewitness who saw someone leaving the taxi area around the time of the crime. The witness was presented with a line-up containing the suspect at the police station the following day. The witness was confident that the perpetrator was not present in the line-up.
Hair	The hair found in the glove near the crime scene was sent to the National Laboratory of Forensic Science for close comparison to the hair of the suspect. The lead scientist on the case has declared that the hair from the crime scene very probably originated from the suspect (85% certainty).	The hair found in the glove near the crime scene was sent to the National Laboratory of Forensic Science for close comparison to the hair of the suspect. The lead scientist on the case has declared that the hair from the crime scene very probably did not originate from the suspect (85% certainty).
CCTV	After obtaining a warrant the police are allowed to inspect the CCTV footage from the turning space. Close observation of the footage shows that someone who matches the physical appearance and clothing of the suspect was at the turning space around the time of the crime.	After obtaining a warrant the police are allowed to inspect the CCTV footage from a bar a few blocks away from the crime scene. Close observation of the footage shows that someone who matches the physical appearance and clothing of the suspect was at the bar at the time of the crime.
Shoeprints	The suspect's shoes were investigated to see whether they matched the shoeprints found around the taxi. The shoeprints matched the shoes the suspect was wearing when he was arrested.	The suspect's shoes were investigated to see whether they matched the shoeprints found around the taxi. The shoeprints did not match the shoes the suspect was wearing when he was arrested.