

# Speak Last and Step-by-Step: The Effect of Order and Response Mode on Evidence Evaluation

Mengxuan Helen Qiao<sup>1</sup> (helen.qiao.18@ucl.ac.uk)

David Lagnado<sup>1</sup> (d.lagnado@ucl.ac.uk)

<sup>1</sup>Department of Experimental Psychology, University College London,  
26 Bedford Way, WC1H 0AP, London, UK

## Abstract

How does evidence presentation order and response mode affect legal judgments? Previous research on order effects in legal decision-making has produced mixed results, possibly due to different response modes adopted in the tasks: End-of-Sequence (EoS) or Step-by-Step (SbS), which might reflect different cognitive models that fact-finders employ during evidence evaluation and integration. This paper investigates how response mode interacts with evidence order to influence judgments of the probability of guilt and verdict. In Study 1 ( $N = 159$ ), no order effects were found in the EoS condition; but a recency effect emerged in the SbS condition. Study 2 ( $N = 95$ ) revealed no order effect when the first set of evidence was judged SbS and the second set EoS. We also found that participants' probability of guilt judgments were generally consistent with Bayesian predictions based on their prior beliefs and perceived strength of new evidence when they responded SbS, but not when they responded EoS. We discuss potential explanations for these findings and their implications for legal decision-making.

**Keywords:** legal decision-making; order effect; response mode; belief updating; Bayesian reasoning

## Introduction

In the UK, the Criminal Procedures Rule – Rule 25.9(2)(k) specifies that ‘the defendant may make final representations’<sup>1</sup> before the verdict is reached, allowing the defence a final opportunity to influence the court’s judgment. In the US, however, closing arguments typically follow a different sequence: the government argues, the defence responds, and then the government presents a final rebuttal argument<sup>2</sup>.

From a normative perspective, fact-finders should form similar conclusions regardless of the order of incriminating and exonerating evidence. Yet extensive psychological research suggests that this ideal is rarely achieved (e.g., Ask et al., 2008; Charman et al., 2016; Higgins, 1996; Imhoff & Nickolaus, 2021). One key factor found to influence the order effect is the response mode—whether judgments are made holistically after considering all evidence (End-of-Sequence [EoS]) or updated incrementally with each piece of new evidence (Step-by-Step [SbS]). Although Hogarth and Einhorn (1992) predicted and observed recency effects in complex tasks regardless of response mode, fact-finders might adopt different cognitive models when responses are elicited by different approaches. The story model describes a process that a coherent narrative explanation is constructed when asked for

an EoS judgment (N. Pennington & Hastie, 1992), while the belief-adjustment model describes a process of sequential adjustments to previous beliefs by the impact of new evidence (Hogarth & Einhorn, 1992).

In this paper, we present two studies that explore how response mode interacts with evidence presentation order to influence legal judgments. Although not designed as a test between these two cognitive models, our study allows us to explore the legal judgment process in greater depth.

## Order Effects

When individuals evaluate a sequence of information, they adjust their judgments based on newly acquired information. If the adjustment is insufficient, previous information disproportionately influences the decision, leading to a *primacy effect*; conversely, if the adjustment is excessive, newer information has a greater impact, resulting in a *recency effect* (Keltz & Adelman, 2008).

In the legal context, some early research identified a *primacy effect*. For example, D. C. Pennington (1982) found that participants were more likely to deliver guilty verdicts when witnesses and testimony indicating guilty were presented first. Tetlock (1983) also observed a primacy effect when evidence suggestive of guilt was first presented. Moreover, Wells et al. (1985) showed that if the defence presented its case early on, participants were more likely to find the defendant not guilty.

In contrast, other studies have identified a *recency effect*. For example, Anderson (1959) observed a recency effect when participants made successive judgments based on prosecution and defence evidence presented in different orders. The same finding was replicated by Furnham (1986) and with various types of evidence, such as alibi and eyewitness evidence (Dahl et al., 2009; Enescu & Kuhn, 2012; Price & Dahl, 2014), DNA and alibi evidence (Charman et al., 2016), and other types of forensic evidence, such as hair, CCTV, and shoe prints (Maegherman et al., 2022). Costabile and Klein (2005) showed a recency effect of critical evidence, in which participants who were exposed to incriminating critical evidence by the last witness rendered more guilty verdicts than those who were exposed to the same piece of evidence by the first witness.

## Response Mode

Response mode, which corresponds to the manner in which judgments are elicited, has been identified as one of the vari-

<sup>1</sup><https://www.legislation.gov.uk/uksi/2020/759/part/25>

<sup>2</sup><https://www.uscourts.gov/rules-policies/current-rules-practice-procedure/federal-rules-criminal-procedure>

ables that influence order effects (Hogarth & Einhorn, 1992). There are two primary response modes: EoS requires a global assessment of all evidence to reach a single final judgment, while SbS involves adjusting one's opinion incrementally with each piece of evidence in a sequence. Hogarth and Einhorn (1992) suggest that in short and simple tasks, the EoS response mode induces a primacy effect and the SbS response mode induces a recency effect. The same pattern is also predicted by Kashima and Kerekes (1994)'s distributed memory model. However, Hogarth and Einhorn (1992) further suggest that recency effects are associated with complex tasks, such as legal judgment tasks, regardless of the response mode. They explain that the process of aggregating multiple items and reaching an EoS judgment demands more from memory and cognitive resources. To save on cognitive load, individuals rely on the most readily accessible and temporally salient information in memory—typically the most recently encountered—thus, people shift to an SbS approach when making an EoS judgment, resulting in a recency effect. There is also empirical evidence supporting that recency effects are detected in EoS responses (e.g., Charman et al., 2016; Costabile & Klein, 2005).

On the other hand, EoS and SbS response modes might reflect different cognitive models that individuals adopt when making legal judgments. During the SbS response mode, individuals adopt the belief-adjustment model where they form an initial judgment and adjust it incrementally based on new evidence (Hogarth & Einhorn, 1992; Kerstholt & Jackson, 1998). By contrast, during the EoS response mode, individuals are likely to adopt the story model, which allows individuals to evaluate and integrate all available evidence to construct a complete, coherent narrative explanation (N. Pennington & Hastie, 1988). The story model posits that fact-finders focus on the causal connections between events within a narrative. Because this process emphasises coherence over sequence, the order in which evidence is presented should have little effect on judgment outcomes. It is worth noting that N. Pennington and Hastie (1992) found the order of evidence presentation did influence jurors' decisions—but specifically when evidence was arranged in a story-consistent order or an issue-based order, to support or disrupt the construction of a coherent causal narrative, respectively. This differs from the present study, which focuses on the presentation order of evidence sets (i.e., prosecution followed by defence vs. defence followed by prosecution), rather than the internal coherence of the narrative structure.

## Overview of Studies

Given the mixed results and inconclusive explanations of how legal judgment outcomes are influenced by response mode and its interaction with evidence order, the current research aimed to directly compare these variables in a legal decision-making task. To our knowledge, this is the first to adopt both EoS and SbS response modes in a single judgment task.

In our studies, we presented a summary of a criminal case to the participants with a set of prosecution evidence and a

set of defence evidence, where the order of sets and response modes were manipulated. We investigated how participants' judgments of the probability of guilt and verdict were influenced by these variables, as well as the Bayesian predictions calculated from their prior beliefs of guilt and the perceived strength of new evidence. N. Pennington and Hastie (1992) found that evidence integration, even in the SbS response mode, often deviated from Bayesian expectations. More recent research suggests that fact-finders' judgments are generally consistent with Bayesian expectations (e.g., Shengelia & Lagnado, 2021; Thompson et al., 2013). Thus, we aimed to determine whether participants' probability of guilt judgments aligned with the perceived strength of the new evidence and reflected Bayesian updating under different response modes.

In Study 1, we compared the judgments between an all EoS context and an all SbS context. In Study 2, we looked at judgments in which the first set of evidence was judged SbS and the second set EoS. All materials, data, and analyses are available at the OSF repository (<https://osf.io/u72fr/>).

## Study 1

In Study 1, we investigated how the order of defence and prosecution evidence, along with the response mode, influenced participants' judgments of the perceived probability of guilt and verdict. We asked participants to make judgments under two conditions: they either provided EoS judgments after viewing the whole sets of evidence (EoS condition) or made sequential judgments after viewing each piece of evidence (SbS condition). We expected a recency effect in the SbS condition, as predicted by the belief-adjustment model. In the EoS condition, if participants adapted to an SbS approach as suggested by Hogarth and Einhorn (1992), a recency effect would be anticipated; on the other hand, if participants adopted the story model to construct a holistic explanation, no order effect should occur.

In addition, we investigated whether participants' probability of guilt judgments aligned with Bayesian predictions for belief updating based on their conditional probability assessments of the new evidence and prior beliefs of guilt. This assessment corresponds to the quantification of the strength of evidence as measured by the likelihood ratio, which is the standard measure in legal and forensic contexts (Good, 1989). Notably, a judgment stage between evidence sets in the EoS condition was implemented to provide a 'prior' for the Bayesian calculation. If participants treated each evidence set as a distinct 'item', a recency effect could still arise in the EoS condition, as observed by Hogarth and Einhorn (1992).

## Methods

**Participants** We recruited 159 participants ( $M_{\text{age}} = 38.3$ ,  $SD_{\text{age}} = 13.3$ ,  $N_{\text{Female}} = 79$ ) from Prolific Academic; 78 participants were assigned to the EoS condition and 81 to the SbS condition. Participants were pre-screened for fluency in English and a platform approval rate between 95-100%. All

participants provided informed consent and were reimbursed £1.50 for completing the 9-minute experiment.

**Design and Materials** The study employed a 2 (order: between-subject)  $\times$  2 (response mode: between-subject)  $\times$  2 (stage: within-subject [EoS condition]) or 8 (stage: within-subject [SbS condition]) mixed design. We constructed a case summary and evidence that was loosely based on the Sally Clark case<sup>3</sup>. The case summary included only objective and factual information. The defence and prosecution evidence sets comprised four pieces of evidence each. Participants either received the defence evidence set followed by the prosecution evidence set (DP order), or the prosecution evidence set followed by the defence evidence set (PD order). The order of evidence within each set was consistent.

We measured participants' perceived probability of guilt of the defendant on a scale from 0% (not likely at all) to 100% (extremely likely) by asking the question 'Based on the given evidence, what do you think is the probability that the defendant is guilty?', and verdict choices of 'Not guilty' and 'Guilty' by asking the question 'Please state your verdict for the defendant.'. Two conditional probability questions were asked at each judgment stage; examples of the conditional probability questions for one piece of prosecution evidence are:

- 'If the defendant is NOT GUILTY, how likely is it that the hypoxic damage to Ben's brain was caused a matter of hours before death?'
- 'If the defendant is GUILTY, how likely is it that the hypoxic damage to Ben's brain was caused a matter of hours before death?'

All conditional probability questions were answered on a scale from 0% (not likely at all) to 100% (extremely likely).

**Procedure** Participants accessed the experiment on Qualtrics ([www.qualtrics.com](http://www.qualtrics.com)). After receiving instructions, participants read the case summary and the sets of defence and prosecution evidence in either the DP or PD order. They were randomly assigned to either the EoS condition or the SbS condition. In the EoS condition, participants provided judgments after viewing each set of evidence. In the SbS condition, participants received the evidence piece by piece and gave judgments after each piece. Participants answered questions about the probability of guilt, verdict, and conditional probabilities of evidence at every judgment stage.

## Results

**Main analysis** We compared participants' probability of guilt ratings in the EoS condition with responses given at stage 4 (end of the first set of evidence) and 8 (end of the second set of evidence) in the SbS condition using a 2 (order)  $\times$  2 (stage)  $\times$  2 (response mode) mixed ANOVA. We found a

main effect of order,  $F(1, 155) = 21.66, p < .001$ , an interaction between order and stage,  $F(1, 155) = 184.45, p < .001$ , and an interaction between order, stage, and response mode,  $F(1, 155) = 8.03, p = .005$ . However, there was no main effect of stage,  $F(1, 155) = 0.05, p = .832$ , or response mode,  $F(1, 155) = 1.37, p = .245$ .

Pairwise comparisons revealed a significant difference in the final probability of guilt ratings between the DP ( $M = 0.67$ ) and PD orders ( $M = 0.48$ ) in the SbS condition,  $t(155) = 3.63, p = .009$ , while no difference was found in the EoS condition between orders,  $t(155) = 0.93, p = .983$  (see Figure 1).

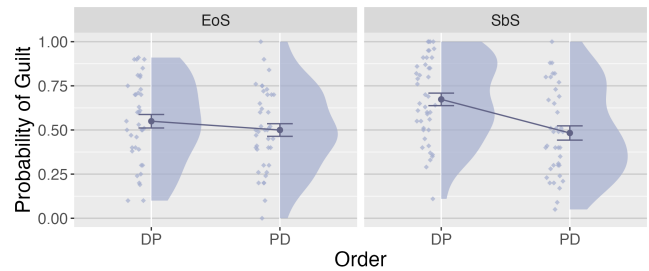


Figure 1: Study 1: Final probability of guilt judgments in DP and PD orders in EoS and SbS conditions.

Further Chi-squared tests showed a significant association between order and final verdict in the SbS condition,  $\chi^2(1, N = 81) = 6.46, p = .011$ ; but no association in the EoS condition,  $\chi^2(1, N = 78) = 0.82, p = .365$  (see Figure 2).

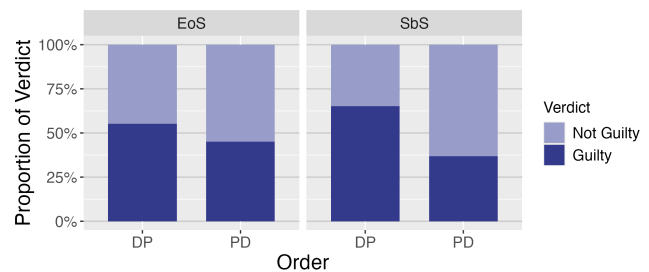


Figure 2: Study 1: Proportion of final verdicts in DP and PD orders in EoS and SbS conditions.

**Observed vs. predicted probabilities** To calculate the individual Bayesian belief updating models, we applied Bayes' rule to predict each participant's posterior probability  $P(H|E)$ , based on their stated conditional probabilities,  $P(E|H)$ ,  $P(E|\neg H)$ , and their prior  $P(H)$  from the previous time-step. In cases where participants gave extreme responses of '0's and '1's for the conditional probability questions, we converted those to '0.001's and '0.999's, respectively, to avoid dividing by zero.

We conducted a 2 (order)  $\times$  2 (judgment type [observed vs. predicted]: within-subject) mixed ANOVA for the EoS

<sup>3</sup><http://www.bailii.org/ew/cases/EWCA/Crim/2003/1020.html>

condition. Stage 1 was not included in this analysis because there were no priors for estimations for stage 1 judgments. We found a main effect of judgment type,  $F(1,76) = 6.77$ ,  $p = .011$ , and an interaction between order and judgment type,  $F(1,76) = 4.32$ ,  $p = .041$ . However, we did not find a main effect of order,  $F(1,76) = 0.04$ ,  $p = .950$ . Pairwise comparisons revealed a significant difference between the observed ( $M = 0.50$ ) and predicted ( $M = 0.63$ ) probabilities in the PD order,  $t(76) = 3.35$ ,  $p = .007$ , but not in the DP order,  $t(76) = 0.37$ ,  $p = .983$  (see Figure 3).

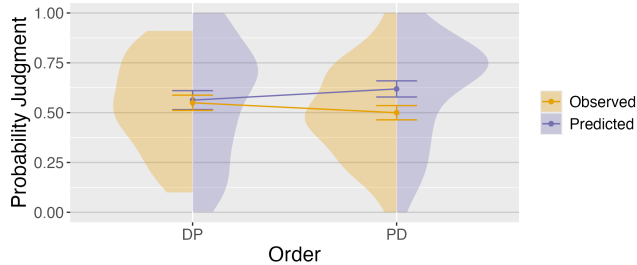


Figure 3: Study 1: Observed and predicted probabilities in the final stage in DP and PD orders in EoS condition.

Further, we conducted a 2 (order)  $\times$  7 (stage)  $\times$  2 (judgment type) mixed ANOVA for the SbS condition. Again, Stage 1 was not included in this analysis. We found a main effect of order,  $F(1,79) = 5.84$ ,  $p = .018$ , a main effect of judgment type,  $F(1,79) = 8.59$ ,  $p = .004$ , an interaction between order and stage,  $F(3.77,298.17) = 55.40$ ,  $p < .001$ , and an interaction between order, stage, and judgment type,  $F(5.80,458.31) = 5.25$ ,  $p < .001^4$ . Pairwise comparisons revealed no significant differences between the observed and predicted values within the same stages in either order (see Figure 4).

## Discussion

In Study 1, we found that the order in which defence and prosecution evidence was presented did not influence participants' judgments of the probability of guilt or verdict in the EoS condition, but a recency effect emerged in the SbS condition. These results suggest that participants did not switch to an SbS response mode in the EoS condition, contradicting Hogarth and Einhorn (1992)'s prediction. Participants might have employed the story model in the EoS condition, evaluating and integrating evidence holistically to form a single, conclusive judgment. The additional judgment stage between evidence sets also did not disrupt the construction of an explanation that informed their final decision.

Notably, in the EoS condition, participants' probability of guilt judgments were relatively in line with Bayesian predic-

<sup>4</sup>Mauchly's test indicated that the assumption of sphericity for the interaction term of 'order' and 'stage' ( $p < .001$ ) and the interaction term of 'order', 'stage', and 'judgment type' ( $p < .001$ ) had been violated, and therefore degrees of freedom were corrected using Greenhouse-Geisser ( $\epsilon = 0.63$ ) and Huynh-Feldt ( $\epsilon = 0.97$ ) estimates of sphericity, respectively.

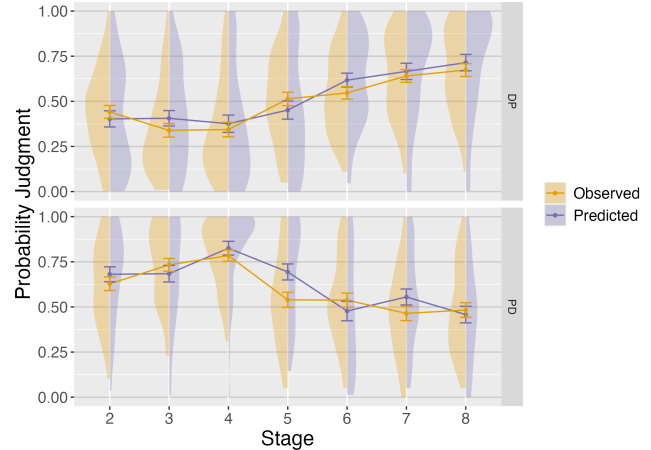


Figure 4: Study 1: Observed and predicted probabilities in DP and PD orders in SbS condition.

tions in the DP order. However, the observed guilt ratings in the PD order were lower than Bayesian predictions, suggesting that after initially considering the prosecution evidence, participants revised their perception of guilt more strongly in favour of innocence than predicted following the later presentation of defence evidence.

While observing a significant main effect of judgment type in the SbS condition, pairwise comparisons showed no significant differences between the observed probability of guilt judgments and Bayesian predictions in any of the stages in either order. We, therefore, propose that participants generally adhered to Bayesian predictions relative to their own prior beliefs of guilt and evaluations of the strength of new evidence when presented with items of evidence SbS. This result is consistent with previous findings (e.g., Shengelia & Lagnado, 2021); it also confirms Schum and Martin (1982)'s claim that people tend to assess evidence accurately on an item-by-item basis rather than integrating it into a single comprehensive judgment.

## Study 2

In Study 1, we observed a recency effect when participants judged evidence in an SbS response mode. In Study 2, we aimed to further confirm whether such an effect was caused by the SbS processing approach of recent evidence. Therefore, in this study, we presented the first set of evidence in an SbS response mode and the second set in an EoS response mode. If later received evidence has more influence on participants' belief only because it was processed SbS, the recency effect would disappear.

Additionally, in Study 1, we found that participants' probability of guilt judgments aligned with Bayesian predictions in the DP order but not in the PD order in the EoS condition, while they were generally consistent in SbS judgments for both orders. In Study 2, we aimed to further assess whether the Bayesian alignment held in a mixed SbS + EoS condition.

## Methods

**Participants** We recruited 95 participants from Prolific Academic, pre-screened for fluency in English and had a platform approval rate of 95-100%. Individuals who have participated in the previous experiment were excluded. Two participants failed the attention check questions, leaving 93 participants ( $M_{\text{age}} = 36.9$ ,  $SD_{\text{age}} = 11.6$ ,  $N_{\text{Female}} = 48$ ) for the analyses. 45 participants were assigned to the DP order and 48 were assigned to the PD order. All participants provided informed consent and were reimbursed £1.50 for the 9-minute experiment.

**Design and Materials** We employed a 2 (order: between-subject)  $\times$  5 (stage: within-subject) mixed design. Stage 1-4 were SbS judgments of pieces of evidence in the first set and stage 5 was an EoS judgment of the second evidence set. The same materials and measurements used in Study 1 were adopted.

**Procedure** After reading the instructions and case summary, participants were randomly assigned to either the DP or PD order. In the DP order, participants viewed the defence evidence piece by piece and provided judgments after each piece; followed by the prosecution evidence presented all at once, after which they gave a final judgement. In the PD order, participants viewed the prosecution evidence piece by piece and provided judgments after each piece, followed by the defence evidence presented all at once and provided a final judgment.

## Results

**Main analysis** A two-sample t-test showed no difference in the final probability of guilt ratings between the DP ( $M = 60.7$ ) and PD ( $M = 55.4$ ) orders,  $t(91) = 1.09$ ,  $p = .280$  (see Figure 5). Further, a Chi-squared test showed no association between order and final verdicts,  $\chi^2(1, N = 93) = 0.57$ ,  $p = .452$  (see Figure 6).

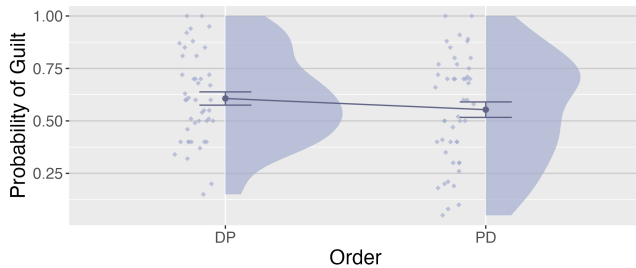


Figure 5: Study 2: Final probability of guilt judgments in the DP and PD orders.

**Observed vs. predicted probabilities** As in Study 1, we applied Bayes' rule to calculate each participant's predicted posterior probability values using their prior beliefs and likelihood judgments of evidence.

We conducted a 2 (order)  $\times$  4 (stage)  $\times$  2 (judgment type)

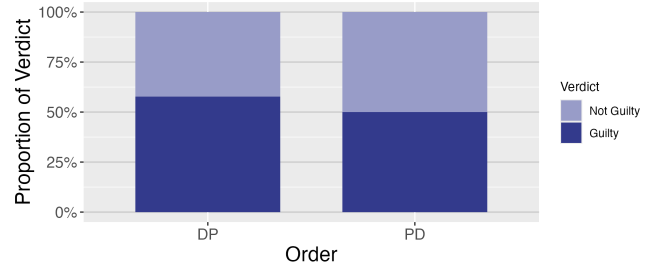


Figure 6: Study 2: Proportion of final verdicts in the DP and PD orders.

mixed ANOVA. The first stage was not included because there were no priors for stage 1 to calculate the Bayesian predictions. We found a main effect of order,  $F(1, 91) = 40.31$ ,  $p < .001$ , an interaction between order and judgment type,  $F(1, 91) = 10.39$ ,  $p = .002$ , an interaction between order and stage,  $F(2.64, 240.12) = 37.93$ ,  $p < .001^5$ , and an interaction between order, stage, and judgment type,  $F(3, 273) = 11.01$ ,  $p < .001$ . Pairwise comparisons revealed a significant difference between observed ( $M = 0.55$ ) and predicted values ( $M = 0.70$ ) in the final stage in the PD order,  $t(91) = 4.50$ ,  $p = .002$ . While there was a descriptive difference between the observed ( $M = 0.61$ ) and predicted judgments ( $M = 0.49$ ) in the final stage in the DP order, the difference was not statistically significant,  $t(91) = 3.42$ ,  $p = .067$ . No significant differences between the observed and predicted values were detected in stages 2 to 4 in either order, where participants made SbS judgments (see Figure 7).

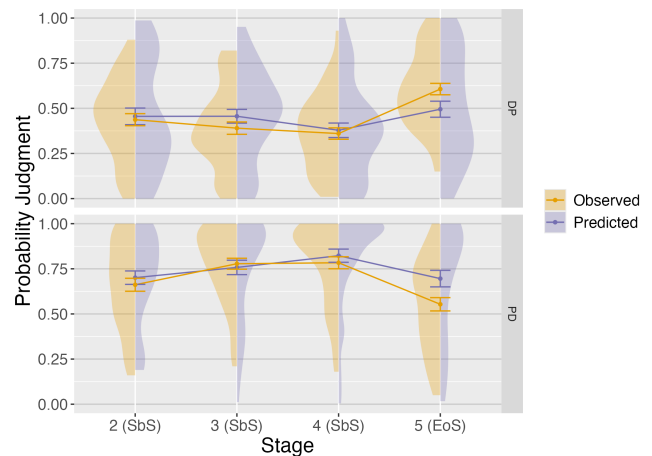


Figure 7: Study 2: Observed probability of guilt ratings and predicted probabilities in DP and PD orders.

<sup>5</sup>Mauchly's test indicated that the assumption of sphericity for the interaction term of order and stage had been violated ( $p < .001$ ), and therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ( $\epsilon = 0.88$ ).

## Discussion

In Study 2, we found that the evidence presentation order did not affect participants' final judgments of the probability of guilt and verdict when the first set of evidence was evaluated SbS and the second set EoS. This result suggests that the recency effect we observed in the previous study was not merely caused by the presentation order of evidence, but by the SbS processing approach of the recent evidence.

Participants' probability of guilt judgments aligned with Bayesian predictions during stages 2 to 4, where they responded SbS. However, after viewing the final evidence set and responded EoS, participants' judgments diverged from Bayesian predictions: they revised their perception of guilt of the defendant significantly lower than the prediction after considering the later presented defence evidence; and they revised their perception of guilt higher than the prediction after prosecution evidence was later presented, although this difference was not statistically significant. This pattern is similar to what we observed in Study 1 in which participants revised the perceived probability of the defendant's guilt downward—shifting toward greater innocence than predicted—after viewing the defence evidence presented later.

### General Discussion

In two studies, we explored the effects of evidence presentation order and response mode on people's judgments in a legal decision-making task. We found that a recency effect emerged only when the last evidence set was processed in an SbS response mode. No order effects appeared when all evidence was processed in an EoS response mode, despite there being an additional judgment stage between the evidence sets. Moreover, no order effects were observed when the initial evidence set was processed SbS and the latter EoS.

Our results suggest that participants integrated evidence holistically in an EoS context and reached judgments that were insensitive to the evidence presentation order. This finding challenges Hogarth and Einhorn (1992)'s claim, which posits that people automatically shift to an SbS response mode when asked to make an EoS judgment in a complex task. Instead, participants may have applied different cognitive approaches depending on whether they evaluated evidence SbS or EoS. Notably, Hogarth and Einhorn (1992) classified simple tasks as those involving only a single item for each piece of evidence with a reasonably assumed task familiarity, while complex tasks involve a large amount of information and/or unfamiliar stimuli. The task in our experiments fits into their classification of a complex task, however, it is possible that our task was not complex enough to tax participants' cognitive load for them to switch the evaluation approach. Future studies are needed to explore how task complexity impacts the cognitive approach people adopt when evaluating evidence and making legal judgments.

The recency effect we observed were reflected both in the probability of guilt judgment and verdicts, a pattern distinct from Maegherman et al. (2022), who reported a recency effect

only in the likelihood of guilty indications but not in the conviction rates. The difference between our results may be due to the larger amount of evidence in our study. There might be a cognitive delay in the integration of evidence and this may have weakened the recency effect in verdicts with smaller evidence sets. Despite the larger amount of evidence in our studies, the task remains a short series task by Hogarth and Einhorn (1992)'s classification, and this may not fully represent realistic criminal cases with extensive evidence. Future studies should incorporate the manipulation of evidence volume and explore whether there is a delay in translating the probability of guilt judgments into verdicts, and how it might interact with order effects and response mode. Additionally, we did not manipulate the internal order of individual pieces of evidence within each set, which may have limited our ability to detect potential anchoring effects. Future research could explore how the sequence of specific items might influence judgments and how this interacts with response mode.

We also found that participants' probability of guilt judgments generally aligned with Bayesian predictions during the SbS response mode, supporting research suggesting that people are capable of updating their beliefs in a Bayesian fashion (Shengelia & Lagnado, 2021; Sommer et al., 2024; Thompson et al., 2013). However, some discrepancies were shown in EoS judgments, especially after viewing the defence evidence set. This might be due to the difficulty of the task of aggregating all evidence to answer the conditional probability questions. This finding can also be explained by the story model, in which N. Pennington and Hastie (1992) claim that people might not consider all evidence when making global judgments, as they may forget to include or downplay inconsistent evidence, leading their guilt judgments to depart from Bayesian predictions. Future studies could explore this further by explicitly prompting participants to construct narratives by collecting qualitative reasoning while manipulating the response mode. Moreover, answering the conditional probability questions might have aided participants in thinking about the probability of guilt ratings more accurately and rationally, which might explain why the values were generally consistent in the SbS stages. Future research can explore whether asking individuals to consider the conditional probabilities of new pieces of evidence assists them in evaluating the probability of guilt.

Our findings answer the question we started this paper with, implying that who speaks the last in the courtroom matters to the outcome of legal judgments. But more importantly, the evaluation process, which might be associated with different cognitive models used by fact-finders plays a more pivotal role. It highlights the importance of evidence organisation by attorneys and investigators when building cases and deciding charges and pleas.

### Acknowledgments

We would like to thank the anonymous reviewers for commenting on a previous version of this manuscript.

## References

- Anderson, N. H. (1959). Test of a model for opinion change. *The Journal of Abnormal and Social Psychology*, 59(3), 371.
- Ask, K., Rebelius, A., & Granhag, P. A. (2008). The 'elasticity' of criminal evidence: A moderator of investigator bias. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 22(9), 1245–1259.
- 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.
- Costabile, K. A., & Klein, S. B. (2005). Finishing strong: Recency effects in juror judgments. *Basic and Applied Social Psychology*, 27(1), 47–58.
- Dahl, L. C., Brimacombe, C., & Lindsay, D. S. (2009). Investigating investigators: How presentation order influences participant–investigators' interpretations of eyewitness identification and alibi evidence. *Law and Human Behavior*, 33(5), 368.
- Enescu, R., & Kuhn, A. (2012). Serial effects of evidence on legal decision-making. *European Journal of Psychology Applied to Legal Context*, 4(2).
- Furnham, A. (1986). The robustness of the recency effect: Studies using legal evidence. *The Journal of General Psychology*, 113(4), 351–357.
- Good, I. J. (1989). *Weight of evidence and the (bayesian) likelihood ratio*. Department of Statistics; Statistical Consulting Center, Virginia . . .
- Higgins, E. T. (1996). Activation: Accessibility, and salience. *Social psychology: Handbook of basic principles*, 133–168.
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive psychology*, 24(1), 1–55.
- Imhoff, R., & Nickolaus, C. (2021). Combined anchoring: Prosecution and defense claims as sequential anchors in the courtroom. *Legal and criminological psychology*, 26(2), 215–227.
- Kashima, Y., & Kerekes, A. R. (1994). A distributed memory model of averaging phenomena in person impression formation. *Journal of Experimental Social Psychology*, 30(5), 407–455.
- Keltz, L. C. I. K., & Adelman, L. (2008). Testing information order effects in a long series of evidence. *Engineering Management Journal*, 20(2), 22–29.
- Kerstholt, J. H., & Jackson, J. L. (1998). Judicial decision making: Order of evidence presentation and availability of background information. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 12(5), 445–454.
- Maegherman, E., Ask, K., Horselenberg, R., & Van Koppen, P. J. (2022). Law and order effects: On cognitive dissonance and belief perseverance. *Psychiatry, psychology and law*, 29(1), 33–52.
- Pennington, D. C. (1982). Witnesses and their testimony: Effects of ordering on juror verdicts. *Journal of Applied Social Psychology*, 12(4), 318–333.
- 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.
- Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the story model for juror decision making. *Journal of personality and social psychology*, 62(2), 189.
- 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.
- Schum, D. A., & Martin, A. W. (1982). Formal and empirical research on cascaded inference in jurisprudence. *Law & Soc'y Rev.*, 17, 105.
- Shengelia, T., & Lagnado, D. (2021). Are jurors intuitive statisticians? bayesian causal reasoning in legal contexts. *Frontiers in Psychology*, 11, 519262.
- Sommer, J., Musolino, J., & Hemmer, P. (2024). Updating, evidence evaluation, and operator availability: A theoretical framework for understanding belief. *Psychological Review*, 131(2), 373.
- Tetlock, P. E. (1983). Accountability and the perseverance of first impressions. *Social psychology quarterly*, 285–292.
- Thompson, W. C., Kaasa, S. O., & Peterson, T. (2013). Do jurors give appropriate weight to forensic identification evidence? *Journal of Empirical Legal Studies*, 10(2), 359–397.
- Wells, G. L., Wrightsman, L. S., & Miene, P. K. (1985). The timing of the defense opening statement: Don't wait until the evidence is in. *Journal of Applied Social Psychology*, 15(8), 758–772.