UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

How does knowledge of detainment affect juror reasoning?

Permalink

https://escholarship.org/uc/item/4w2463z3

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors

Dewitt, Stephen H Glatzel, Sammy Lagnado, David

Publication Date

2023

Peer reviewed

How does knowledge of detainment affect juror reasoning?

Stephen H. Dewitt (s.dewitt@ucl.ac.uk)

Department of Experimental Psychology, University College London, London WC1H 0AP, United Kingdom

Sammy Glatzel (sammy.glatzel.17@alumni.ucl.ac.uk)

Department of Experimental Psychology, University College London, London WC1H 0AP, United Kingdom

David A. Lagnado (d.lagnado@ucl.ac.uk)

Department of Experimental Psychology, University College London, London WC1H 0AP, United Kingdom

Abstract

Recent work suggests that the decisions to detain defendants before trial increase the likelihood of conviction. One reason may be that knowledge of detainment makes jurors more likely to convict. Previous work has claimed this as an example of 'bounded rationality' i.e., due to a simple bias. We argue that this inference represents sophisticated causal reasoning e.g. about information hidden from jurors such as criminal history. We examine whether the effect of detainment knowledge on conviction depends on rational inference by presenting participants with a legal vignette in a 2x2 design: a defendant either has or has not been detained, and this detainment decision is either (1) not explained or (2) explained as due to an iron clad rule always used for this class of crimes. We find an effect of detainment when it is not explained, but either no or a limited effect when explained, providing evidence against the 'bounded rationality' view. We provide qualitative extracts of participants' reasoning, demonstrating sophisticated and nuanced inferences from detainment to hidden information when the decision is not explained.

Keywords: causal reasoning; legal reasoning; heuristics and biases; bounded rationality; uncertainty

If a defendant is detained prior to their main trial (i.e. not granted bail), are they more likely to be convicted? While a correlation has been reported in the literature (e.g. Phillips, 2008; Spohn, 2009), the causal structure of the relationship has not been definitively shown. A correlation between detainment and subsequent conviction would be expected even if there were no direct causal effect of detainment on conviction because factors associated with the defendant and the nature of their crime may, rightly or wrongly, impact upon both (e.g. the strength and weight of the evidence against them, age, sex, race etc.) This 'common cause' theory can be seen in Figure 1.



Figure 1. A causal diagram depicting the 'common cause' theory

Notably, in Figure 1 there is no direct causation between pre-trial detainment and the main trial decision. This means that a particular individual's chances of being convicted at the main trial do not change depending on whether a particular judge decides to detain them or not. Recent work (Lee, 2019) has challenged this model. Lee used propensity score matching to test for the relationship between pre-trial and main-trial decision while 'controlling' for a large number of the common factors involved in the 'common cause' theory such as gender, age, race, the nature of the crime, and the individual's criminal history. In this way they intended to test the theory that, as well as the common cause effect, there exists a 'direct cause' effect from detainment decision to main trial, as can be seen in Figure 2. Their conclusion, in line with previous work (e.g. Tartaro and Sedelmaier, 2009) was that when controlling for all of the common factors, an effect of detainment on main trial decision was still present i.e. that as well as the common cause effect, there is some direct effect of detainment on main-trial decision.

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*. ©2023 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).



Figure 2. A causal diagram depicting the 'direct cause' theory

Assuming a direct effect does exist, it is important to continue to probe the causal mechanisms at play to understand how we can intervene. Lee (2019) cites two possible reasons for the direct effect and called for more research to determine the causal mechanisms at play. These are illustrated in the lower part of Figure 3. The first reason (2a) is systemic and known as the 'Domino effect' (Tartaro and Sedelmaier, 2009). When individuals are detained pretrial this has a range of knock-on effects which reduce their chances at the main trial. For example, they will likely have less access to their lawyer and capacity to assist in developing their defence case. The second reason (2b) is psychological. Jurors are often aware of the detainment decision, and the detained defendant may even appear in court in prison attire. It has been suggested therefore that jurors' decision making may be inappropriately affected by this.



Figure 3. A causal diagram depicting two 'direct cause' models.

The question of whether and why (2b) occurs is the focus of the present paper. This effect, if it exists, also has multiple possible explanations. Lee (2019) explains this psychological effect in terms of 'bounded rationality' (Alboneti, 1991; Simon, 1956; Tversky & Kahneman, 1974), with an implication that such use of detainment information is in some way either non-rational or only rational because of the bounded nature of human cognition. In line with changing thought around this (e.g. Gigerenzer & Goldstein, 1996; Hahn & Harris, 2014: Mackenzie, 2003: Tuckett & Nikolic, 2017), we would argue a different narrative. We propose that the detainment decision provides valuable information to jurors, and their use of it shows sophisticated causal inference (Fenton, Lagnado & Neil, 2013; Lagnado & Gerstenberg, 2017; Pearl, 2009). It isn't in doubt that the 'common cause' effect depicted in Figure 1 exists. Therefore, if a pre-trial judge bases their decision on information about the crime and the defendant, then that decision legitimately conveys information about those things: we know from Lee's (2019) and Phillips' (2009) data that someone who is detained is more likely to have committed a more serious crime and to have a criminal history. Excluding legal norms, if jurors make this inference, this is simply making good use of available information in a highly uncertain environment.

While controlled-observational studies like Lee (2019) have the advantage of studying real world data, the result relies on the always-questionable assumption that you have controlled for everything i.e. severed the solid black lines in Figure 3 entirely. This is because you can only control for those factors you know about / have thought of (unknown factors may slip through). Perhaps for this reason, randomised controlled trials, which use stronger methods to sever causal links, typically show a smaller effect size than controlled-observational studies across a range of fields (e.g. Pearl, 2019; Rothman, 2012). Experimental trials of course suffer from the problem of ecological validity, so these two methods complement each other well, and indeed we seek to complement the findings of Lee (2019) by conducting an experimental test of the research question of whether there is a direct effect of pre-trial detainment on main-trial conviction via an influence on jurors' beliefs about the guilt of the defendant. Furthermore, if this effect is found, we hope to shed light on why this might be the case.

In the present study we aim to experimentally sever the solid lines in Figure 1 by explicitly explaining that the detainment decision was not caused by factors such as crime severity or criminal history. Instead of listing all the common factors, and thus succumbing to the same possibility of leaving out an unknown factor, as with Lee (2019), we instead cut all possible links by 'explaining away' the pretrial detainment decision as being entirely caused by something else: notably, an iron clad rule by this particular judge to always detain for the current class of crimes. This also has the benefit of nullifying the inference that the pretrial judgement indicates anything about the judge's beliefs about the defendant. If the judge detains 100% of the time for cases like this, their decision to detain in this case should convey no information about the judge's beliefs about the defendant or about the characteristics of the case or the defendant. This version of the situation can be seen in Figure 4.



Figure 4. A causal diagram depicting the structure of the experimental manipulation.

If participants are succumbing to a simple irrational bias whereby knowledge about detainment influences their main trial decision, regardless of how rational that inference is, they should still be more willing to convict a defendant who has been detained when faced with the situation in Figure 4. However, if they are sensitive to the rationality of this inference, detainment will only increase conviction estimates at the main trial when they are in the situation depicted by Figure 1 i.e. when they are told that the judge decided on this occasion to detain the defendant, based on the information available. In the present study we seek to compare participant judgements across these two different scenarios. We cross this with the judge's decision: i.e. to detain or not to detain, creating a 2 (explained vs unexplained) x 2 (detained vs notdetained) design. In this way we aim to determine firstly whether participants are influenced in their conviction judgements by pre-trial detainment and secondly, if they do make this judgement, whether that is influenced by how 'rational' it is under the circumstances. We have three quantitative expectations:

- 1. In the unexplained conditions, conviction judgements will be higher in the detained condition than the not-detained condition.
- 2. In the explained conditions, conviction judgements will not be different between the detained and the not-detained conditions.
- 3. As a result of (1) and (2) we will find an interaction effect, such that the difference between the detained and not-detained conditions will be smaller in the explained conditions than the unexplained conditions.

Beyond these quantitative hypotheses, we also intend to ask participants to explain their reasoning in an open-ended text and will apply qualitative analysis to better understand the nature of the inferences participants are making in each condition.

Method

Design

The study presented participants with a vignette describing the result of a pre-trial decision about whether a defendant was detained (not granted bail) or not. The study used a 2 (Not detained vs Detained) x 2 (Unexplained vs Explained) design. The key quantitative dependent variable was the participant's percentage estimate of the defendant being convicted. Participants were also asked to explain their reasoning in an open text box.

Participants

Three hundred and sixty-three participants were recruited from Prolific Academic and were paid £9 per hour. No participants were excluded. The only requirement placed on the sample was that they were fluent in English. Unfortunately, a survey error prevented participants from accessing our demographic questions. However, Prolific Academic provides extensive demographic details of the population that we randomly sampled from. The population has good representation across European countries (39% UK, ~20% other European countries) and the US (31%), but has little representation from either Asian, African, or South American countries. The population has a fairly good spread of age ranges: while 48% of the population are aged 20-30, 26% are 30-40, 11% are 40-50, 5% are 50-60 and 2% are 60-70. They also have a nearly equal sex split, with 45% reporting male and 55% reporting female.

Materials & Procedure

After proceeding through the consent page, participants were presented with the below scenario (all materials and data are available at https://osf.io/46byc/?view_only=79776dba91e740559344c9 ea2e2f0600):

You are part of a jury in a case in which a man, X, has been accused of a crime. X has been accused of burglary. X attends a pre-trial hearing with Judge Armstrong to determine if he will be detained up until his trial

They were then presented with one passages of text in Table 1, depending on their condition.

Table 1. Text describing the decision of the pre-trial judge for all four conditions.

	Unexplained	Explained
Not detained	The judge DECIDES that X will NOT be detained up until his trial.	Judge Armstrong NEVER detains suspects for burglary cases and therefore X will NOT be detained up until his trial.
Detained	The judge DECIDES that X WILL be	Judge Armstrong ALWAYS detains suspects for burglary cases and therefore X

detained up	WILL be detained up
until his trial.	until his trial.

Following this, participants were asked to provide a percentage estimate to the following question: 'At this early stage, what do you personally think is the probability that X will be found guilty of burglary at trial?' They were then presented with an open text box and asked to explain their reasoning.

Results

Quantitative

The basic quantitative results can be seen in Table 2 and these were analysed using linear regression. Looking only at the unexplained conditions (first column), a model with 'Detainment' as the predictor found that participants estimated the chance of guilt as 22.1% higher when the defendant was detained vs not detained (t[1, 182]=10.1, η 2=0.36, p<.001). Looking only at the explained conditions (second column), a model with 'Detainment' as the predictor found that participants estimated the chance of guilt as 7.5% higher when the defendant was detained vs not detained (t[1,177]=2.4, n2=0.03, p=.019). Detainment therefore seems to produce a 14.6% larger effect within the unexplained conditions than the explained conditions. This was assessed with a regression model predicting 'Conviction' from 'Detainment', 'Explained' and an interaction term. The interaction term showed a highly significant effect (t[1, 359]=-3.8, η 2=0.03, p<.001). Overall, we can therefore see that while explaining detainment reduced the effect of detainment on prosecution estimations substantially (-14.6%), a potential lingering effect (7.5%) of detainment still remains.

Table 2. Quantitative results by condition (standard errors in grey brackets)

Difference	22.1% (p<.001)	7.5% (p=.019)	-14.6% (p<.001)
Detained	64.8% (1.3%)	57.0% (2.2%)	-7.8% (p=.004)
Not detained	42.6% (1.7%)	49.5% (2.2%)	6.8% (p=.013)
	Un- explained	Explained	Difference

Before we present the main qualitative results, it is important to introduce a particular response that was found which may affect the interpretation of these quantitative results. We found that 23 individuals in the explained conditions interpreted the study in an interesting manner but which we didn't intend. Specifically, they believed that the same judge who gave the detainment decision would also be conducting the main trial. Based on this assumption, they thought that the judge's pre-trial decision 'rule' shows a bias against these kinds of crimes and so they would likely continue to be harsh at the main trial. We labelled these participants 'confounds' as their interpretation of the study represented a confound, but which, as will be seen in the main qualitative section, further illustrates the sophistication and nuance of our participants' reasoning. A few responses from these participants can be seen below:

P353: "Because the judge always detains burglary suspects I believe he may treat them more harshly until proven innocent. He probably has disdain for them and would like to see them in jail."

P339: "Since the judge always detains suspects for a relatively minor crime like burglary it feels like he's very strict."

Interestingly, these participants have therefore re-created, or found, a possible new 'common cause' connection between the pre-trial and main trial, despite our efforts to sever all such connections in the explained conditions. This therefore resurrects the same causal structure as in Figure 1, and relies upon the same inference, except that instead of some attribute of the defendant having a causal impact upon both the pre-trial and main trial, it is the attitudes and bias of this particular judge. This can be seen in Figure 5.



Figure 5. The causal structure representing participant reasoning in the 'confound' response

Given their assumptions, it makes sense that these participants would expect the defendant who is detained to be more likely to be prosecuted at trial, because the judge's detainment decision could indicate a bias against people who commit these kinds of crimes, or a general harshness. Interestingly, it seems that these participants are responsible for a large part of the 'lingering' effect of detainment in the explained pair of conditions. With the 'confound' participants removed, the difference between detainment and non-detainment in the explained pair of condition is reduced to 4.8% (t[1, 156]=1.4, p=.16) rather than 7.5% with the 'confound' participants (Table 2).

Qualitative

The qualitative reasoning data was coded blind to condition by the first author. The first author's codebook was then emailed to the second author, with definitions of each code (which can be seen at the public repository), and the second author coded all the responses independently and blind to condition according to this codebook. Over 95% agreement was found, and the remaining responses were resolved through discussion. Generally, a conservative approach was taken by both coders, where if they were not confident in the code it was assigned 'Unclassified / Other'. The conservative nature of the approach can be seen in the high number (39.7%) of cases classified this way.

Table 3. Percentage of participants assigned each code by condition. No Info/Evidence = 'No information / Evidence', Innocent-UPG = 'Innocent until proven guilty', Unclassified/Oth = 'Unclassified / Other'. N-D = Not detained, D = Detained. The most frequent code assigned other than 'Unclassified / Other' is made bold for each condition.

	Unexplained		Explained	
	N-D	D	N-D	D
n	93	91	84	95
Guilt more likely	-	54.9	1.2	8.4
Guilt less likely	39.8	-	3.6	-
Confound	-	-	8.3	13.7
No Info/Evidence	23.7	13.2	36.9	25.3
Innocent-UPG	1.1	3.3	1.2	4.2
Unclassified/Oth	35.5	28.6	48.8	48.4

Guilt more likely

The most common code in the 'Unexplained-detained' condition was for participants to say that they believed the judge's decision revealed that the judge must have some information about the defendant which the participant didn't know about, but which indicated a greater likelihood of guilt. For example:

P120: "Because [the] judge decided to detain him, so there is [a] higher chance that he is guilty, maybe he said something that gave [the] judge a reason to keep him or maybe he has no alibi."

P122: "Since the judge chose to detain X until trial, [it] might be because the judge possibly thinks X is guilty and is afraid that X might do it again or run." P130: "If the judge decided to detain him it might be because they have many proofs of him being guilty already, and there is a big probability he will do it again or maybe he has a history of previous offenses and there is a reason to think he might not show up for the trail or do something bad in the meantime."

Effectively these participants are reasoning across the causal model in Figure 1. They think that the pre-trial decision is indicative of some factor which influenced that decision, such as their criminal history, or some aspect of their crime, and which will similarly have a direct influence on the main trial.

Guilt less likely

Exactly mirroring this, the most common code in the 'Unexplained-not-detained' condition was to state that the judge's decision revealed some factor indicating a lower likelihood of guilt. For example:

P4: "He is allowed to be out of jail until the trial, which gives me reason to believe that the judge thinks he is a good man."

P6: "Maybe this means he is not dangerous, they would [have] detained him if he was"

No Information / Evidence

The most common code in both 'explained' conditions was 'No Information / Evidence' (36.9% when not detained, 25.3% when detained). Participants giving this response stated either explicitly that the detainment decision didn't provide any information, or simply stated they didn't have any information to make the judgement, without directly mentioning the detainment decision. For example, P52 said "There is absolutely no relevant information to go on. Therefore, it is impossible to tell either way." Finally, and perhaps reflecting a similar thought process to the 'No Information / Evidence' code, a small number of participants also cited the 'innocent until proven guilty' principle in their response, for example P115 who said "Innocent until proven guilty. Need more information"

Discussion

The overarching research questions driving this paper are whether and why pre-trial detainment increases conviction rates. Evidence that detainment directly affects conviction rates has been provided by the real-world research previously conducted (e.g. Lee, 2019, Phillips, 2008) however the present work has been able to complement this by not only systematically controlling certain key factors, but also extracting reasoning processes from participants to further our understanding of why this might occur. This design has yielded several important and nuanced answers to this question which could be informative in considering protocols around the way detainment information is revealed to jurors. It also has bearing on the ongoing debate about whether human inference is best thought of as a collection of crude heuristics or as a sophisticated inference engine.

First, when the reason for detainment is not explained, participants may infer that detainment is (1) based upon a judge's decision which (2) is based upon their assessment of the body of evidence against the defendant and (3) therefore detainment implies that the body of evidence is more likely to suggest guilt than if they were not detained. There are multiple important ramifications of this. First, many of our participants stated that this indicated that the judge had information suggesting guilt, which the participant didn't have access to. Clearly our participants had less information available than a real juror, however it is common for information to be withheld from jurors in order not to prejudice them, and jurors may believe, rightly or wrongly, that the judge making the detainment decision would have access to this information. Jurors may therefore use a detainment decision as a backdoor route to infer about the nature of this withheld information. Secondly, even if it were stated that the judge making the detainment decision only had the same information as jurors, a similar inference could still be made based on the assumption that the pre-trial judge has expertise that the jurors do not have i.e. the pre-trial decision is an indicator of the judge's belief about the defendant's guilt, based upon their extensive experience and legal knowledge, and therefore may still be considered valuable information by an uncertain juror.

Second, we consider our participants' inferences rational and we do not find evidence for a simple bias. When we explained the pre-trial decision, severing the causal links making the above inference possible, the difference between detainment and no-detainment conditions dropped substantially. While a lingering difference was detected, which could have indicated a simple bias, we have shown that this was in fact due to participants making a separate, unintended on our part, but reasonable, inference (that the pre-trial judge will also be the main trial judge and seems harsh / biased). This again highlights the point that while there is often a tendency to label an effect of information on behaviour where we believe there is no causal connection as a bias, there may well be a subtle inference lying behind it that the researcher hasn't considered. After removing these participants from our analysis, we saw no impact of detainment on beliefs about guilt when detainment was explained. Furthermore, the most common qualitative response in these conditions was to say that there was no relevant information available. This has clear ramifications for communication about pre-trial detainment. Firstly, the ideal situation (but which may not be possible) would be to not communicate whether the defendant was detained or not, as it seems that many people, under the inherent uncertainty of the meaning of this detainment, will make a range of inferences that, even if rational based on their knowledge state, could lead to ethically and legally undesirable effects upon their decision making. Crucially, while we have described these inferences as 'rational', it is important to note that jurors may well make this inference regardless of

whether it is true or not in a particular case because they are reasoning from a position of uncertainty. Therefore, if there is a legitimate explanation of detainment which genuinely mitigates its impact on guilt, such as it being routine for the current class of crimes, our research suggests that jurors will be able to mitigate the impact of detainment upon their judgements (i.e. they do not appear to be prone to a simple prejudice), and it would be better to communicate this reason. We are not confident that telling jurors not to take detainment information into account, even where that detainment information does carry information, would be effective.

There are limitations in our study that should be taken into account. While we have presented our study as complementary to Lee (2019) this only applies to the study of the effect of detainment decision on jurors (which we find, in line with them, does exist, when detainment is not explained, as it usually isn't). However, Lee did not study, and we are not aware of a non-experimental study of why detainment has this impact on jurors. Lee proposed 'bounded rationality' but did not investigate this, and it is likely to be difficult to investigate outside of an experimental setting. Inferences from our study should therefore be tempered by considerations of ecological validity. Our study is highly simplistic, with no actual stakes. In a real legal situation jurors may feel stressed by the complexity of the case and the importance of their decision for the defendant's life. In such a situation, especially if they feel highly uncertain, it is possible that they will be more prone to resort to simple heuristics such as 'go with the detainment decision' even when it clearly carries no information. Our participants had very minimal information (only the detainment decision) and the study was not cognitively demanding so participants were unlikely to feel the strain of overwhelming amounts of information that can occur in real cases. Furthermore, precisely in order to keep things simple we presented the situation to the participants as being 'before' the main trial, asking them for their expectations of the defendant's guilt. While this does neatly demonstrate that communication of detainment may mean that jurors 'enter' a case already thinking the defendant more likely to be guilty, we were unable to study how this expectation interacts with the evidence they receive. It is possible that they start with this expectation but that the effect of it is minimal by the time they have to make their final decision. Alternatively, it may interact with future information in a confirmation-bias like Nickerson, 1998), influencing their manner (e.g. interpretation of the evidence such that they become increasingly convinced of their guilt. Further studies building upon the current work with the same design, but which examine how participants go on to interpret the same set of evidence would be valuable. Similar study setups but in a more ecologically valid setting, such as in a mock court would also be valuable.

References

Albonetti, C. (1991). An integration of theories to explain judicial discretion. Social Problems, 38, 247-266.

- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. Psychological Review, 103(4), 650–669.
- Fenton, N. E., D. Lagnado and M. Neil (2013). "A General Structure for Legal Arguments Using Bayesian Networks." Cognitive Science 37, 61-102.
- Hahn, U., & Harris, A. J. L. (2014). What Does It Mean to be Biased. Motivated Reasoning and Rationality. In Psychology of Learning and Motivation - Advances in Research and Theory (1st ed., Vol. 61).
- Lagnado, D. A., & Gerstenberg, T. (2017). Causation in legal and moral reasoning. The Oxford Handbook of Causal Reasoning, 565-601.
- Lee, J. G. (2019). To Detain or Not to Detain? Using Propensity Scores to Examine the Relationship Between Pretrial Detention and Conviction. Criminal Justice Policy Review, 30(1), 128–152.
- McKenzie, C. R. M. (2003). Rational models as theories not standards – of behavior. Trends in Cognitive Sciences, 7(9), 403–406.
- Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. Review of General Psychology, 2(2), 175–220.
- Pearl, J. (2009). Causality. Cambridge university press.
- Pearl, J., & Mackenzie, D. (2018). The book of why: the new science of cause and effect. Basic Books.
- Phillips, M. T. (2008). Pretrial detention and case outcomes, part 2: Felony cases. New York: New York City Criminal Justice Agency.
- Rothman, K. J. (2012). Epidemiology: an introduction. Oxford university press.
- Simon, A. H. (1956). Rational choice and the structure of environment. Psychological Review, 63, 129–138.
- Spohn, C. (2009). Race, sex, and pretrial detention in federal court: Indirect effects and cumulative disadvantage. University of Kansas Law Review, 57, 879-901.
- Tartaro, C., & Sedelmaier, C. M. (2009). A tale of two counties: The impact of pretrial release, race, and ethnicity upon sentencing decisions. Criminal Justice Studies, 22(2), 203–221.
- Tuckett, D., & Nikolic, M. (2017). The role of conviction and narrative in decision-making under radical uncertainty. Theory and Psychology, 27(4), 501– 523.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science (New York, N.Y.), 185(4157), 1124–1131.