Maastricht University

You don't know: knowledge as supportive alibi evidence

Citation for published version (APA):

Nieuwkamp, R., Horselenberg, R., & van Koppen, P. (2023). You don't know: knowledge as supportive alibi evidence. *Psychiatry Psychology and Law*, *30*(5), 695-712. https://doi.org/10.1080/13218719.2022.2116608

Document status and date:
Published: 03/09/2023

DOI:
10.1080/13218719.2022.2116608

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

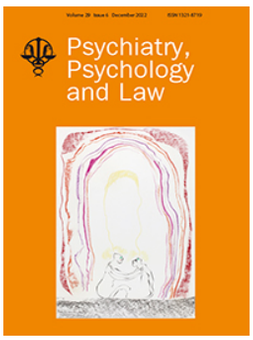
www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 20 Apr. 2024



Psychiatry, Psychology and Law

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tppl20

You don't know: knowledge as supportive alibi evidence

Ricardo Nieuwkamp, Robert Horselenberg & Peter van Koppen

To cite this article: Ricardo Nieuwkamp, Robert Horselenberg & Peter van Koppen (2022): You don't know: knowledge as supportive alibi evidence, Psychiatry, Psychology and Law, DOI: 10.1080/13218719.2022.2116608

To link to this article: https://doi.org/10.1080/13218719.2022.2116608

Published online: 18 Sep 2022.

Submit your article to this journal

Article views: 92

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tppl20

Psychiatry, Psychology and Law, 2022
Vol. 0, No. 0, 1–18, https://dx.doi.org/10.1080/13218719.2022.2116608

Routledge
Taylor & Francis Group

Check for updates

You don't know: knowledge as supportive alibi evidence

Ricardo Nieuwkamp[a], Robert Horselenberg[b] and Peter van Koppen[b,c]

[a]*Federal Judicial Police Antwerp, Belgium;* [b]*Faculty of Law, Maastricht University, Maastricht, The Netherlands;* [c]*Faculty of Law, VU University Amsterdam, Amsterdam, The Netherlands*

Until now, supportive evidence for alibis has been conceptualised into two distinct types: witness and physical evidence. The present study examined whether knowledge, as a third type of supportive evidence, can contribute to the understanding of evidence for alibis. Three experiments were conducted in which police detectives, laypersons and undergraduate students were asked to evaluate four alibis with witness, physical or knowledge supportive evidence, or with no supportive evidence. The results from the three experiments show that knowledge evidence is equally believable as strong witness evidence. We also found that not all items of strong physical evidence are evaluated as equally strong and believable. We therefore suggest adjusting the criteria to determine the strength of physical evidence and conducting more research on knowledge evidence.

Key words: alibi; knowledge; police detectives; supportive evidence; taxonomy.

1. Introduction

When the police services question a suspect who denies any involvement in the alleged crime at least two options arise: either the suspect is lying about his involvement or the suspect is genuinely not involved in the crime. At a certain point in time, these two scenarios need to be evaluated in court, by a jury and/or a judge (Nieuwkamp & Mergaerts, 2022). In both scenarios a suspect may present an alibi, stating he could not be involved in the alleged crime as he was at another location (Olson & Wells, 2004). At first, the suspect will have to report where he was at the time the alleged crime was committed and will potentially support his account with evidence. Based on this account and the reported evidence, its believability can be assessed. In proven wrongful conviction cases, where an alibi was presented by the defendant, the alibi was often not believed (about 70% of the first 250 DNA exoneration cases in the USA; Garrett, 2011). Most of the wrongful convicted people presented an alibi that was supported by no or weak evidence (Cardenas et al., 2022; Wells, 1998). These results illustrate that the strength of the supportive evidence plays an prominent role in determining the alibi believability and ultimately whether the court believes that the suspect is innocent or guilty.

Previous research on alibi evaluation has concluded that the presence and strength of supportive evidence, categorised as physical or witness evidence, are the most important factors in determining alibi believability (e.g. Allison, 2022; Pozzulo et al., 2015). In the alibi literature, however, not all types of supportive evidence are captured. In the present paper we focus on a type of supportive evidence not yet studied, namely knowledge evidence. Knowledge evidence refers to

Correspondence: Ricardo Nieuwkamp, Federal Judicial Police Antwerp, Noordersingel 25–27, 2140, Belgium. E-mail: ricardo.nieuwkamp@police.belgium.eu

© 2022 The Australian and New Zealand Association of Psychiatry, Psychology and Law

information that could only be known to a person if that person was at a certain event.

1.1. Supportive evidence

Classically, two types of supporting alibi evidence are studied: witness evidence and physical evidence. The evidence for an alibi can consist of either type or a combination of both types (Culhane et al., 2013; Nieuwkamp et al., 2017). To determine the strength of supportive evidence, Olson and Wells (2004) created an alibi taxonomy consisting of the two classic types of supportive evidence. The strength of witness evidence is based on the degree to which the witness might be motivated to lie in the suspect's favour (Sullivan, 1971). It is expected that if the relational distance between the suspect and witness is large (e.g. a previously unknown stranger), the witness is assumed to be the least inclined to lie for the suspect. In the taxonomy, such a witness is referred to as an *unmotivated other witness*, considered the strongest type of witness evidence (Olson & Wells, 2004). The witness's testimony is viewed as weaker evidence when the relational distance is small (e.g. the suspect's girlfriend) and is referred to as a *motivated familiar other witness.*

Physical evidence is considered strong supportive evidence when three criteria are met, the evidence: (a) is linked to the suspect; (b) contains an indication of the date and time; and (c) contains location information (Olson & Charman, 2012). That is, the evidence has to place the suspect at a specific place at a specific time other than the crime scene. CCTV recordings are an example of strong physical evidence, as they are hard to alter (Olson & Wells, 2004). In contrast, a receipt is an example of weak physical evidence, for a receipt lacks a solid link to whom purchased the goods. The rationale underlying the three criteria is that the more difficult it is for the suspect to fabricate the physical evidence or to obtain it from someone else, the stronger the evidence is.

The presence of physical evidence has a great impact on the alibi believability. An alibi supported by the weakest type of physical evidence (e.g. a receipt) is still more believable than an alibi supported by the strongest type of witness evidence (i.e. an unmotivated familiar other witness; Olson & Wells, 2004). So far, there was little variation in the items presented as physical evidence. Typically, strong evidence consists of the recordings of CCTV cameras, and weak evidence consists of a receipt in alibi research (e.g. Allison & Brimacombe, 2010; Jung et al., 2013), although a credit card receipt or an airline ticket stub are also examples of strong physical evidence (Olson & Charman, 2012).

A suspect can also present a combination of both types of evidence for his alibi (Culhane et al., 2013; Nieuwkamp et al., 2017). The strength of a combination of supportive evidence can also be determined using the taxonomy of Olson and Wells (2004) in such that it depends on the strength of the witness and physical evidence separately. For example, weak witness evidence combined with weak physical evidence is of equal strength to strong physical evidence without witness evidence (Olson & Wells, 2004). However, to date it remains unclear how a combination within one type of evidence should be assessed. For example, are two motivated familiar witnesses stronger evidence than one unmotivated stranger witness? That question needs to be addressed in future research.

1.2. Knowledge as supportive evidence

Although the taxonomy helps to understand the evaluation of alibi believability, it is unclear whether all types of supportive evidence are captured in the taxonomy. Based on the well-known Dutch *Deventer* murder case (van Koppen & Nieuwkamp, 2017; van Koppen & van Koppen, 2010) it is argued that knowledge could be a novel type of supportive alibi evidence. In the *Deventer* murder case, the convict presented an alibi supported by evidence that cannot be found in the alibi

literature. The convict said that he could not have committed the murder since he was in a traffic jam around 30 km from the murder site. His knowledge of the traffic jam could support his alibi. The police verified that at the time of the murder there was indeed a traffic jam where the convict reported it to be, and that the jam had not been mentioned in the radio traffic reports. The question that remained was whether the convict knew about the traffic jam from his own experience or if he got that knowledge through another source (Haan & University Press, 2009). Haan and University Press (2009) argued that the convict could have seen the traffic jam from another location or could have talked to colleagues who were stuck in that traffic jam. Regardless of how the convict obtained the knowledge, the concept of knowledge as supportive evidence for an alibi can be valuable for the alibi literature.

Knowledge evidence can consist of any kind of information that the suspect has to support his alibi for what he was doing at the time the alleged crime was committed. The strength of knowledge evidence could perhaps be assessed based on the same rationale as physical evidence: the more difficult it is to fabricate or obtain, the stronger the evidence is. In contrast, the easier it is to obtain particular knowledge from another source (e.g. talking to someone else) or to base the alibi on prior knowledge of the suspect (e.g. being able to tell the plot of a classic movie like Harry Potter and the Philosopher's Stone that was broadcast on television on the night of the crime), the weaker the evidence is. Unique knowledge evidence is therefore the strongest sort of knowledge evidence and can be defined as knowledge that could only be known to the person if that person was present at a certain event. Other examples could be: while driving from home to work being passed by an ambulance at a specific location (verifiable) or got flashed for speeding by an automatic speed camera (verifiable), seeing fireworks in a city park when going out for an evening run in the summer (verifiable but could be obtained from another source); watching a live television programme where there was a power failure for a brief moment (verifiable but could be obtained from another source), and so on. Based on these examples and identical categories that apply to physical evidence, the alibi taxonomy (Olson & Charman, 2012; Olson & Wells, 2004) could be updated towards the taxonomy presented in Table 1.

Knowledge evidence has the potential of being a meaningful addition to the two classical types of supportive alibi evidence because it has not yet been researched but could be a valuable concept. The present paper examines how knowledge evidence is evaluated by diverse groups of participants.

1.3. Police detectives as participants

Our understanding of alibi evaluation has grown over the years, but most of that research has been conducted with undergraduate students as participants (e.g. Sauerland et al., 2016). In practice, the police detectives' role is very important in evaluating alibis in criminal pre-trial investigations in both adversarial and inquisitorial legal systems (Nieuwkamp & Mergaerts, 2022). Depending on whether they consider the alibi to be believable, the police officers may invest more or less time and energy investigating the suspect's potential involvement in the alleged crime. By doing so, they influence subsequent procedures. Nevertheless, the use of police detectives in alibi evaluation research is rare (Culhane & Hosch, 2012; Dysart & Strange, 2012; Eastwood et al., 2016; Nieuwkamp et al., 2018) compared to the body of alibi research (for an overview see: Behl and Kienzle (2022)). Research shows that important differences exist between students and police detectives in alibi evaluation. When both groups of participants are asked to evaluate alibis on five aspects (e.g. the quantity of alibi witnesses and the quality of the relation between the suspect and the alibi witness), police detectives focus on more aspects (2.5 aspects) than students (1.7 aspects; Eastwood et al., 2016). In

Table 1. Alibi taxonomy.

Witness evidence	Knowledge or physical evidence			
	None	Weak evidence: easy to fabricate (or to obtain from another source)	Moderate strong evidence: not easy or difficult to fabricate (or to obtain from another source)	Strong evidence: difficult to fabricate (or to obtain from another source)
Non-motivated familiar other	*Least believable*			*Most believable*
Non-motivated other				
Motivated other				
None				

addition, police detectives find the number of alibi witnesses more important than the relational distance with the suspect (Eastwood et al., 2016). Furthermore, police detectives focus more on the verifiability of the alibi (see: Verschuere et al., 2021, for a meta-analysis) rather than the perceived strength of the reported evidence as students do (Nieuwkamp et al., 2018). Police detectives also come to more guilty judgements than students (Culhane & Hosch, 2012), which can be explained by their distrust of suspects in general (e.g. Kassin et al., 2003; Meissner & Kassin, 2002; Vrij, 2008) and of their alibis in particular (Dysart & Strange, 2012). When a suspect changes his alibi, for example, about 80% of the police detectives think that the change is made due to deception rather than error (Dysart & Strange, 2012), although such misrecollections are common occurrences (Crozier et al., 2017; Strange et al., 2014).

1.4. Present studies

The present experiments were designed to determine how knowledge evidence is evaluated in terms of believability and strength compared to witness, physical and no supportive alibi evidence among police detectives, undergraduate students and laypersons.

In general, we expected an alibi supported by any type of evidence to be more believable than an alibi without supportive evidence (Hypothesis 1) based on earlier research (e.g. Olson & Wells, 2004). In line with previous studies (e.g. Pozzulo et al., 2015), we expected knowledge evidence to be equally strong and believable as witness evidence (Hypothesis 2), but physical evidence to be the strongest and most believable (Hypothesis 3).

2. Experiment 1

2.1. Method

2.1.1. Design

The experiment consisted of a 2 (group: laypersons vs. police detectives) × 4 (evidence:

no evidence; witness evidence; physical evidence; and knowledge evidence) mixed-factorial design. The between-subjects factor consisted of two groups of participants. The within-subjects factor was the type of supportive evidence presented.

2.1.2. Participants

One group of laypeople and two groups of police detectives were recruited for the experiment. The police detectives consisted of one group of police detectives from the regular police force ($n = 48$) and one group of detectives working for the Dutch Fiscal Information and Investigation Service ('Fiscale inlichtingen- en opsporingsdienst', FIOD; $n = 40$). The police detectives from the regular force were recruited during a theme day in the Netherlands. They were between 29 and 59 years old ($M = 45.6$, $SD = 7.8$). The second group was recruited by contacting the FIOD. The FIOD provided a list of 58 employees based somewhere in the country who were highly likely to have experience in evaluating alibis based on their previous position in the service. A total of 41 FIOD detectives completed an online version of the questionnaire. One participant was excluded because he had never had to evaluate an alibi. The age of the 40 remaining FIOD detectives ranged from 35 to 62 years ($M = 49.8$, $SD = 6.6$).

The laypersons ($n = 56$) were recruited in public areas in two average-sized cities in the Netherlands. Their ages varied between 19 and 62 years ($M = 33.8$, $SD = 14.5$).

No significant differences were observed between the two groups of police detectives in regard to demographics and their scores on the dependent variables ($p > .05$). The two groups of police detectives were therefore combined. Two participants from the group of police detectives had at least one missing value for the alibi evaluation; they were excluded from the analysis. Hence, the final group of police detectives consisted of 86 participants (see Table 2 for an overview).

All participants completed the questionnaire individually and received no compensation. The standing ethical committee of the Faculty of Psychology and Neuroscience at Maastricht University approved the study.

2.1.3. Materials and procedure

The FIOD detectives received a digital copy of the questionnaire using www.surveygizmo.com. A digital version was created because the detectives were located throughout the country. Potential participants were contacted via email with the request to independently complete the questionnaire. All other participants completed the questionnaire individually using a paper-and-pencil form.

After consent was received, the participants were asked for some demographic information (e.g. gender and age), and the detectives were also asked how often they evaluate an alibi on average per month. All participants were then presented with a short description of a mock armed robbery that was committed the previous Friday at 10.05am at a small grocery store at the train station of a Dutch city, Amersfoort. The police arrested four men who all presented the same alibi (i.e. 'I took the train at 9.26am in Amersfoort to arrive at 10.25am in Amsterdam').[1] One suspect presented the alibi without supporting evidence, and the other three suspects presented witness, physical or knowledge supportive evidence. The participants were asked to evaluate the presented alibis in a randomised order to preclude any order effects. Each suspect statement and subsequent questions were printed on a new page.

The witness and physical evidence consisted of the strongest type of evidence according to the taxonomy of Olson and Charman (2012). The alibi witness was an unmotivated other witness, namely the ticket collector who recognised the suspect on the train since the suspect often takes the train from Amersfoort to Amsterdam. The physical evidence consisted of the records of the suspect's personal public transportation card, the so-called OV

Table 2. Participant characteristics in percentages.

	Group of participants			
	Laypeople ($n = 56$)	Police detectives specialised in sexual delinquency ($n = 48$)	FIOD detectives ($n = 40$)	Average[b]
Sex				
Male	46.4	41.7	87.5	58.5
Female	53.6	58.3	12.5	41.5
Highest completed education[a]				
Elementary school	0.0	0.0	5.0	1.7
Secondary school	36.4	35.5	40.0	37.3
Intermediate Vocational Education	14.5	33.3	35.0	27.6
College or University	49.0	31.3	20.0	33.4
Alibi evaluation in current job				
Never		0.0	12.5	7.3
Less than once a month		54.2	70.0	61.1
Less than once a week		22.9	10.0	16.5
About once a week		18.8	2.5	10.7
Two to four times a week		2.1	5.0	7.1
About once a day		2.1	0.0	1.1

Note: The dependent variable is bold. FIOD = Dutch Fiscal Information and Investigation Service ('Fiscale inlichtingen- en opsporingsdienst').
[a]One missing value; only the valid percentages are presented.
[b]Average percentage and total number of participants are only displayed for the two groups of police detectives when it comes to alibi evaluation.

chip card. The Dutch OV chip card is a personal card containing the passenger's name, date of birth and a photograph. A passenger typically checks in with the OV chip card electronically at the departure station and checks out at the arrival station. The records of the suspect's OV chip card show that the card was used at 9.20am to check in at Amersfoort central and to check out at 10.31am in Amsterdam central station. These records thus fulfil the requirements for strong physical evidence as: (a) it is the suspect's personal card and is directly linked to him, and it also contains (b) time and (c) location information. The knowledge evidence in the present experiment consists of unique knowledge evidence. The suspect said that on his train journey someone pulled the emergency brake that resulted in an emergency stop of the train. After standing still for a couple of minutes, the train proceeded and managed to arrive on time in Amsterdam. The Dutch railways never publicised information about the incident.

All alibis and supportive evidence were rated on three dependent measures: (a) alibi believability; (b) evidence strength; and (c) the degree to which the alibi and its supportive evidence contributed to the innocence of the suspect. The participants rated the dependent measures on a Visual Analogue Scale (VAS; Luria, 1975) by marking their answer on a horizontal bar of 10 cm in length. The VAS consisted of two values below the horizontal bar at the beginning and the end of the bar. The value at the beginning was the lowest value (e.g. 'not at all believable'), and the end

point consisted of the highest value (e.g. 'completely believable').

We created two scales of dependent measures: one only about the alibi and the other consisting of the average of the other two items referring to the presented evidence. Because the supportive evidence was different in each presented alibi, separate scales (evidentiary value hereafter) were constructed per type of presented evidence. The α values for the evidentiary value varied between .83 and .93.

Lastly, the participants were asked which type of supportive evidence for an alibi was most important to them in general. They could choose from four options: witness evidence, physical evidence, knowledge evidence, or that all types of evidence were equally important. They were then fully debriefed and thanked for their participation. Along with the three general formulated hypotheses, we also expected police detectives to attribute lower believability scores to all presented alibis (Hypothesis 4) in line with previous studies (Culhane & Hosch, 2012).

2.2. Results

The data were analysed using general linear model (GLM) repeated measures analyses for the alibi believability measure and the evidentiary value. The within-subjects factor consisted of the supportive evidence, and the between-subjects factor was the group of participants (i.e. police detectives or laypersons). The results of these analyses showed that the assumption of sphericity was violated for the alibi believability measure and the evidentiary value. A Huynh–Feldt correction of the degrees of freedom was therefore applied ($\varepsilon = .95$ for alibi believability, $\varepsilon = .89$ for the evidentiary value).

For believability, no interaction between the groups and the type of evidence was found. A main effect of evidence was observed, $F(2.85, 399.25) = 58.51$, $p < .001$, $\eta_p^2 = .30$. An alibi without supportive evidence was found to be the least believable compared to alibis with supportive evidence, $p < .001$, supporting Hypothesis 1. No differences were observed between the types of supportive evidence for the alibi believability, $p > .05$.

No interaction was observed between the groups in regard to evidentiary value, but a main effect of supportive evidence was found, $F(2.68, 378.24) = 91.70$, $p < .001$, $\eta_p^2 = .39$. Similar to the believability measure, the alibi without supportive evidence obtained the lowest evidentiary value compared to the alibis supported with evidence, $p < .001$. No differences were observed between the types of evidence on the scale ($p > .05$), meaning that, in contrast to Hypothesis 3, Hypothesis 2 was supported. See Table 3 for an overview.

Most participants found all types of evidence equally important (44.8%), although a considerable percentage (34.4%) indicated that knowledge evidence was most important. A smaller proportion of the participants indicated that they found physical (14.7%) or witness evidence (6.3%) the most critical to support an alibi. There were no differences between detectives and laypersons on these ratings.

2.3. Discussion

There was no difference between detectives and laypersons in alibi evaluation and believability. Also, knowledge evidence turned out to be a useful addition to the value of alibis. The results even suggest that knowledge evidence has the same effects on alibi evaluation as the other types of supportive evidence. This was unexpected, as previous research has shown that physical evidence is considered to be stronger supportive evidence than witness evidence (e.g. Allison & Brimacombe, 2010; Jung et al., 2013). These deviating results could be explained by two limitations.

For the first time both laypersons and police detectives participated in an alibi study. Typically, undergraduate students are asked to evaluate the alibis, and therefore most of our understanding of alibi evaluation research comes from how students evaluate alibis. The selection of laypersons and police officers in

Table 3. Mean scores and standard deviations for the alibi believability measure and the evidence scale measure.

Supportive evidence	*M*	*SD*
Believability measure		
No evidence[a]	3.70	2.32
Witness evidence	6.46	2.46
Physical evidence	6.09	2.41
Knowledge evidence	6.18	2.30
Evidentiary value		
No evidence[a]	2.19	1.76
Witness evidence	5.82	2.68
Physical evidence	5.21	2.60
Knowledge evidence[b]	5.09	2.41

Note: $N = 144$. The dependent variable is bold.
[a]Significant at $p < .001$ compared to the other forms of supportive evidence.
[b]One missing value.

the experiment could possibly explain why our results are not consistent with previous findings. On one hand, police detectives are thought to have a sceptical attitude towards alibis (Dysart & Strange, 2012; Portnoy et al., 2020) and come to more guilty judgements than students when evaluating alibis (Culhane & Hosch, 2012). On the other hand, there is no difference in the assessment of alibi believability when comparing the alibi evaluations of police detectives and students (Culhane & Hosch, 2012). Similar to these results, we did not find any difference in the evaluation of alibis in terms of believability and the evidentiary value between police detectives and laypersons in the present experiment.

Second, the police detectives did not perceive the physical evidence as strong evidence. When they were debriefed, they indicated that although the OV chip card is directly linked to the suspect, it does not necessarily mean that he himself used the card on the day of the robbery. From their experience, they know that the ticket collector is more interested in verifying whether the passenger has paid for his journey than whether the card truly belongs to the passenger. The police detectives stated that they would have evaluated the alibis differently if CCTV recordings were presented showing that the suspect entered the train in Amersfoort and exited the train in Amsterdam. Their remarks might aid in understanding why the physical evidence was evaluated as weaker than we expected.

3. Experiment 2

To control for the first limitation in Experiment 1, we only included students in the second experiment. To control for the second limitation of Experiment 1, half of the participants received another type of strong physical evidence (i.e. CCTV recordings) while the other half of the participants received the same physical evidence (i.e. an OV chip card) as that in Experiment 1. CCTV recordings were included based on the feedback of the detectives and also because CCTV recordings are typically used as strong physical evidence in alibi research (Allison & Brimacombe, 2010; Jung et al., 2013). These two changes were made to make our results more comparable to previous research to determine the added value of knowledge evidence.

For the second experiment, we expected that knowledge evidence would not differ in its believability and the evidentiary value from witness evidence and physical evidence when it consisted of the OV chip card (Hypothesis 1a). When the physical evidence consisted of

CCTV recordings, we expected knowledge evidence to be evaluated similarly to witness evidence but yet weaker than the presented physical evidence (Hypothesis 1b). As in Experiment 1, we expected the alibi without supportive evidence to obtain the lowest score on both measures (Hypothesis 2). For the physical evidence, we expected CCTV recordings to score higher on both measures than the OV chip card (Hypothesis 3) as well as higher on both measures than witness evidence (Hypothesis 4).

3.1. Method

3.1.1. Design

The second experiment consisted of a 2 (item of physical evidence: OV chip card vs. video recordings) × 4 (supportive evidence: witness vs. physical vs. knowledge vs. none) mixed factorial design.

3.1.2. Participants

The 85 participants were undergraduate students from the Faculty of Psychology and Neuroscience at Maastricht University. They were recruited by means of flyers and received a research credit in return for their participation. Their ages varied between 19 and 53 years ($M = 22.08$, $SD = 3.91$), and most of them were female (76.5%). The standing ethical committee of the Faculty of Psychology and Neuroscience at Maastricht University approved the study.

3.1.3. Materials and procedure

The supportive physical evidence was changed for half of the participants compared to the materials used in Experiment 1. Half of the participants received the same materials as those in Experiment 1, while the other half of the participants read that the suspect's alibi was supported by CCTV recordings. On the recordings, the suspect was visible on the train platform around 9.20am in Amersfoort entering a train, and around 10.30am the suspect was again visible on the train platform in *Amsterdam* exiting a train. All of the other materials and procedures in Experiment 2 were identical to those in Experiment 1. The α values for the evidentiary value ranged between .88 and .94.

3.2. Results

The data were analysed using GLM repeated measures analyses for the alibi believability measure and the evidentiary value. The within-subjects factor consisted of the supportive evidence, while the between-subjects factor consisted of the item of physical evidence (i.e. OV chip card or CCTV recordings). The assumption of sphericity was violated for both the believability measure and the evidentiary value; thus a Huynh–Feldt correction was applied to the degrees of freedom ($\varepsilon = .90$ for the alibi believability, $\varepsilon = .94$ for the evidentiary value).

3.2.1. Believability measure

No interaction was observed, but a main effect of supportive evidence was found, $F(2.58, 214.31) = 56.72$, $p < .001$, $\eta_p^2 = .41$. All alibis with supportive evidence were rated to be more believable than the alibi without supportive evidence ($p < .001$), supporting Hypothesis 2. In addition, when the supportive evidence consisted of witness evidence, the alibi was rated to be more believable ($M = 6.95$, $SD = 2.04$) than when knowledge evidence was reported ($M = 6.25$, $SD = 2.34$), $p = .028$. The results of a one-way analysis of variance (ANOVA) revealed that video recordings and the OV chip card were found to be equally believable: $F(1, 83) = 0.003$, $p > .05$, which did not support Hypotheses 1a and 1b.

3.2.2. Evidentiary value

Depending on which physical evidence was presented, the evidence was rated differently by the participants, $F(2.82, 234.38) = 2.88$, $p = .040$, $\eta_p^2 = .03$. The participants rated the evidentiary value of CCTV recordings higher ($M = 7.19$, $SD = 2.61$) than the OV chip card

Table 4. Mean scores and standard deviations for the alibi believability measure and the evidence scale measure displayed for the OV chip card and video recordings as supportive evidence.

	Item of physical evidence					
	Personal public transportation card		CCTV recordings		Average	
Supportive evidence	*M*	*SD*	*M*	*SD*	*M*	*SD*
Believability measure						
No evidence[a]	3.41	2.38	3.78	2.47	3.60	2.42
Witness evidence[b]	7.02	2.13	6.88	1.97	6.95	2.04
Physical evidence	6.75	2.31	6.78	2.40	6.77	2.35
Knowledge evidence	6.21	2.12	6.30	2.52	6.25	2.34
Evidentiary value						
No evidence[a]	1.80	1.62	1.96	1.82	1.88	1.72
Witness evidence	5.95	2.46	6.00	2.41	5.97	2.42
Physical evidence[c]	5.84	2.58	7.19	2.61	6.55	2.67
Knowledge evidence	5.36	2.48	4.84	2.76	5.09	2.63

Note: $N = 85$. The dependent variable is displayed in bold.
[a]Significant at $p < .001$ compared to the other forms of supportive evidence.
[b]Significant at $p = .028$ compared to knowledge evidence.
[c]Significant at $p < .001$ compared to knowledge evidence when the physical evidence consisted of the personal public transportation card.

($M = 5.84$, $SD = 2.58$), $F(1, 83) = 5.73$, $p = .019$, $\eta_p^2 = .07$, supporting the third hypothesis. The evidentiary values were therefore compared per item of physical evidence using pairwise comparisons.

When the physical evidence consisted of the OV chip card, an alibi without supportive evidence received the lowest score compared to the alibis with supportive evidence ($p < .001$). No differences were observed between the types of supportive evidence ($p > .05$), similar to the results in Experiment 1. Furthermore, the participants rated the alibis supported by knowledge and witness evidence as comparable in evidentiary value, supporting Hypothesis 1a.

When the physical evidence consisted of CCTV recordings, an alibi without supportive evidence still obtained the lowest score compared to alibis supported with evidence ($p < .001$), supporting the second hypothesis. However, the participants now rated the evidentiary value of witness and physical evidence higher than knowledge evidence ($p < .001$), supporting Hypothesis 1b. Thus, knowledge evidence was considered to be weaker than witness and physical evidence, which leads to partial rejection of Hypothesis 1b. Although physical evidence yielded a higher average score ($M = 7.19$) than witness evidence ($M = 6.00$) this difference was just not significant ($p = .051$). Therefore, Hypothesis 4 is rejected (see Table 4 for an overview).

Most participants found all types of evidence equally important (47.6%), although a quarter of the participants (25.0%) indicated that knowledge evidence was most important. A smaller proportion of the participants indicated that they viewed physical evidence (17.9%) or witness evidence (9.5%) as the most important to support an alibi. There were no differences between the participants who received the OV card and those who received the CCTV recordings as physical evidence.

3.3. Discussion

The aim of the second experiment was to determine whether the atypical results in Experiment 1 could be explained by the

selection of participants and/or the type of physical evidence that was presented. Although both items of strong physical evidence were equally believable, we found that CCTV recordings obtained a higher evidentiary value than the OV chip card. When the physical evidence consisted of the OV chip card, no differences were observed for the alibi evaluation compared to the types of supportive evidence, similar to the results in Experiment 1.

The results of this study show that the students' results are in line with the results in Experiment 1. Students also rated all types of evidence as equally strong and believable. Yet, when the physical evidence consisted of another type of strong evidence, differences between the type of evidence were apparent. Hence, the fact that no differences are observed in the evaluation in Experiment 1 and in the present study when the OV chip card was presented can be explained by the selection of the physical evidence in both studies and is less likely due to the selection of participants. One could, however, argue that students more frequently use public transport than police detectives, which could have affected the results. However, that does not imply that the detectives are not familiar with this card. As stated in Experiment 1, the role of police detectives is very important in the criminal pre-trial investigation. Depending on whether or not they consider the alibi to be believable, the police officers may invest more or less time and energy investigating the suspect's potential involvement in the alleged crime. By doing so, they influence further procedures. It is therefore important to determine how they evaluate knowledge evidence in relation to CCTV recordings. Therefore, as a final experiment, Experiment 2 was replicated using police detectives.

4. Experiment 3

The last study was designed to determine how police detectives would evaluate knowledge evidence compared to both types of physical evidence, witness evidence and no supportive evidence. For this experiment we expected, based on the findings in Experiments 1 and 2, that an alibi without supportive evidence would be the least believable and obtain the lowest evidentiary value (Hypothesis 1). Second, we expected that CCTV recordings would hold a higher evidentiary value than the OV chip card. In addition, we expected CCTV recordings to receive a higher evidentiary value than witness evidence, in line with the results of previous research (Hypothesis 3), although no difference was observed between witness and physical evidence in Experiment 2. Third, we expected knowledge evidence to be rated comparable to witness evidence on both measures when the physical evidence consisted of the OV chip card based on the results in Experiment 1 (Hypothesis 4). Furthermore, when the physical evidence consisted of CCTV recordings, we expected a lower evidentiary value for knowledge evidence than for witness evidence based on the results in Experiment 2 (Hypothesis 5).

4.1. Method

A total of 402 police detectives were asked to fill out the same questionnaires as those described in Experiment 2. The questionnaires were handed out to all of the police detectives at a national police conference in April 2014 in Almere, the Netherlands, of whom 108 detectives filled out the questionnaire. Five participants had to be excluded as their answers were incomplete. Two additional participants were excluded for they reported to have never evaluated an alibi. The sample thus consisted of 101 police detectives. Fifty-eight participants filled out the questionnaire in which the OV chip card was presented as physical evidence; 43 participants completed the questionnaire in which the physical evidence consisted of CCTV recordings. The two versions of the questionnaire were randomly distributed among the attendees; however, more people completed the questionnaire in

Table 5. Participants' characteristics in percentages.

Sex	
Male	50.5
Female	49.5
Highest completed education	
Elementary school	0.0
Secondary school	23.8
Intermediate Vocational Education	25.7
College or University	50.5
Alibi evaluation in present job[a]	
Never	0.0
Less than once a month	50.0
Less than once a week	10.9
About once a week	27.2
Two to four times a week	3.3
About once a day	5.4
More than once a day	2.2
More than three times a day	1.1

Note: The dependent variable is displayed in bold.
[a]Nine missing values; only the valid percentages are presented.

which the OV chip card was reported as physical evidence. The participants' ages varied between 24 and 64 years ($M = 42.44$, $SD = 11.94$), and on average they evaluated an alibi once a month or less (45.5%) in their present job (see Table 5 for demographics). The α values for the evidentiary value varied between .80 and .95.

4.2. Results

The data were analysed using GLM repeated measures analyses for the alibi believability measure and the evidentiary value. The within-subjects factor was the supportive evidence, and the between-subjects factor consisted of the type of physical evidence (i.e. OV card or CCTV recordings).

For the believability measure, the assumption of sphericity was violated, and therefore the degrees of freedom were adjusted using the Huynh–Feldt correction ($\varepsilon = .97$).

4.2.1. Believability measure

Similar to the previous experiments, an alibi without supportive evidence was rated to be the least believable compared to those with supportive evidence, $F(2.91, 287.74) = 64.53$, $p < .001$, $\eta_p^2 = .40$. No other difference was observed between the types of supportive evidence in terms of believability, supporting Hypothesis 1. However, the results of a one-way ANOVA revealed that participants rated the CCTV recordings as more believable physical evidence ($M = 6.70$, $SD = 2.12$) than the OV chip card ($M = 5.42$, $SD = 2.11$), $F(1, 99) = 9.02$, $p = .003$, $\eta_p^2 = 0.89$.

4.2.2. Evidentiary value

Depending on the type of physical evidence that was presented, the evidence was rated differently in evidentiary value: $F(3, 297) = 7.31$, $p < .001$, $\eta_p^2 = .07$. The video recordings obtained a higher evidentiary value ($M = 6.92$, $SD = 1.87$) than the OV chip card ($M = 4.65$, $SD = 2.45$), $F(1, 99) = 25.86$, $p < .001$, $\eta_p^2 = .21$, supporting Hypothesis 2. Again, regardless of which type of physical evidence was presented, all alibis supported with evidence obtained a higher score than those without supportive evidence ($p < .001$). In contrast to Experiment 2, CCTV recordings received a higher score than knowledge

Table 6. Mean scores and standard deviations for the alibi believability measure and the evidence scale measure displayed for the OV chip card and video recordings as physical evidence.

	Form of physical supportive evidence					
	Personal public transportation card		Video recordings		Average	
Supportive evidence	*M*	*SD*	*M*	*SD*	*M*	*SD*
Believability measure						
No evidence[a]	3.11	1.57	3.80	2.11	3.40	1.84
Witness evidence	5.88	2.29	6.32	2.04	6.07	2.19
Physical evidence[b]	5.42	2.11	6.70	2.12	5.97	2.20
Knowledge evidence	5.92	1.74	6.05	1.95	5.97	1.82
Evidentiary value						
No evidence	1.89	1.35	2.09	1.60	1.97	1.46
Witness evidence[c]	5.23	2.33	5.80	2.34	5.47	2.34
Physical evidence	4.65	2.45	6.92	1.87	5.62	2.48
Knowledge evidence[d]	4.94	2.07	5.38	2.08	5.13	2.08

Note: The dependent variable is displayed in bold. $n = 101$.
[a]Significant at $p < .001$ compared to the other forms of supportive evidence.
[b]Significant at $p = .003$ between the personal public transportation card and video recordings.
[c]Significant at $p = .038$ compared to physical evidence but only when it consisted of video recordings.
[d]Significant at $p = .001$ compared to physical evidence but only when it consisted of video recordings.

evidence ($p = .001$) and witness evidence ($p = .038$), while no difference was observed between knowledge and witness evidence. In contrast, no differences were observed between the OV chip card and the other types of supportive evidence, supporting Hypotheses 3 and 4 but not Hypothesis 5. For an overview, see Table 6.

Most participants found all types of evidence equally important for a believable alibi (55.5%), although one in three participants (33.0%) indicated that knowledge evidence was most important. A smaller proportion of the participants indicated that they found physical evidence (9.0%) or witness evidence (3.0%) to be the most important to support an alibi. There were no differences between the participants who received the OV card and those who received the CCTV recordings as physical evidence.

4.3. Discussion

In the third experiment, we found that the alibi evaluation by police officers differed depending on which physical evidence was presented. When the physical evidence consisted of the OV chip card, the physical evidence was evaluated in a similar manner to witness and knowledge evidence. However, when the physical evidence consisted of CCTV recordings, the physical evidence appeared to be more believable and obtained a higher evidentiary value score than the OV chip card. Knowledge evidence was equally believable as witness and both items of physical evidence. In addition, knowledge evidence obtained a similar evidentiary value as witness evidence when the OV chip card was reported. We can conclude that the evaluation of alibis is also affected by which item of physical evidence is presented and that knowledge evidence is a relevant new type of supportive evidence for alibis. Both items of strong physical evidence used in the experiment meet the criteria for strong evidence, yet are evaluated differently.

5. General discussion

In the experiments, laypersons and police detectives rated knowledge evidence to be

equally believable as witness and physical evidence. Both students and police detectives evaluated knowledge evidence similarly in regard to its evidentiary value compared to the OV chip card as physical evidence. The results offer a valuable addition of the knowledge evidence to the alibi literature. The results suggest that the Olson and Wells (2004) taxonomy could be expanded to include the new type of evidence in addition to witness and physical evidence.

If knowledge evidence is added to the taxonomy, rules of thumb have to be established to determine the relative strength of knowledge evidence. For physical evidence, the strength is determined by the ease by which it can be fabricated or obtained from someone else. For knowledge evidence, the strength could perhaps be determined based on the uniqueness of the knowledge (i.e. how easily the suspect could have obtained the knowledge from another source). It can be argued that having knowledge of a well-known movie's plot like 'Harry Potter and the Philosopher's Stone' can be expected to be weaker evidence than the knowledge of a new movie that was released a week ago in the theatre. That assumption should be tested in future studies to determine whether knowledge evidence could be added to the existing alibi taxonomy (Olson & Charman, 2012; Olson & Wells, 2004).

Besides adjusting the taxonomy by adding knowledge evidence, the taxonomy could possibly be improved by revising the criteria to determine the strength of physical evidence. In the present studies, all types of supportive evidence were chosen based on the strongest items of evidence according to the taxonomy (Olson & Wells, 2004), using the criteria of Olson and Charman (2012). According to the criteria, strong physical evidence ' . . . contains time and place information and could be linked to the specific participant (such as a time card from an employer, a credit card receipt, or an airline ticket stub)' (Olson & Charman, 2012, p. 461). To our surprise, the OV chip card records are weaker supportive evidence than CCTV recordings, although both items meet the academic criteria for strong physical evidence. The police detectives argued that the records of the OV chip card do not necessarily confirm that the suspect used that card. A suspect's friend could have used the card to provide the suspect with evidence for the false alibi. The police detectives' argument also applies to the examples of strong supportive evidence provided by Olson and Charman (2012). The purchases registered on the suspect's credit card on the day and time of the alleged crime do not necessarily imply that the suspect himself used the card at that moment. At present, physical evidence is classified into three categories: strong, moderate and weak. The OV chip card records seem to fall between strong and moderate evidence. Furthermore, there is little variation in the presented items of physical evidence in alibi evaluation research to date. In most studies, CCTV recordings are used as strong physical supportive evidence (e.g. Allison & Brimacombe, 2010; Olson & Wells, 2004). In only a few studies has the presented physical evidence consisted of weaker evidence, such as a receipt or a time-stamped movie ticket (e.g. Jung et al., 2013; Pozzulo et al., 2015). In addition, there is a lack of studies in which a combination of evidence is reported (e.g. Dahl et al., 2009; Hosch et al., 2011; Sargent & Bradfield, 2004), although in reality one third of non-offenders offer a combination of evidence when asked for their alibi (Nieuwkamp et al., 2017). Given the impact of physical evidence on the evaluation of an alibi (Olson & Wells, 2004), we suggest conducting more research on the evaluation of different items of physical evidence.

In addition, nowadays most people possess a smartphone,[2] which also affects what evidence can be expected from a suspect when carrying or using a smartphone compared to the start of alibi research (Olson & Wells, 2004). For example, the registration of every cell phone with mobile providers via IMEI

numbers, the connection of a mobile device to a private or public WiFi network, the use of mobile data (4G, 5G), location information via the mobile device, and so on, add to the possibilities of providing physical evidence to support an alibi. In cases where the suspect claims to have been home alone in the absence of witness or physical evidence, these data may help to validate the alibi. Whereas such an alibi would remain unsupported before these technological advancements, nowadays these data can serve as supportive alibi evidence. Police detectives more often report these new types of physical evidence to support their alibis than do laypeople (Nieuwkamp, 2018). Therefore, it can be argued that police detectives are more aware of the presence of such physical evidence when they are asked for their whereabouts than are laypeople. This might explain why detectives put such high demands on a suspect's alibi (Dysart & Strange, 2012) before they find it believable (Nieuwkamp, 2018).

5.1. Further research

It should be noted that the students in Experiment 2 evaluated CCTV recordings as equally believable as the OV chip card, while the police detectives considered the CCTV recordings to be more believable than the OV chip card. It could therefore be concluded that perhaps the criteria to determine the strength of the supportive evidence only applies for students and not for police detectives. It raises doubt as to how valid the criteria of the taxonomy are for determining the strength of physical evidence in practice. As stated above, most of the research on alibis has been conducted using students as participants, although recent results suggest that alibi evaluation differs between police detectives and students (Culhane & Hosch, 2012; Dysart & Strange, 2012; Eastwood et al., 2016). The results of the present experiments add to those recent findings. We therefore call for more alibi research involving police detectives as participants in order to better understand how they evaluate alibis in practice, especially since they are the first to encounter and evaluate the alibis in practice (Culhane & Hosch, 2012; Nieuwkamp, 2018). Taking these reflections into account, further research should focus on adapting the criteria for determining the strength of physical evidence to correspond to the way in which alibis are evaluated in practice in combination with newer evidence from cell phones. In addition, in future research these technological advancements should be taken into account determining whether the current criteria for establishing the strength of supportive evidence (Olson & Charman, 2012) still apply. One could argue that laypeople are more aware of the presence of technological evidence to support their alibis as that technology is more visible in our daily lives or at least police detectives are aware of the presence of such evidence (Nieuwkamp, 2018) and it is therefore of interest to study in more detail.

5.2. Limitations

In the present three experiments, the same methodology was applied in which the four suspects reported the same alibi, but the alibis were supported with a different type of supportive evidence or no supportive evidence at all. A limitation could therefore be that the alibis were too similar, which could have affected the alibi evaluation. We tried to minimise that limitation by randomising the reported alibis among all participants. In addition, because we wanted to determine the added value of knowledge evidence, our experiments were inspired by a study by Olson and Wells (2004) in which the alibis were kept constant, and only the supportive evidence varied among the suspects.

5.3. Conclusion

We found that the taxonomy does not align with how alibis are evaluated in practice because certain items of physical evidence that police detectives come across in practice do not fully fit today's taxonomy. In addition, the

set criteria to determine the strength of evidence do not align with how alibis are evaluated in practice. This raises doubts about the applicability of the alibi taxonomy in practice and how applicable our knowledge is to the field. Therefore, more refinement is needed to determine how various items of physical evidence are evaluated both in the lab and in practice. Furthermore, our results demonstrate the value of adding knowledge evidence to the alibi evaluation literature. Jeremy Bentham once stated: 'The field of evidence is no other than the field of knowledge' (Twining, 2019, p. 23). Based on the results of these experiments, we can conclude that the opposite also seems to be true for the evaluation of alibis: the field of knowledge is no other than the field of evidence.

Ethical standards

Declaration of conflicts of interest

Ricardo Nieuwkamp has declared no conflicts of interest

Robert Horselenberg has declared no conflicts of interest

Peter J. 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 and/or national research committee [the standing ethical committee of the Faculty of Psychology and Neuroscience at Maastricht University[3]] 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.

Notes

1. A journey of about 1 hour and 62 km.
2. In 2021, 87% of all the inhabitants in the Netherlands possessed a smartphone (source: https://opendata.cbs.nl/#/CBS/nl/dataset/84888NED/table?ts=1655812656197, consulted 21 June 2022) and 85% in the United States of America (source: https://www.statista.com/statistics/219865/percentage-of-us-adults-who-own-a-smartphone/, consulted 21 June 2022).
3. ECP-124 11_01_2013 & ECP-124 11_01_2013_A1

Acknowledgements

We would like to thank Kim Michelle Rietman and Michelle Huygen for their help in collecting the data.

Funding

This study was financially supported by the Netherlands Organisation for Scientific Research [NWO, grant number 404-10-349].

ORCID

Ricardo Nieuwkamp http://orcid.org/0000-0001-7101-0052

References

Allison, M. (2022). Alibi believability. In J. D. Behl & M. R. Kienzle (Eds.), *Alibis and corroborators: Psychological, criminological, and legal perspectives* (pp. 37–53). Springer.

Allison, M., & Brimacombe, C. (2010). Alibi believability: The effect of prior convictions and judicial instructions. *Journal of Applied Social Psychology*, *40*, 1054–1084. https://doi.org/10.1111/j.1559-1816.2010.00610.x.

Behl, J.D. & Kienzle, M.R. (2022). *Alibis and corroborators: psychological, criminological, and legal perspectives*. Springer.

Cardenas, S. A., Crozier, W. E., & Strange, D. (2022). The effect of erroneous alibis. In J. D. Behl & M. R. Kienzle (Eds.), *Alibis and corroborators: Psychological, criminological, and legal perspectives* (pp. 125–147). Springer.

Crozier, W. E., Strange, D., & Loftus, E. F. (2017). Memory errors in alibi generation: How an alibi can turn agaist us. *Behavioral Sciences and the Law*, *35*, 6–17. https://doi.org/10.1002/bsl.2273

Culhane, S. E., & Hosch, H. M. (2004). An alibi witness's influence on mock jurors' verdicts. *Journal of Applied Social Psychology*, *34*, 1604–1616. https://doi.org/10.1111/j.1559-1816.2004.tb02789.x.

Culhane, S. E., & Hosch, H. M. (2012). Changed alibis: Current law enforcement, future law enforcement, and layperson reactions. *Criminal Justice and Behavior*, *39*, 958–977. https://doi.org/10.1177/0093854812438185.

Culhane, S. E., Hosch, H. M., & Kehn, A. (2008). Alibi generation: Data from U.S. Hispanics and U.S. non-Hispanic Whites. *Journal of Ethnicity in Criminal Justice*, *6*, 177–199. https://doi.org/10.1080/15377930802243395.

Culhane, S. E., Kehn, A., Horgan, A. J., Meissner, C. A., Hosch, H. M., & Wodahl, E. J. (2013). Generation and detection of true and false alibi statements. *Psychiatry, Psychology and Law*, *20*, 619–638. https://doi.org/10.1080/13218719.2012.729018.

Dahl, L. C., Brimacombe, C. A., & Lindsay, D. S. (2009). Investigating investigators: How presentation order influences participant-investigators' interpretations of eyewitness identification and alibi evidence. *Law and Human Behavior*, *33*, 368–380. https://doi.org/10.1007/s10979-008-9151-y.

Dysart, J. E., & Strange, D. (2012). Beliefs about alibis and alibi investigations: A survey of law enforcement. *Psychology, Crime and Law*, *18*, 11–25. https://doi.org/10.1080/1068316X.2011.562867.

Eastwood, J., Snook, B., & Au, D. (2016). Safety in numbers: A policy-capturing study of the alibi assessment process. *Applied Cognitive Psychology*, 260–269. https://doi.org/10.1002/acp.3200.

Garrett, B. L. (2011). *Convicting the innocent*. Harvard.

Hosch, H. M., Culhane, S. E., Jolly, K. W., Chavez, R. M., & Shaw, L. H. (2011). Effects of an alibi witnesse.g.s relationship to the defendant on mock jurors' judgments. *Law and Human Behavior*, *35*, 127–142. https://doi.org/10.1007/s10979-010-9225-5.

Jung, S., Allison, M., & Bohn, L. (2013). Legal decision-making on crimes involving an alibi. *Applied Psychology in Criminal Justice*, *9*, 45–58.

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*, 187–203. https://doi.org/10.1023/A:1022599230598.

Luria, R. E. (1975). The validity and reliability of the visual analogue mood scale. *J Psychiatr Res*, *12*, 51–57. https://doi.org/10.1016/0022-3956(75)90020-5.

Meissner, C. A., & Kassin, S. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior*, *26*, 469–480. https://doi.org/10.1023/A:1020278620751.

Nieuwkamp, R. (2018). *Where I was and how I will prove it. On the believability of alibis*. Gildeprint.

Nieuwkamp, R., Horselenberg, R., & van Koppen, P. (2018). True and false alibis among prisoners and their detection by police detectives. *Psychiatry, Psychology, and Law*, *25*, 902–921. https://doi.org/10.1080/13218719.2018.1482570

Nieuwkamp, R., Horselenberg, R., & Van Koppen, P. J. (2017). The illusion of the perfect alibi: Establishing the base rate of non-offenders' alibis. *Journal of Investigative Psychology and Offender Profiling*, *14*, 23–42. https://doi.org/10.1002/jip.1470.

Nieuwkamp, R., Mergaerts, L. (2022). Alibi assessment and believability across different legal systems and cultural contexts. In J. D. Behl & M. R. Kienzle (Eds.), *Alibis and corroborators: Psychological, criminological, and legal perspectives* (pp. 75–86). Springer.

Olson, E. A., & Charman, S. D. (2012). 'But can you prove it?' – Examining the quality of innocent suspects' alibis. *Psychology, Crime and Law*, *18*, 453–741. https://doi.org/10.1080/1068316x.2010.505567.

Olson, E. A., & Wells, G. L. (2004). What makes a good alibi? A proposed taxonomy. *Law and Human Behavior*, *28*, 157–176. https://doi.org/10.1023/B:LAHU.0000022320.47112.d3.

Portnoy, S., Hope, L., Vrij, A., Ask, K., & Landström, S. (2020). Beliefs about suspect alibis: A survey of lay people in the United Kingdom, Israel, and Sweden. *The International Journal of Evidence & Proof*, *24*, 59–74. https://doi.org/10.1177/1365712719873008

Pozzulo, J. D., Pettalia, J. L., Dempsey, J. L., & Gooden, A. (2015). Juvenile offenders on trial: does alibi corroboration evidence and defendant age interact to influence jurors' perceptions and verdicts? *Psychiatry, Psychology and Law*, *22*, 224–234. https://doi.org/10.1080/13218719.2014.937518.

Sargent, M. J., & Bradfield, A. L. (2004). Race and information processing in criminal trials: Does the defendant's race affect how the facts are evaluated? *Personality and Social*

Psychology Bulletin, *30*, 995–1008. https://doi.org/10.1177/0146167204265741.

Sauerland, M., Mehlkoph, S., Krix, A. C., & Sagana, A. (2016). Deceiving suspects about the content of their alibis: Consequences for truthful and untruthful suspects. *Journal of Forensic Practice*. https://doi.org/10.1108/JFP-10-2014-0042

Strange, D., Dysart, J. E., & Loftus, E. F. (2014). Why errors in alibis are not necessarily evidence of guilt. *Zeitschrift für Psychologie*, *222*, 82–89. https://doi.org/10.1027/2151-2604/a000169.

Sullivan, T. P. (1971). Presentation of the defense. *Journal of Criminal Law, Criminology, and Police Science*, *62*, 153–172. https://doi.org/10.2307/1141875

Twining, W. (2019). Bentham's theory of evidence: Setting a context. *Journal of Bentham Studies*, *18*, 20–37. doi: https://doi.org/10.14324/111.2045-757X.047.

University Press. Haan, B. (2009). *De Deventer Moordzaak. Het complot ontrafeld* [The Deventer murder case. The conspiracy unraveled]. Nieuw Amsterdam Uitgevers.

van Koppen, M. V., & Nieuwkamp, R. (2017). Alibi's [Alibis]. In P. J. van Koppen, J. W. de Keijser, R. Horselenberg, & M. Jelicic (Eds.), *Routes van het Recht: Over de Rechtspsychologie* [Routes of the law: About legal psychology] (pp. 249–260). Boom Juridisch.

van Koppen, M. V., & van Koppen, P. J. (2010). Alibi's [Alibis]. In P. J. van Koppen, M. Merckelbach, M. Jelicic, & J. W. de Keijser (Eds.), *Reizen met mijn rechter: Psychologie van het Recht* [Traveling with my judge: Psychology of the law*/* (pp. 439–448). Kluwer.

Verschuere, B., Bogaard, G., & Meijer, E. (2021). Discriminating deceptive from truthful statements using the verifiability approach: A meta-analysis. *Applied Cognitive Psychology*, *35*, 374–384. https://doi.org/10.1002/acp.3775

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. John Wiley.