



# “Blood, Bucks and Bias”’: Reliability and biasability of crime scene investigators’ selection and prioritization of blood traces

Moa Lidén<sup>a,b,c,\*</sup>, Mohammed A. Almazrouei<sup>a,b,d,e</sup>

<sup>a</sup> Department of Security and Crime Science, University College London, 35 Tavistock Square, London WC1H 9EZ, UK

<sup>b</sup> Centre for the Forensic Sciences, University College London, 35 Tavistock Square, London WC1H 9EZ, UK

<sup>c</sup> Law Faculty, Uppsala University, Munken 1, Trädgårdsgatan 20, 753 09 Uppsala, Sweden

<sup>d</sup> Yale Decision Neuroscience Lab, Yale University, 300 George St, New Haven, CT 06511, USA

<sup>e</sup> Forensic Evidence Department, Abu Dhabi Police General Headquarters, Abu Dhabi 253, United Arab Emirates

## ARTICLE INFO

### Keywords:

Crime scene

Blood

Trace

Bias

Reliability

Triage

## ABSTRACT

When crime scene investigators (CSIs) encounter crime scenes with large volumes of blood, some selection and prioritization is often needed, and this will impact on what blood is and is not available for forensic analysis. What factors influence CSIs decision making process is largely unknown. This study examines the effects of awareness of limited resources and irrelevant contextual case information indicating either a homicide or a suicide on CSIs collection of blood traces. To this end, two scenario-based experiments with CSIs and novices were conducted. Overall, the results suggest that even when CSIs decisions are made under identical conditions, their trace selection varies both when it comes to numbers and locations. Furthermore, awareness of limited resources made CSIs collect fewer traces and their selections also varied following the contextual case information, showing similarities and differences with novices. Since blood traces can be used to establish both activity and identity the findings can have important implications for the subsequent investigation as well as trial.

## 1. Introduction

In previously published literature, CSIs have sometimes been referred to as ‘forensic dustmen’ [1] or ‘evidence collectors’ [2], which may imply they only mechanically assemble traces from the scene. However, this phrasing greatly understates their role as active agents in creating the crime scene evidence. Through selection and prioritization, CSIs transform traces into evidence that is potentially useful for the criminal investigation and proceedings [3]. While words such as ‘create’ and ‘transform’ seem to imply that CSIs consciously alter or make up evidence, this is not the intended meaning. Rather, CSIs *make sense* of the crime scene [4] by delimiting the potentially meaningful from the meaningless and making necessary prioritizations under uncertainty, time pressure, limited resources, stress etc. Hence, collecting traces from a crime scene and making prioritizations as regards which should be sent off for forensic analysis, are indeed discretionary tasks.

The selection of traces at the scene and the prioritization of traces for forensic analysis are logically and chronologically separate processes. However, many CSIs are aware, as they select the traces, that the forensic labs have limitations as to the number of traces they will

analyze. The mere knowledge of the available resources may influence the trace selection process. Across different jurisdictions, it varies to what extent such limitations exist and whether they are explicit/implicit and binding/non-binding. Yet, since all criminal investigations are restrained in terms of time and money, the existence of such limitations is a common and continuous denominator in criminal investigations globally. Today, it is unknown how awareness of such limitations impact on the work conducted by the CSIs while still at the scene. Concerns have been raised that such awareness restricts CSIs trace collection to primarily or only comprehend the traces that, given what is known at the time, seem to have the greatest potential in bringing the investigation forward e.g. because they would be useful for a prosecutor when pressing charges and/or arguing the case in Court [5]. However, as the investigation proceeds or as a case is reopened, new circumstances or evidence may shed new light on what traces would have been most relevant to secure from the scene. Yet, these traces may no longer be available since they were never secured from the scene which by then has been released.

Crime scenes can potentially contain a range of traces, that is, physical remnants of criminal activity which later on in the process will

\* Corresponding author at: Department of Security and Crime Science, University College London, 35 Tavistock Square, London WC1H 9EZ, UK.

E-mail address: [moa.liden@jur.uu.se](mailto:moa.liden@jur.uu.se) (M. Lidén).

become evidence to be presented in a Court of law [6,7]. However, blood is a common type of trace, not the least in cases of suspected violent crimes. Blood can be informative for activity-level and identity questions during investigations. The activity will be inferred from bloodstain pattern analysis (BPA), that is, categorization of bloodstain patterns based on the blood-shedding events which is believed to have given rise to the patterns [8–10]. Today, human sources of error pertaining to BPA has received fairly extensive attention by researchers and it is well established that there are a number of issues with current methods of bloodstain pattern classification [8–10]. Clearly, this can cause issues in the understanding of whether a crime has been committed at all, and if so, what crime.

Additionally, collection and preservation of blood is also essential because it enables typing of the blood or DNA analysis, that is, identification of the individuals likely to have been present at the scene. This identification process is also important for criminal investigations. It has, nevertheless, received far less attention when it comes to the potential human sources of error, compared to the determination of what activity has taken place at the scene. At high volume blood scenes, CSIs will be unable to collect and preserve all blood traces and this constitutes a risk both of false positives and false negatives. For example, blood from the victim and another individual may have been identified. Given that this other individual is the only individual known to have some kind of involvement with the crime, he/she may easily become the suspect [11] and also be convicted of the crime, while, in fact, there was also another or several other individuals present at the scene who are the real perpetrators (i.e. a false positive outcome). Likewise, it could be that only the victim's blood is identified at the scene and based on this alone or together with other circumstances, it is presumed that the individual has committed suicide. Yet, in fact, it was not a suicide but a homicide and had the blood at the scene been collected and prioritized differently, the presence of another individual at the scene would have become apparent to investigators (i.e. a false negative outcome).

Following the discretionary nature of CSIs selection and prioritization of blood traces, it would be unsurprising if different crime scene investigators examining *identical* crime scenes would make different observations and also draw different conclusions regarding which blood traces should be selected and prioritized for forensic analysis. This is here referred to as lacking *Between CSI Reliability*. This term stems from an application of Dror's Hierarchy of Expert Performance (HEP) in the context of crime scene investigation, noting that the framework describes expert reliability in a wider sense [12]. Furthermore, what traces are collected from the scene may in practice be a result of an interaction between limited investigative resources ('bucks') and contextual case information that cause CSIs to focus one-sidedly on one type of trace rather than others ('bias'). In this research, we examine how both the 'bucks' and 'bias' variables can impact on the selection and prioritization of blood traces from a high-volume, ambiguous blood scene (i.e. ambiguous as regards whether the deceased has committed suicide or been the victim of a homicide). Specifically, we examine whether knowledge of a case hypothesis or that only a limited number of swabs can influence CSIs decision making. This is here referred to as *Between CSI Biasability* also following an application of the HEP framework [12].

Hence, the purpose of this research is to examine *Between CSI Reliability* and *Biasability* in the selection of blood traces from a crime scene as well as the prioritization of the traces for forensic analysis. To this end, two experiments were conducted to examine the effect of two variables, 1) Resources: Unlimited v. Limited, and 2) Context: Homicide v. Suicide on, for example, the number and location of blood traces that are collected from crime scenes. For **Experiment 1**, data was collected from experienced CSIs operating in Sweden, Norway and the UK. For **Experiment 2**, using the same experimental design as Experiment 1, data was instead collected from novices. Experiment 2 provided baseline data for how individuals without the relevant expertise would behave. The data from each experiment was analyzed separately, and also compared to each other to see whether and how the behaviors of novices

were different from that of experts.

### 1.1. *Between CSI reliability*

CSIs may examine *identical* crime scenes, yet they may make different observations and also draw different conclusions [12,13]. Disagreement between different CSIs regarding which blood traces should be collected from the scene and prioritized for forensic analyses can have different explanations, whereof *inherent task difficulty* is one [14]. Crime scenes are often both complex and ambiguous, and decisions are often made at early stages when relatively little is known about what might have occurred. Disagreement that stems from inherent task difficulty is important disagreement that, ideally, should be integrated into the decision-making process. For example, in some high-volume blood scenes where deceased individuals with multiple stab wounds are encountered, it can be incredibly difficult to decide whether the deceased has committed suicide or become the victim of a homicide. While suicide by self-stabbing is fairly unusual [15–20], there may be circumstances such as hesitation marks [21–23], primarily horizontal stab wounds [24,25], documented suicidal tendencies [26], and lack of signs of struggle or defense wounds [15,16,27] that may indicate suicide. Yet, in these situations, the sheer volume of the blood and the number of stab wounds could make it questionable whether the harm could be self-inflicted, despite drugs, mental illness etc. [15,17,26,28–33]. However, in an applied setting, disagreement stemming from *task difficulty* is not necessarily easy to distinguish from other sources of disagreement such as *unstandardized methods* [14,34] which is likely to be applicable to BPA which lacks a discipline standard in the methodology, and some BPAs may even struggle to articulate the methodology they employ [35,36]. Furthermore, there may be *individual evaluator differences* [14,37,38], that is, patterns of stable individual differences among evaluators, as opposed to mere inaccuracy or random variation, which contributes to diverging opinions. This can be due to, for example, *differences in training, experience or theoretical starting points* as well as different decision thresholds, which, in turn, may be linked to variations in personality and/or attitudes [14,39–42]. Disagreement between forensic experts examining identical evidence has been noted in a wide range of fields including for example age estimation [43], legal sanity opinions [44–46], sharp bone trauma diagnosis [47] as well as bloodstain pattern analysis (BPA) [48,49].

It is also clear that disagreement between BPA analysts does not only occur in controlled laboratory settings but also in real criminal cases, such as *Aguirre-Jarquín v. State of Florida* [50] and *Camm v. State* [51]. In both of these high profile cases BPA analysts were at the center of controversial arguments relating to the explanation of the mechanisms that produce very small bloodstains on clothing. When disagreeing experts are pitted against each other in Court, this often forces judges to decide on scientific issues about which the judges normally lack expertise [43] but, if disagreement is never detected or communicated, this is likely to pose an even larger threat to the administration of justice. To turn expert disagreement into an asset rather than a threat to the administration of justice, platforms for such disagreement need to be created, for example by having more than one CSI examining the scene. If CSIs regularly disagree on which traces should be collected and/or prioritized, then having two or possible three CSIs work together may result in productive conflict and ultimately a wider more comprehensive crime scene investigation. Against this background, it is noteworthy that no research up until today has examined the prevalence and extent of *Between CSI Reliability* when it comes to using blood for identification, that is which blood traces are secured and sent off for forensic analyses. If there is disagreement also in this process, on top of the classification of bloodstain patterns originating in certain activities, then the crime scene investigation may be more or less like a lottery both when it comes to identifying whether a criminal act has been committed (activity) and which individuals were at the scene (identification). In other words, the result of the crime scene investigation, that is, which traces are or are not

collected/prioritized, may be more dependent on the specific CSI than the characteristics of the scene. This can clearly have important consequences not only for the focus of the subsequent investigation but also the Court proceedings and the odds of a wrongful conviction or acquittal. Similarly, if a wrongfully convicted individual applies for judicial review claiming for example that the real perpetrator's blood must have been left at the crime scene, the possibilities to evaluate such claims are practically zero if those samples were never collected from the crime scene in the first place.

### 1.2. Between CSI biasability

There is potentially a range of circumstances which can have a biasing effect on CSIs. Knowledge of case hypothesis is one such factor which has been shown to bias Dutch CSIs selection and interpretation of traces from a mock crime scene [52]. Similarly, it has a documented biasing effect in relation to BPA analysts pattern classifications both on rigid non-absorbent surfaces [53] and fabric surfaces [54], with higher overall error rates following the biasing information in the latter (23.00 %) than the former (13.10 %) category [54]. It can and has been argued that contextual information is a necessary part of BPA as patterns are analyzed in the context of a case with the objective to assist with the reconstruction of events [55] but if the methodology is unreliable, integrating case information may create a fertile breeding ground for *confirmation bias* [56]. Among experienced BPA analysts from New Zealand and Australia who were asked to think aloud as they classified patterns from two homicide cases and who could request items of contextual information, most analysts requested pathology reports (e.g., to determine the mechanism for blood deposition or to confirm suspect blood-letting injuries) and scene photographs (e.g. in order to put the scene in context and to understand the totality of the scene [55]). Analysts were often reluctant to conduct their analyses based solely on the pattern itself and many also stated they wanted the contextual information because of routine, that is, because it is usually accessible to them. Especially in situations where data from a bloodstain pattern is limited, it may be tempting for analysts to seek out contextual information in order to help reach a decision. Contextual bias is most likely to occur when data are ambiguous [57–60].

Furthermore, there are often routine approaches to trace collection. In other words, knowledge of previous cases allows targeting the most promising kinds of forensic case data in the context of a series or a typical crime phenomenon, often referred to as intelligence-led crime scene processing [61]. For example, when CSIs recognize a situation as belonging to a typical phenomenon, such as a house burglary, they would generally collect latent biological material at a commonly identified point of contact (on the window of the premises) [62]. This is due to an assumption that the specific *modus operandi* used by perpetrators in such situations creates exchange of biological material at this specific point. Hence, swabbing at this point is part of a routine approach when processing such scenes. Such routine approaches can be more or less explicit, sometimes even dictated by applicable guidelines, but are fairly common as noted for example in a sample of Swiss Crime Scene Examiners (CSE) ( $N = 73$ ), among whom 68 % reported that they often to always base the selection of traces to be exploited on the available intelligence [61,63]. While the validity of such approaches can be questioned in themselves, they are conscious strategies and therefore not in themselves a bias. However, if the initial classification of a situation triggers a routine approach, this is likely to bias CSI selection and prioritization of traces, even if the scene may very well contain other traces suggesting that the initial classification is incorrect. The ability to notice such traces may be hampered by the initial classification, thus resulting in a confirmation bias [52,64]. Thus, the use of crime intelligence in crime scene investigation immediately resonates with cognitive biases [65–68] since CSIs, in accordance with the crime intelligence, would generally tend to focus on what they or their organization know and prioritize. Along these lines, although not specifically researched yet, it

is likely that limited investigative resources as well as time pressure resulting from organizational priorities can bias CSIs decision making. Typically, volume crime scene examinations take place within tighter budgetary and resource limits than those of major crime and this can affect whether a CSI attend a crime scene at all, the time available to examine the scene, what is collected and which of the forensic objects removed from the scene, if any, are submitted for laboratory analysis [69,70]. However, also with major crimes there are limitations as to how much evidence can be analyzed forensically, and typically CSIs are aware of such limitation. Knowledge of limited investigative resources may, for example, lead CSIs to engage in hypothetical prospective reasoning about which traces will be selected for forensic analysis and then they only or primarily collect those traces. This is not necessarily due to laziness but also rational considerations regarding the perceived usefulness and quality and nature of the trace, or odds that a trace will be capable of bringing the investigation forward, e.g. what the trace can possibly prove [4,63]. Yet, in the process of assessing usefulness etc. it is likely that knowledge of case hypotheses will have an impact.

## 2. Experiment 1: Crime scene investigators

### 2.1. Method

#### 2.1.1. Participants

Study participants were 37 crime scene investigators (CSIs), 19 (51.40 %) men and 18 (48.60 %) women, from Sweden (11), Norway (13) and the UK (13). The participants' ages varied between 23 and 68 years ( $M = 43.16$ ,  $SD = 10.80$ ) and their length of experience as CSIs ranged from 1 and 32 years ( $M = 10.49$ ,  $SD = 8.60$ ). For participants to be included in the study (i.e., the selection criteria) they needed to be "crime scene investigators, actively examining cases, including bloody scenes. Among the CSIs, 14 (37.80 %) had no formal training on bloodstain pattern analysis (BPA), 15 (40.50 %) had basic BPA training (40hrs or less), and 8 (21.60 %) had advanced BPA training (more than 40hrs).

The CSIs were approached through personal contact with the heads of staff or other representatives from within units in police organizations where crime scene investigators work. This study was approved by UCL Research Ethics Committee 17415/001 and the Swedish Ethics Committee 2019–03878 prior to commencement of data collection.

### 2.2. Design

The experiment had a 2 (Resource limitations: yes vs. no)  $\times$  2 (Context: homicide vs. suicide) Mixed Subjects Design. The Resource limitations variable was varied within participants whereas the Context variable was varied between participants.

### 2.3. Material

The participants conducted their assessments on the basis of photographs from a real scene which was originally investigated in a national jurisdiction. The photos depicted a high-volume blood scene; more specifically a male body inside a toilet cubicle of a public bathroom. Inside the cubicle, a man had been found deceased and there was also broken glass inside the cubicle.

The photos were high quality color photos, which were incorporated into Qualtrics software. The program enabled participants to zoom in and out on different aspects of the photos. Initially, in the homicide condition, participants were exposed to a photo of the deceased's neck, with deep stab wounds visible. Participants in the suicide condition were instead shown a photo of the deceased's left wrist showing non-deep marks consistent with hesitation marks. Thereafter, all participants were exposed to eleven photos of the scene. They were only asked to make assessments in relation to four of these photos, while the remaining photos displayed more details of specific parts of the scene as

a means to facilitate the participants’ decision making. The participants marked directly in the four photos from where they wanted to collect blood traces. This produced a heat map illustrating overall trends in the collection of blood traces.

The choice of photographs as stimulus material was preferred over setting up a mock crime scene for four reasons. Firstly, also in real criminal investigations, assessments are often made on the basis of photos [55], not the least in released and cold cases [71]. Secondly, it facilitated compliance with COVID-19 regulations since all data collection could take place online and participants could maintain social distancing without complications. Thirdly, the conditions in the photographs would clearly remain exactly the same even if several other CSIs had examined them before, unlike with a mock crime scene. Fourthly, the usage of photographs enabled effective tracking of where the CSIs secured blood traces, as indicated by their markups in the photos.

2.4. Procedure

The procedure followed four steps; the manipulation step (Step 1), the trace collection under unlimited resources step (Step 2), the trace collection under limited resources step (Step 3) and the supplementary questions step (Step 4).

In Step 1; the manipulation step, participants were randomized via Qualtrics to either the homicide or the suicide condition. All participants

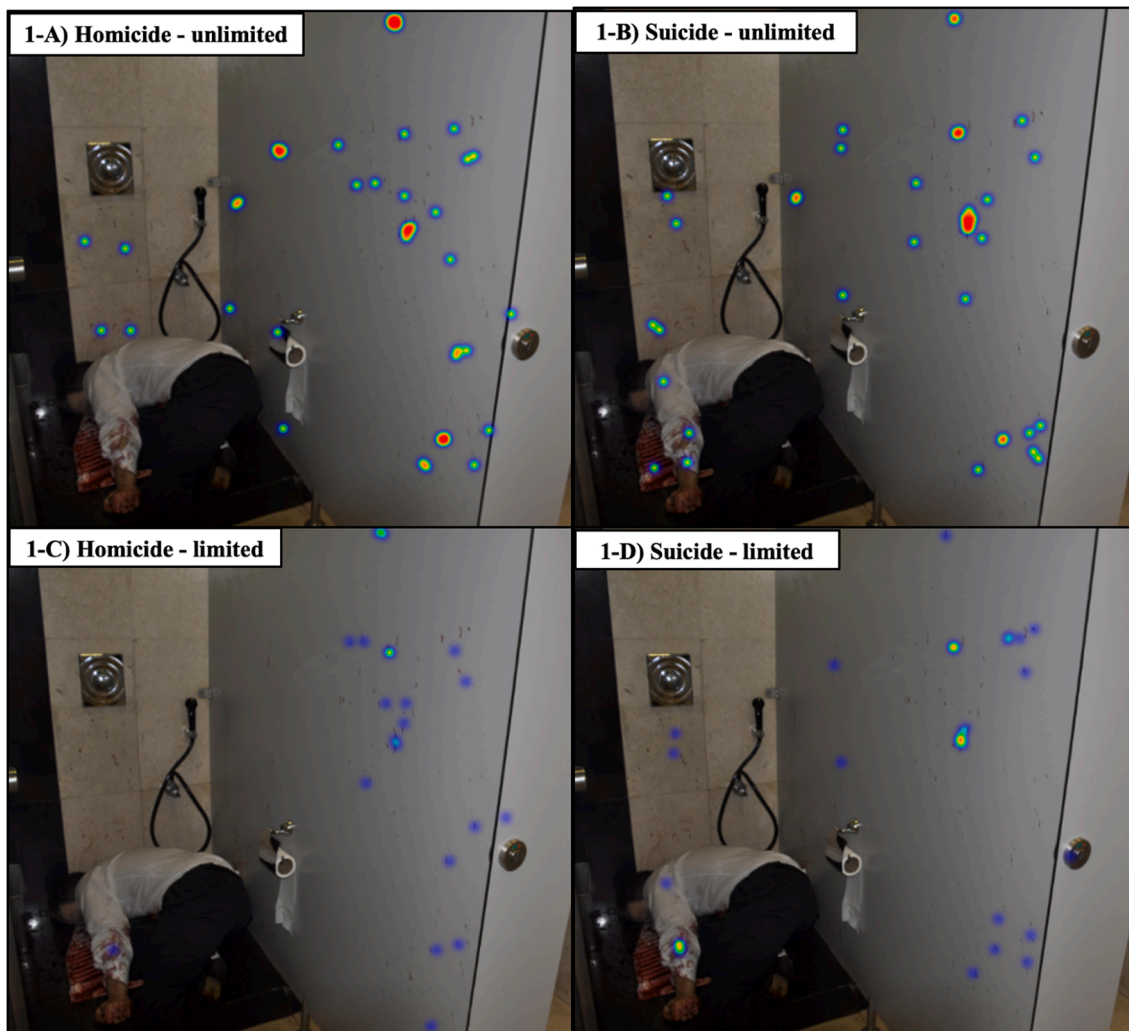
were asked to collect blood from a scene in which a deceased person had been found.

In the homicide condition, participants were told that: 1) On the way to the scene, a police officer provided information that they were going to a murder scene with high volumes of blood, and 2) the detective on site said that a witness had heard two men who appeared to be fighting in the bathroom. One of the men had shouted to the other “I will kill you, bastard!”. Also, in this condition, the photo of the deceased man’s neck with deep stab wounds was presented. For clarity, the photo of the neck with stab wound was only shown in the homicide condition. Similarly, the photo of the wrist with hesitation marks was only shown in the suicide condition.

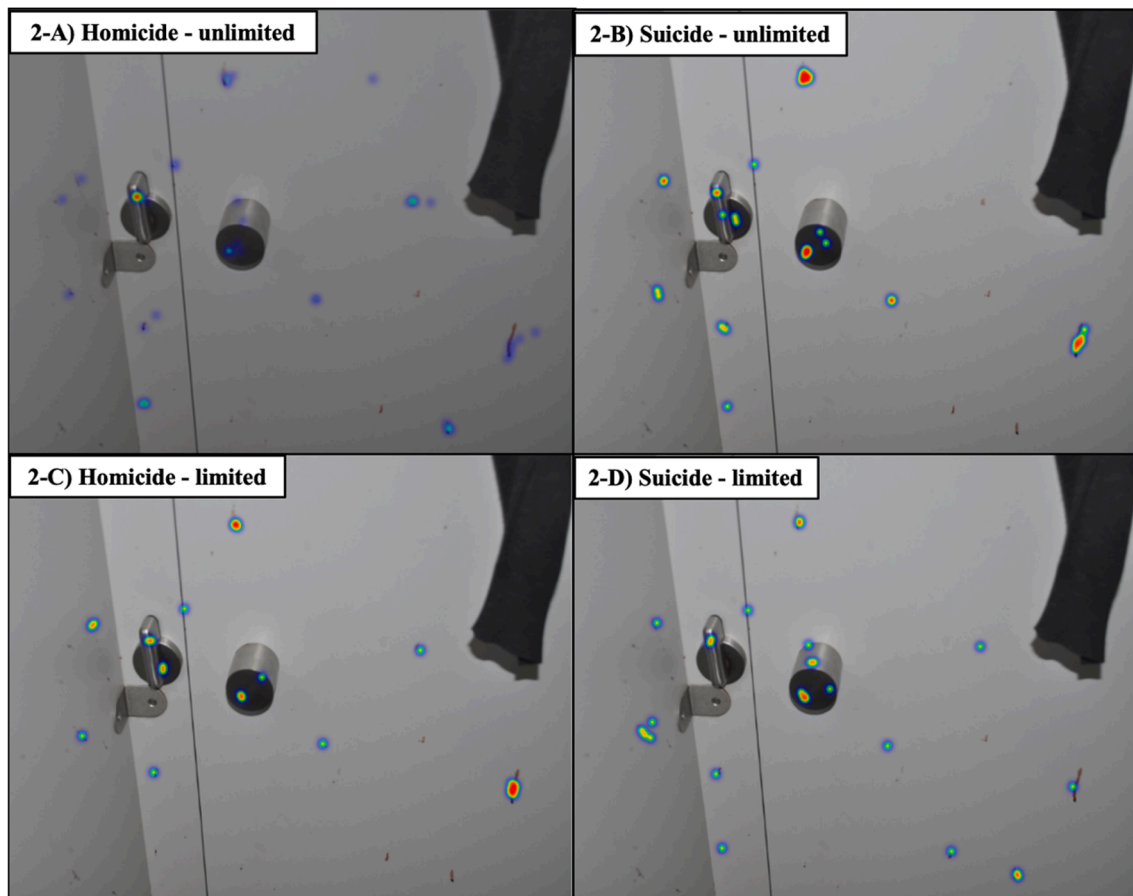
In the suicide condition, participants were told that: 1) On the way to the scene, a police officer provided information that they were going to a site where a man had committed suicide and the site had high volumes of blood, and 2) the detective on site said that a witness had heard a man inside the bathroom who appeared to be on the phone while he was saying “I love you, good bye” several times. In this condition, the photo of the deceased’s left wrist showing non-deep marks consistent with hesitation marks was shown.

All participants were informed that they could take an unlimited number of swabs from the scene (Unlimited Resources).

In Step 2; the trace collection under unlimited resources step, CSIs were exposed to the 11 crime scene photos. They could watch the photos for as long as they wanted but the time taken to complete the task was



Heat map 1. Right Wall of Entrance and Body on the Floor. Location of Markups in the Conditions: Homicide and Unlimited Resources (1-A), Suicide and Unlimited Resources (1-B), Homicide and Limited Resources (1-C) and Suicide and Limited Resources (1-D).



**Heat map 2.** indicates that participants in the Suicide conditions (2-B and 2-D) collected somewhat more blood from the handle on the inside of the toilet cubicle door, compared to participants in the Homicide conditions (2-A and 2-C). This difference appears to have been largest between the Suicide Unlimited (2-B) and the Homicide Unlimited (2-A) conditions.

recorded and correlated with the dependent measures. They were then asked to mark up, in four of the photos, the areas from which they wanted to take their swabs. Once the markup was done, they could not change their decisions. The CSIs markups were saved in the program, producing the heat maps which enabled observations of differences in their collection of blood traces, hot spots etc. Hence, in this step, the program registered three measures: 1) where the CSIs did their markups, 2) how many markups they did and 3) how long they took to complete the task.

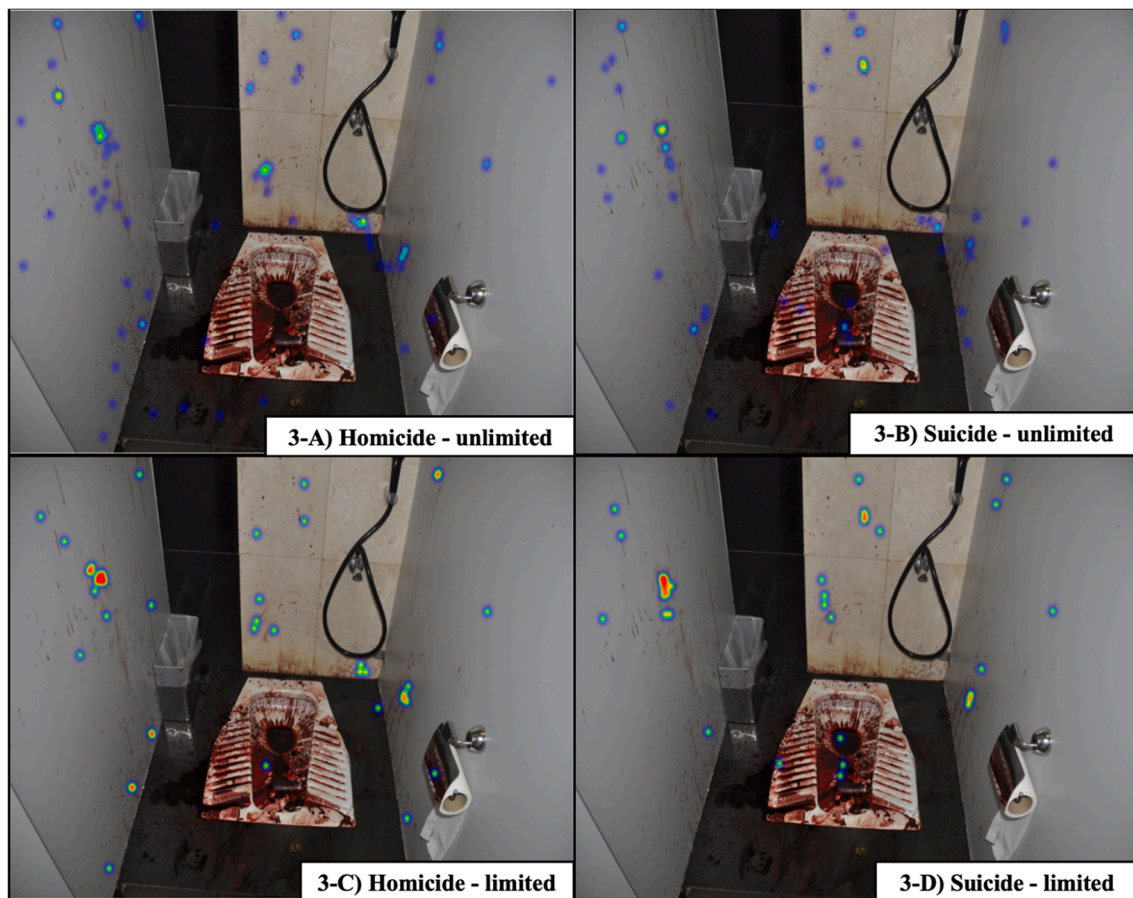
In **Step 3**; the trace collection under limited resources step, all participants were asked to decide again, on the basis of the same four photos, from where they wanted to collect blood. However, this time expert participants were told that they were asked to prioritize and collect given that a maximum of five (5) bloodstains could be analyzed forensically. Hence, Step 3 required CSIs to collect and prioritize simultaneously. This was in order to see whether knowledge of limited resources when it comes to the number of forensic analyses that could be conducted would impact on e.g., from where they collected blood and how many samples they took. Hence, also in this step, the program also registered the same three measures: 1) where the CSIs did their markups, 2) how many markups they did and 3) how long they took to complete the task.

In **Step 4**; the supplementary questions step, all participants were asked five questions in order to gain some insight into their impressions and thought processes while collecting the blood traces. The first two questions pertained to the participants' impressions, more specifically:

- (1) What is your impression: was the deceased killed by another person (i.e., a homicide/manslaughter) or did he inflict injuries to himself (i.e., a suicide)? Participants could choose one of the following options: The deceased was killed by another person (i.e. homicide/manslaughter), he inflicted injuries to himself (i.e. suicide), or Other (with free text),
- (2) On a scale from 0 to 100 %, how confident are you that your impression is correct? (0–100 %). On the scale “0 %” means “totally unsure”, “50 %” means “neither unsure nor sure” and “100 %” means “totally sure”.

The latter three questions were more focused on understanding the participants' thought processes. They were all free response questions and while all participants had to respond, they could provide as little or as much information they wanted. These questions were:

- (3) How did you decide from *where* to secure blood?
- (4) How did you decide *how many* swabs to collect and send for lab analysis?
- (5) Are there rules or recommendations in the organization where you are currently working which tell you from where to secure blood as well as how many swabs to take and send for analysis? If yes, is this related to the type of crime suspected to have taken place? Participants' responses to the five questions were analyzed statistically (questions 1 and 2) and qualitatively using thematic analysis (questions 3–5).



**Heat map 3.** indicates that participants in both conditions collected blood from the walls, although to different extents. Specifically, participants in the Homicide conditions (3-A and 3-C) collected somewhat more blood from what appears to be a smear stain right above the toilet than participants in the Suicide conditions (3-B and 3-D). The difference appears more clearly in the Unlimited (3-A v. 3-B) than the Limited (3-C and 3-D) conditions, although it is likely to still be a relatively small difference.

## 2.5. Data analysis

The data was analysed using IBM SPSS Statistics version 28.0.1.0. For significant results obtained using ANOVA, partial eta squared ( $\eta_p^2$ ) was used as an effect size measure. Because partial eta squared is a measure of the variance that a variable explains which is not explained by other variables in the analysis it, unlike eta squared, partials out effects of other variables and is therefore preferred [72,73]. The associated cut-off levels used for interpretation are  $\eta_p^2$ .

$\geq 0.01$  (small), 0.06 (medium) and 0.14 (large). Furthermore, for chi-square analysis, odds ratio was used as an effect size measure. The higher the odds ratio, the stronger the association. An odds ratio of 1.0 indicates that the odds are the same as those in another condition relevant for comparison and an odds ratio greater than 1.0 indicates that the odds are greater [72]. Because it is usually easier to understand than e.g. Cramér's  $V$ , the odds ratio is usually preferred as an effect size measure. Also, for correlations, the cut-off levels suggested by Cohen [74] which are widely accepted in the research community [72] were used. This entails the following values:  $r \geq 0.10$  (small), 0.30 (moderate) and 0.50 (large). In the Results section, the applicable effect size measures and cut-off values are reported for significant outcomes.

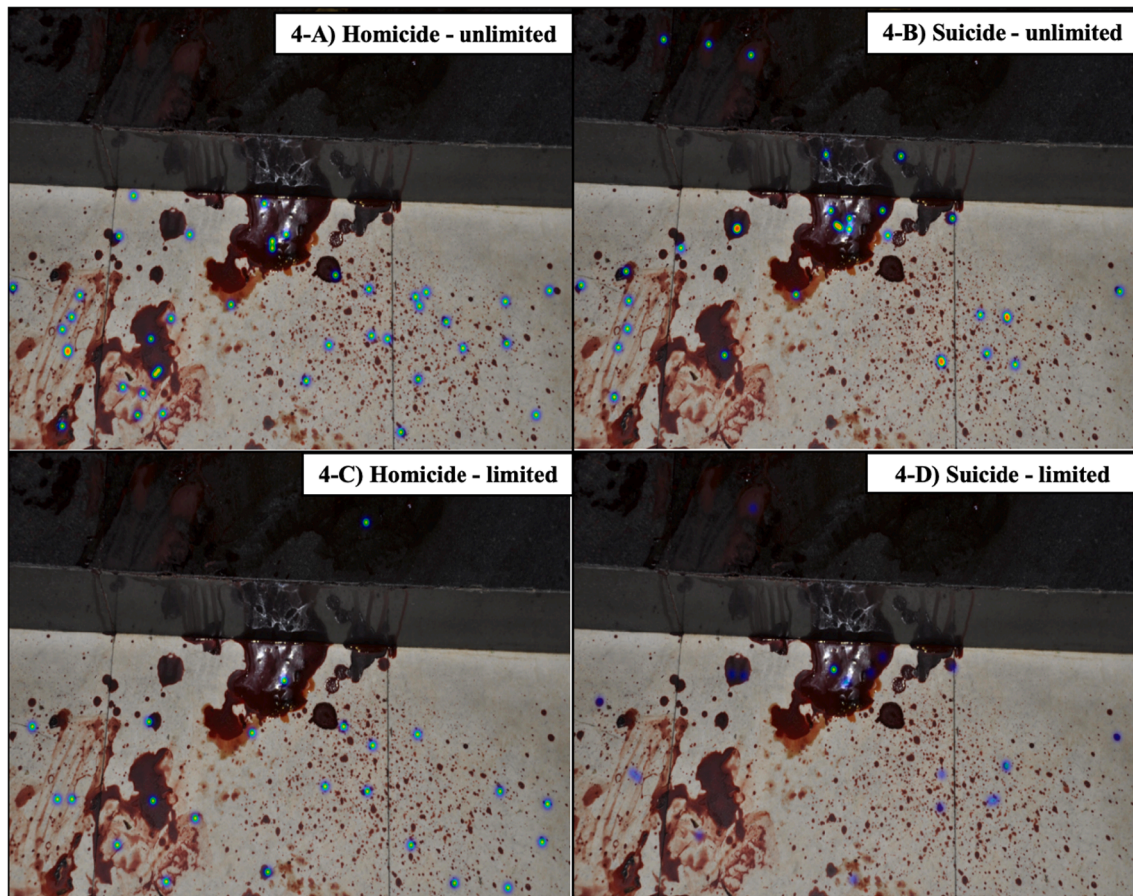
## 3. Experiment 1: Results

### 3.1. Background variables

Exploratory analyses were conducted in order to examine whether any of the background variables, that is, biological sex, age, years of experience and level of training were significantly related to participants' impressions of the scene or their confidence levels.

For impressions of the scene, the following tests were conducted: chi-square for biological sex and level of training, correlational analysis for age as well as years of experience. For confidence levels the following tests were conducted:  $t$ -test for biological sex, correlational analysis for age as well as years of experience and Between Subjects ANOVA for level of training.

As illustrated by the analysis in the Online [Supplementary Material](#), none of the background variables were significantly related to participants' impressions of the scene or their confidence judgments.



**Heat map 4.** Bloodstains below the Toilet. Location of Markups in the Conditions: Homicide and Unlimited Resources (4-A), Suicide and Unlimited Resources (4-B), Homicide and Limited Resources (4-C) and Suicide and Limited Resources (4-D).

### 3.2. Location of markups

#### 3.2.1. Heat maps for homicide and suicide conditions with Unlimited and Resources

In the following, we present and comment on [Heat maps 1-5](#). Note that a higher intensity of the color in the heatmaps indicates a higher number of markups selected by the participants (i.e. intensity goes from lowest to highest as follows: blue, green, yellow and red).

The locations from which CSIs decided to secure blood was different across the conditions, see [Heat maps 1 and 2](#). To facilitate comparisons the maps for the respective conditions are presented together.

As illustrated in [Heat map 1](#), in the Homicide conditions, there was no apparent collection of blood from the deceased's left arm (1-A and 1-C), to be compared to the concentrated collection of blood traces from the left arm in the Suicide conditions (1-B and 1-D). Furthermore, in the Homicide conditions, there is some collection from the external side of the toilet door (1-A and 1-C), while no collection from this location is apparent in the Suicide conditions (1-B and 1-D). However some similarities between the conditions are also apparent, as participants in both conditions appeared to have collected blood from the right wall.

As illustrated in [Heat map 4](#), there appears to have been more variation and spread in the collection made by participants in the homicide condition compared to in the suicide condition, particularly with unlimited resources. Possibly, this is an indication that participants in the homicide condition, to a larger extent than those in the suicide condition, wanted to ensure they did not leave behind any blood from a perpetrator at the scene.

It is also notable that when comparing the limited and unlimited conditions, it appears that the limited condition is not always merely a selection of the samples selected in the unlimited condition, but also

other blood samples are taken. See, for example, how CSIs in the limited condition appear to gather more blood from inside the toilet seat in [Heat map 3](#) (3C&D v. 3A&B). This indicates that one and the same CSI, when faced with limited resources, will not simply do less of the same he/she would do with unlimited resources but it seems like the strategy itself may also change.

Furthermore, to examine variation in how CSI's collected blood in identical conditions, individual responses were compared. Among these individual responses, two CSIs with similar background characteristics (age, years of experience, etc.) were identified. The collection of these two CSIs respectively are presented in [Heat map 5](#). They were both in the homicide condition. The comparison refers to the collection they did with unlimited resources. Participant A collected in total 15 samples while Participant B collected in total 38 samples.

#### 3.2.2. Location of markups – Explanations

In the homicide condition, participants' explanations as to why they decided to secure blood from specific locations were divided into 4 themes. Three of the themes were reoccurring. Most participants, 7 out of 18 (38.89 %), decided on the basis of different blood patterns. Also, 5 out of 18 participants (27.78 %) used a maximizing strategy, i.e. to do as many markups as possible from different blood patterns in different parts of the cubicle. Furthermore, 5 out of 18 participants (27.78 %) made assessments of what blood could be informative about the activity in the cubicle and/or a perpetrator. In this assessment, they would take into account, for example, that the act had been carried out with a sharp object. Another less common strategy, used by 1 out 18 (5.56 %) was to collect blood only from the small patterns of blood in order to avoid mixed DNA and to enable more reliable determinations of activity and the presence of a perpetrator.



Heat map 5. Collection of Blood by Participant A and Participant B Blood in Identical Conditions; Homicide and Unlimited Resources.

In the suicide condition, 8 out of 19 (42.11 %) participants decided to secure blood traces they thought were informative about what happened (activity) and who was there, e.g. from the handle or broken glass or areas where the blood was swiped or smeared. Also, it was fairly common to do as many markups as possible from different types of patterns in different parts of the room (maximizing strategy), which 4 out of 19 (21.05 %) participants mentioned. Some participants, 2 out of 19 (10.53 %) focused on the blood furthest away from the victim to make sure there was not a perpetrator and others – also 2 out 19 (10.53 %) – based their decisions on both blood patterns and surface area of deposition.

Individual participants (1 out 19, 5.26 %) focused on small patterns of blood in order to avoid mixed DNA, or on areas where blood would be unexpected, or where contamination would be least likely.

### 3.3. Number of markups

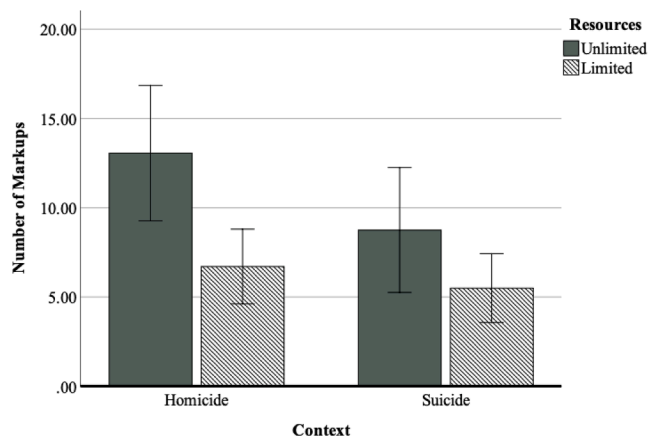
Table 1 outlines the descriptive statistics for the number of markups. As illustrated in Table 1, participants sometimes decided to not collect any blood traces at all. Possibly, this is because they did not consider it relevant to collect blood. Alternative explanations are that



**Table 1**

Descriptive statistics (M,SD, min, max.) for number of markups in the Homicide and Suicide Conditions and Unlimited and Limited Conditions.

| Unlimited | Homicide |      |      |      | Suicide |      |      |      |
|-----------|----------|------|------|------|---------|------|------|------|
|           | M        | SD   | Min. | Max. | M       | SD   | Min. | Max. |
|           | 13.06    | 9.79 | 0    | 38   | 8.75    | 5.35 | 0    | 17   |
| Limited   | M        | SD   | Min. | Max. | M       | SD   | Min. | Max. |
|           | 6.71     | 5.00 | 4    | 20   | 5.50    | 3.49 | 2    | 20   |
| Total     | 9.89     |      | 0    | 38   | 7.13    |      | 0    | 20   |



**Fig. 1.** Number of Markups by Resources and Context in Experiment 1.

they did not fully understand the instructions or that they did not fully understand how to use the heat maps.

A mixed subjects ANOVA was conducted to test whether the Resources (Unlimited v. Limited) and the Context (Homicide v. Suicide) impacted significantly on the number of markups made. There was a significant main effect of Resources,  $F(1) = 11.41, p = .002, \eta_p^2 = 0.246$  (large effect). However, there was no significant main effect of Context,  $F(1) = 3.47, p = .071$ , and no significant interaction effect,  $F(1) = 1.191, p = .283$ . This is illustrated in Fig. 1. As illustrated in Fig. 1, the main effect of Resources appear to have been driven by differences in the homicide condition.

Also, to test whether participants' impressions were significantly related to the number of markups they made, we conducted a between subjects ANOVA. This analysis showed no significant difference,  $F(2) = 0.105, p = .901$ . Notably, since participants were only asked about their impressions after they had done their markups, it would be unclear how the two variables relate to one another, even if there had been a significant relation.

### 3.3.1. Number of markups - explanations

In the homicide condition, 10 out of 18 (55.55 %) of the participants wanted to do as many markups as possible to avoid contamination and improve chances of saying something about activity and identity. Others, 4 out of 18 (22.22 %) decided based on blood patterns and location, or on blood that might originate from the perpetrator (2 out of 18, 11.11 %). Individual participants (1 out of 18, 5.55 %) decided on the basis of the volume of the blood or the injuries on the victim's body.

In the suicide condition, 7 out of 19 (36.84 %) participants wanted to do as many markup as possible and/or based their decisions on blood patterns and locations (5 out 19, 26.32 %). Also, 2 out 19 (10.53 %) participants stated they would submit samples that might originate from someone other than the victim. Individual participants (1 out 19, 5.26 %) based the number of markups on the volume of the blood (5.88 %), or stated that the number of markups did not matter since the blood was expected to come from the deceased regardless of whether it was a homicide or suicide. Furthermore, individual participants suggested that

the number is for the detective not the CSI to decide, or that 5 to 7 swabs should be taken considering the wet/dry method used in the participant's jurisdiction (based on e.g. lab capacity).

## 3.4. Impressions and confidence

### 3.4.1. Other impressions

Among participants who answered "other" for their impressions of the scene, thematic analysis was conducted for each condition. In the homicide condition, participants who said their impression was "other" explained that it was too early and/or that they had too little information available to form any impression regarding how the deceased man had passed away (80 %), or that they thought it was not for them to decide (20 %).

In the suicide condition, participants also stated that the information provided was insufficient for deciding (60 %), some mentioned other explanations such as medical conditions or illnesses (20 %) and some stated that while they would want to know more, for example, regarding the prevalence of blood outside the cubicle, the information spoke more for suicide than homicide (20 %).

As illustrated in Table 2 the percentage of CSIs whose first impressions were that the scene was a homicide scene was higher in the homicide condition (52.90 %) than in the suicide condition (30.00). Also the percentage of CSIs who stated "other" as their first impression was lower in the homicide condition (29.40 %) than in the suicide condition (55.00 %). However, there was no significant association between the condition participants were in and their impressions of the scene,  $\chi^2(2) = 2.624, p = .269$ .

A Between Subjects ANOVA was conducted to examine whether the condition (Homicide v. Suicide) and the Impression (Homicide v. Suicide) impacted significantly on confidence judgments. There were no significant main effects of condition,  $F(1) = 0.054, p = .818$ , or impression,  $F(1) = 1.185, p = .319$  and neither a significant interaction effect,  $F(1) = 0.787, p = .464$  on confidence levels.

## 3.5. Organizational rules and recommendations

When asked to answer whether there are rules or recommendations on securing blood in their respective jurisdictions/workplaces, the participants answered in the following way. Most participants, 23 out of 37 (62.16 %) said there are neither rules nor recommendations but that this is up to the CSI's critical thinking and own judgements on what are relevant traces. A few participants, 7 out of 37 (18.92 %) answered that there are such limitations with less serious but frequent crimes, like burglaries. Four participants in this group mentioned three swabs as a limit in the less serious crimes but that there usually wouldn't be any limits for more serious crimes. One participants mentioned that there are rules when it comes to examine vehicles involved in e.g. robbery, and another participant mentioned that if a deceased body is found there are rules stating from where on the body to secure traces. Yet another participant stated that while there are no limitations as to how many swabs can be taken a CSI is aware of limitations on how many swabs will then be analyzed forensically. Furthermore, a single participants mentioned that while no recommendations are in place, if the medical doctor and CSI believe that they are dealing with a suicide, no blood samples will be collected.

To test whether these differences significantly impacted on participants' decisions in this study we conducted chi-square tests and Between Subjects ANOVAs to test for any associations with the dependent variables. As illustrated by the analysis available in the Online Supplementary Material, the differences in practices were not significantly associated with any of the dependent variables.

### 3.6. Time to complete task

The total time to complete the task was correlated with participants' background variables as well as the dependent variables. This analysis showed two significant correlations. Firstly, there was a significant positive correlation between the time to complete the task and the number of years of experience as a crime scene investigator,  $r = 0.387$  (moderate),  $p = .018$ . Secondly, there was a significant positive correlation between the time to complete the task and the age of the CSIs,  $r = 0.369$  (moderate),  $p = .025$ . There was also a significant positive correlation between the number of years of experience and the age of the CSIs,  $r = 0.748$ ,  $p < .001$  (strong). Hence, it is uncertain whether longer experience, higher age or the two in interaction are associated with longer time to complete the task. However, there were no significant correlations with the dependent measures, that is, the impression of the scene  $r = 0.198$ ,  $p = .240$  or confidence judgments,  $r = 0.318$ ,  $p = .065$ .

### 3.7. Country where CSIs operate

A chi-square analysis was run to examine whether CSI operating in Sweden, Norway and the UK differed significantly in their impressions of the scene, and this analysis showed no such difference,  $\chi^2(4) = 2.33$ ,  $p = .676$ . Furthermore, there were no significant differences in participants' confidence levels when analyzed by the country in which they operate,  $F(2) = 0.379$ ,  $p = .687$ . Neither was there any significant difference between the number of markups CSIs from these respective countries made,  $F(2) = 28.32$ ,  $p = .457$ . Also, the participants did not vary significantly in their responses as to whether their respective organizations had rules and recommendations for trace collection  $\chi^2(2) = .83$ ,  $p = .661$ , or in the time taken to complete the task,  $F(2) = 0.44$ ,  $p = .645$ .

## 4. Experiment 2: Novices

### 4.1. Method

#### 4.1.1. Participants

Participants were 80 individuals from the UK recruited via Prolific, 42 (52.50 %) females and 37 (46.30 %) males. Their ages varied between 18 and 57 ( $M = 25.400$ ,  $SD = 7.477$ ). They were offered 3 Euros each for their participation in line with Prolific guidelines. The inclusion of novices generated baseline data to be compared with the expert data in Experiment 1.

#### 4.1.2. Design, material and procedure

Just like Experiment 1, Experiment 2 had a 2 (Resource limitations: yes vs. no)  $\times$  2 (Context: homicide vs. suicide) Mixed Subjects Design. The Resource limitations variable was varied within participants whereas the Context variable was varied between participants. Also the material and the procedure was identical to in Experiment 1, with the exceptions of what questions were posed to participants (Step 4). Given that participants in Experiment 2 were novices they were not asked how many years of experience, what type of BPA training they have and also not whether there are rules or recommendations in their respective organizations about collecting blood from high volume blood scenes.

**Table 2**

Descriptive statistics for overall impressions (%) and confidence judgments (M,SD) in Experiment 1.

|            | Condition          |                    |                    |                   |                   |                    |
|------------|--------------------|--------------------|--------------------|-------------------|-------------------|--------------------|
|            | Homicide           |                    |                    | Suicide           |                   |                    |
|            | Homicide           | Suicide            | Other              | Homicide          | Suicide           | Other              |
| Impression | 52.90              | 17.60              | 29.40              | 30.00             | 15.00             | 55.00              |
| Confidence | 69.889<br>(15.909) | 66.667<br>(11.547) | 49.800<br>(35.358) | 55.00<br>(13.784) | 70.00<br>(10.000) | 55.455<br>(27.955) |

## 5. Results: Experiment 2

### 5.1. Background variables

Exploratory analyses were conducted to just like in Experiment 1. As illustrated by the analysis available in the Online [Supplementary Material](#) none of the background variables were related to the dependent variables.

### 5.2. Location of markups

#### 5.2.1. Heat maps for homicide and suicide conditions with Unlimited and Resources

Unlike with the experts, the locations from which novices decided to secure blood did not seem to vary across the conditions. This is illustrated in [Heat maps 6-9](#). To facilitate comparisons the maps for the respective conditions are presented together.

#### 5.2.2. Location of markups – Explanations

In the homicide condition, participants' explanations as to why they decided to secure blood from specific locations were divided into 7 themes. Most participants, 17 out of 40 (42.50 %) decided to secure blood from places where there was a lot of blood, and some, 15 out of 40 (37.50 %) from where they expected to find blood from another person than the victim. A few participants, 6 out of 40 (15.00 %), were attempting to secure blood from all over the cubicle. Yet others, 3 out of 40 (7.50 %), in places that were not contaminated or were closest to the victim (also 3 out of 40, 7.50 %). Individual participants (1 or 2) secured blood on the basis of different patterns, e.g. blood that seemed to come from different angles or they just guessed (2.439 %). Note that some participants mentioned more than one theme.

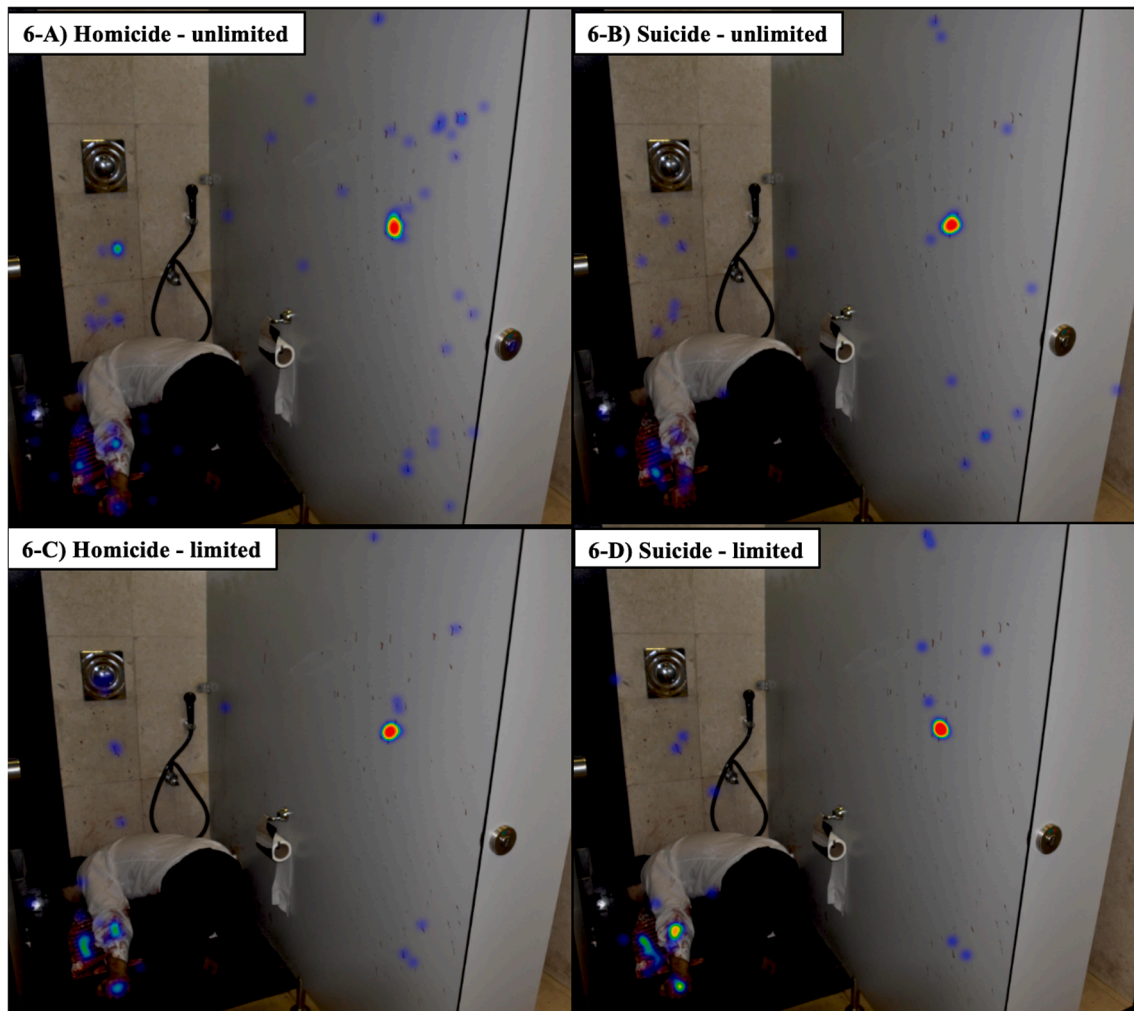
In the suicide condition, most participants, 17 out of 40 (42.50 %) secured blood from places where there was a lot of blood or, 8 out of 40 (20.00 %) tried to spread the collection all over. A few novices, 5 out of 40 (12.50 %) decided to secure blood from close to the victim and/or where they expected to find blood from someone else other than the victim. Individuals participants (1 or 2) focused on places where the blood wouldn't be contaminated, where there was just a little blood, based on patterns or just followed their intuition/guessed.

Hence, roughly the same themes were mentioned in both conditions, although with different frequencies, and there was one additional theme in the suicide condition (where there was just a little blood) which was not mentioned in the homicide condition.

### 5.3. Number of markups

[Table 3](#) outlines the descriptive statistics for the number of markups.

A mixed subjects ANOVA was conducted to test whether the Resources (Unlimited v. Limited) and the Context (Homicide v. Suicide) impacted significantly on the number of markups made. There was a significant main effect of Resources,  $F(1) = 21.489$ ,  $p < .001$ ,  $\eta_p^2 = 0.218$  (large effect). However, there was no significant effect of Context,  $F(1) = 2.638$ ,  $p = .108$  or significant interaction effect,  $F(1) = 3.589$ ,  $p = .062$ . As illustrated in [Fig. 2](#), the main effect appears to have been driven by a difference in the homicide condition, in which participants



**Heat map 6.** Right wall of Entrance and Body on the Floor. Location of Markups in the Conditions: Homicide and Unlimited Resources (6-A), Suicide and Unlimited Resources (6-B), Homicide and Limited Resources (6-C) and Suicide and Limited Resources (6-D).

collected significantly more traces in the unlimited condition than the limited. Also, the number of traces collected in the homicide condition with unlimited resources was significantly higher than in the suicide condition with limited resources.

#### 5.3.1. Number of markups – Explanations

In the homicide condition, participants' explanations as to how they decided on how many markups to make entailed the following 5 themes. Most participants, 27 out of 40 (67.50 %) had a maximizing strategy, i.e. they wanted to get as many markups as possible to make sure they did not miss anything important. Some, on the contrary, wanted to make as few markups as possible with the best results (7 out of 40, 17.50 %). Others, 3 out of 40 (7.50 %) decided on the basis of where it was likely that the involved would have touched. Individual participants (1 or 2) did not think much about the number or only marked up closest to the broken glass, referred to as "the weapon".

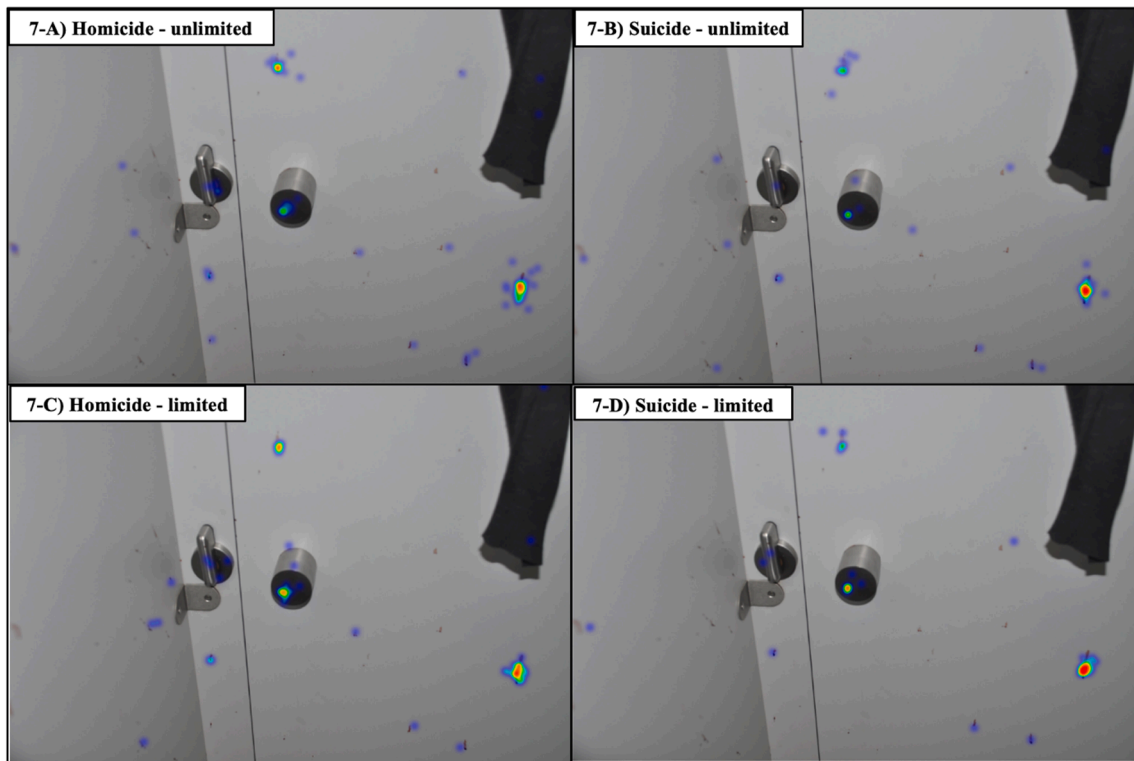
In the suicide condition, 14 out of 40 (35.00 %) participants decided to markup based on the number of blood stains they could see and/or (11 out of 40, 27.50 %) the lowest number necessary to obtain a good result. A few, 8 out of 40 (20.00 %) did not think much about the number. Others tried to do as many markups as possible (7 out of 40, 17.50 %). Hence, none of these participants decided on the basis of where individuals were likely to have touched or closest to the "weapon".

#### 5.4. Impressions and confidence

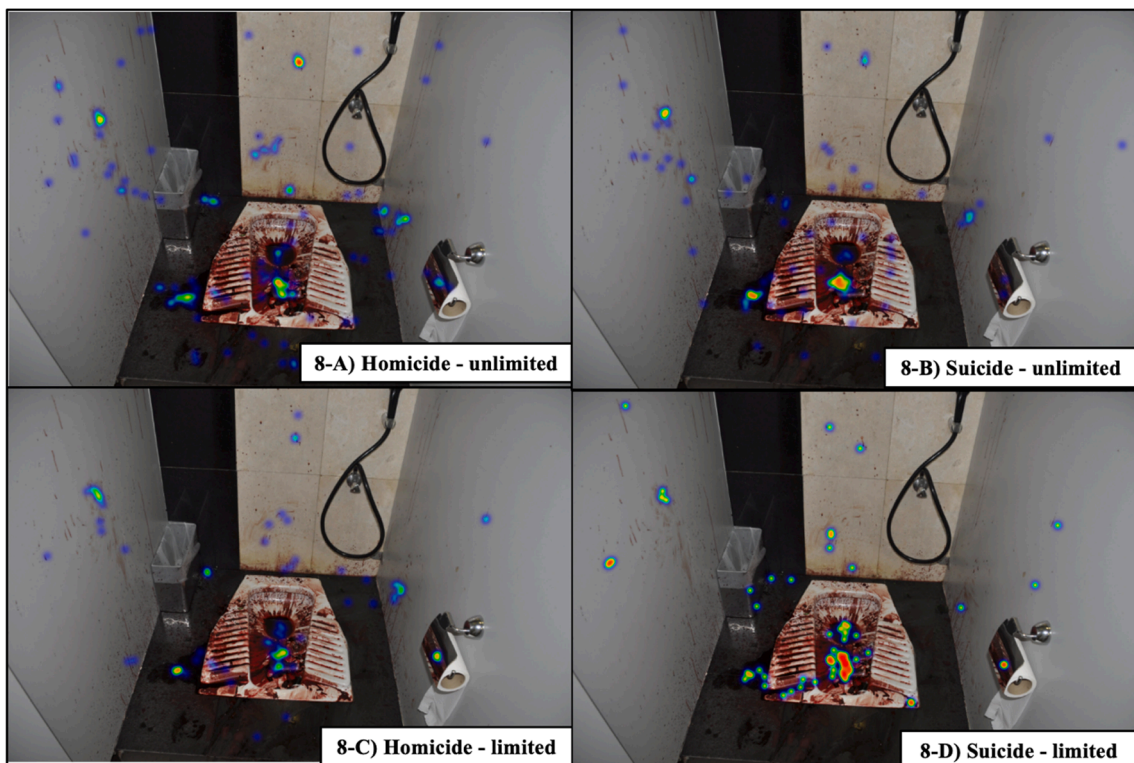
Descriptive statistics for participants overall impressions of the scene and confidence judgments are displayed in [Table 4](#).

There was a significant association between the condition the participants were in and their impressions of the scene,  $\chi^2(1) = 14.687, p < .001$ . The odds that participants perceived of the scene as a homicide scene were 1.522 times higher if they were in the homicide condition than the suicide condition. Vice versa, if participants were in the suicide condition, the odds of them perceiving of the scene as a suicide scene were 14.718 times higher than if they were in the homicide condition. Because the odds ratio was higher in the suicide condition (14.718) than in the homicide condition (1.522), it seems like the suicide manipulation had more of an impact on novice participants. However, this is probably at least in part explained by that many novices, regardless of manipulation, appears to have thought that a homicide had taken place, see [Table 4](#) (97.60 % in homicide condition and 64.10 % in the suicide condition). In other words, had it not been for the suicide manipulation, most participants probably would not have thought of this possibility themselves.

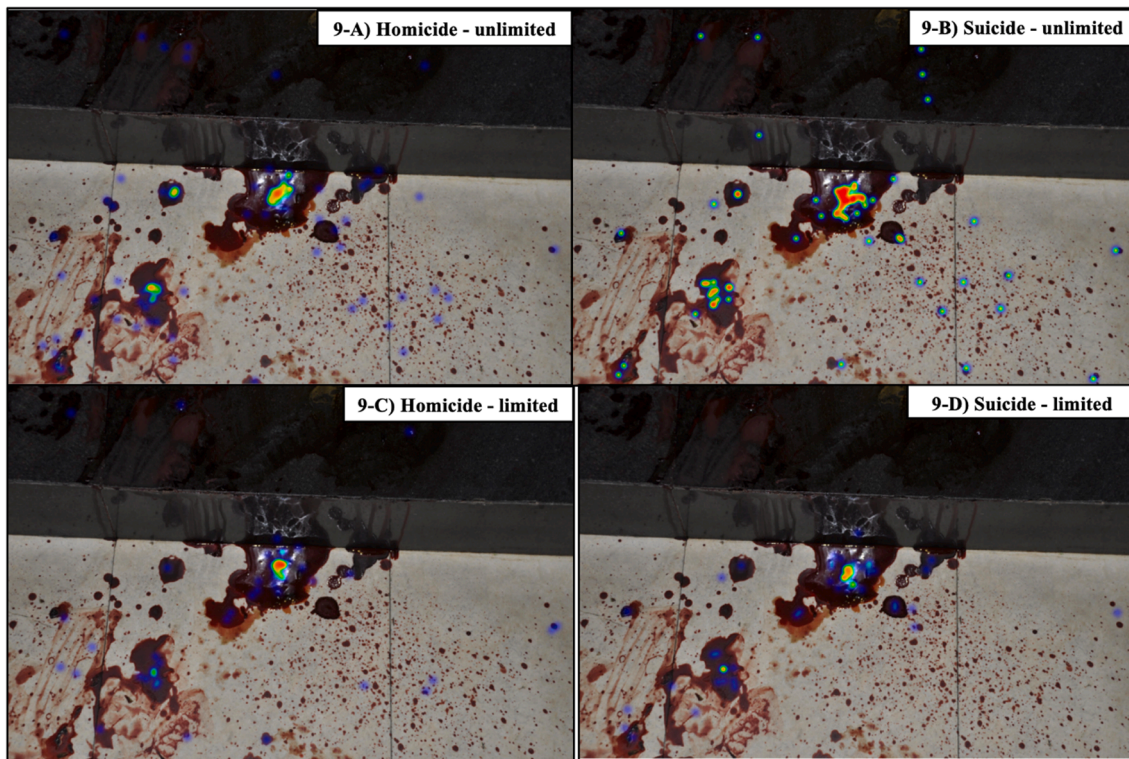
A Between Subjects ANOVA was conducted to examine whether the condition (Homicide v. Suicide) and the Impression (Homicide v. Suicide) impacted significantly on confidence judgments. There were no significant main effects of conditions,  $F(1) = 0.247, p = .620$  or impression,  $F(1) = 1.411, p = .239$ . There was a significant interaction



**Heat map 7.** Close-up to Bloodstains around the Door Handle of the Toilet Cubicle Seen from Inside. Location of Markups in the Conditions: Homicide and Unlimited Resources (7-A), Suicide and Unlimited Resources (7-B), Homicide and Limited Resources (7-C) and Suicide and Limited Resources (7-D).



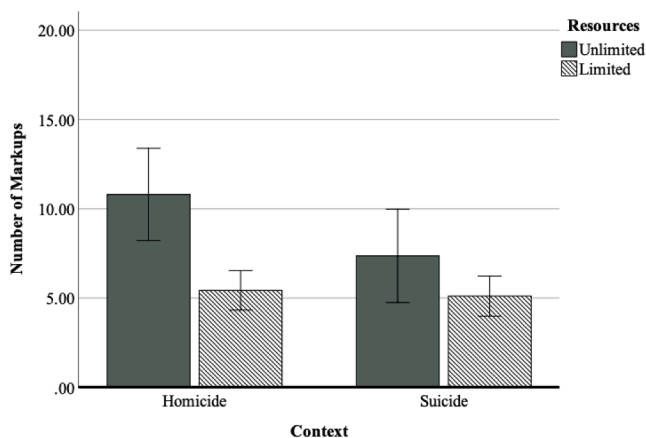
**Heat map 8.** Overview of Bloodstains on the Walls and the Floor. Location of Markups in the Conditions: Homicide and Unlimited Resources (8-A), Suicide and Unlimited Resources (8-B), Homicide and Limited Resources (8-C) and Suicide and Limited Resources (8-D).



**Heat map 9.** Bloodstains below the Toilet. Location of Markups in the Conditions: Homicide and Unlimited Resources (9-A), Suicide and Unlimited Resources (9-B), Homicide and Limited Resources (9-C) and Suicide and Limited Resources (9-D).

**Table 3**  
Descriptive statistics (M,SD, min, max.) for number of markups in the Homicide and Suicide Conditions and Unlimited and Limited Conditions.

|              | Homicide |       |      |      | Suicide |       |      |      |
|--------------|----------|-------|------|------|---------|-------|------|------|
|              | M        | SD    | Min. | Max. | M       | SD    | Min. | Max. |
| Unlimited    | 10.800   | 9.455 | 0    | 40   | 7.359   | 6.698 | 0    | 40   |
| Limited      | 5.425    | 3.594 | 0    | 20   | 5.103   | 3.432 | 0    | 19   |
| <b>Total</b> |          |       |      |      |         |       |      |      |



**Fig. 2.** Number of Markups by Resources and Context in Experiment 2.

**Table 4**  
Descriptive statistics for overall impressions (%) and confidence judgments (M, SD) in Experiment 2.

|            | Condition       |         |                 |                |
|------------|-----------------|---------|-----------------|----------------|
|            | Homicide        |         | Suicide         |                |
|            | Homicide        | Suicide | Homicide        | Suicide        |
| Impression | 97.60           | 2.40    | 64.10           | 35.90          |
| Confidence | 80.425 (18.230) | 50.00   | 65.680 (17.382) | 74.00 (15.059) |

effect between condition and impression,  $F(1) = 4.335, p = .041$ . However, the more conservative post hoc tests came out non-significant.

**5.5. Time to complete task**

Correlational analysis was conducted to test whether the total time to complete the task was correlated with participants' background

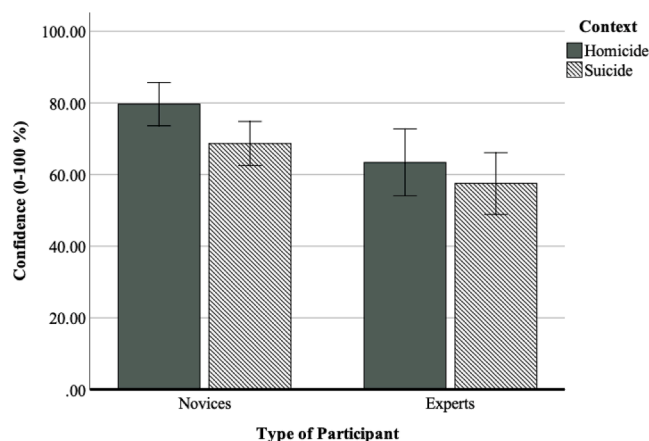


Fig. 3. Confidence Judgments by Type of Participant and Context.

variables or the dependent variables. This analysis showed no significant correlations,  $r = 0.117$ ,  $p = .301$  (biological sex),  $r = 0.011$ ,  $p = .921$  (age),  $r = -0.076$ ,  $p = .501$  (impressions of scene),  $r = 0.085$ ,  $p = .454$  (confidence judgments).

## 6. Comparison of novices and experts

### 6.1. Impressions and confidence

There was a significant association between whether the participant was a novice or an expert and what their impression of the scene was,  $\chi^2(2) = 40.817$ ,  $p < .001$ . Overall, across both the homicide and the suicide condition, novices were 6.35 times more likely than experts to state that their impression of the scene was homicide. Within the homicide condition specifically,  $\chi^2(2) = 26.380$ ,  $p < .001$ , novices were 35.56 times more likely than experts to state homicide as their impression. Within the suicide condition specifically,  $\chi^2(2) = 26.380$ ,  $p < .001$ , novices were 4.17 times more likely than experts to state homicide as their impression. Notably, none of the novices stated “other” as their impression, while a total of 16 experts – 5 in the homicide condition and 11 in the suicide condition – did so. Overall, these results suggest that novices were more likely to believe that a homicide had taken place, regardless of what condition they were in. As indicated by the high odds ratios, particularly 35.56, this was a considerable difference.

A Between Subjects ANOVA was conducted to examine whether the Context (Homicide v. Suicide) and the Type of Participant (Novice v. Expert) impacted significantly on confidence judgments. There were significant main effects both of the Context,  $F(1) = 4.759$ ,  $p = .031$ ,  $\eta_p^2 = 0.040$  and the Type of Participant,  $F(1) = 12.503$ ,  $p = .001$ ,  $\eta_p^2 = 0.100$ . However, there was no significant interaction effect,  $F(1) = 0.433$ ,  $p = .512$ . As illustrated in Fig. 3 participants in the homicide condition appeared more confident than participants in the suicide condition and the novices were more confident than experts.

## 7. Number of markups

A visual comparison of Figs. 1 and 2 suggest that novices and experts showed similar behaviors in the number of markups they did across the conditions, although experts appeared to collect somewhat more traces in the homicide unlimited condition. This difference did not come out significant,  $t(55) = -0.817$ ,  $p = .418$ .

## 8. Discussion

### 8.1. Between expert Biasability

The heat maps contain a few relevant findings for the discussion

pertaining to Between Expert Biasability. As illustrated by Heat map 1, expert participants in the suicide condition seemed to have focused more on the deceased’s arm when collecting blood, compared to participants in the homicide condition who do not appear to have collected virtually any blood from the man’s arm. The difference remains even when participants in the respective conditions had limited resources. The difference is likely a result of the context manipulation, including the picture of the hesitation marks and the witness testimony. The context was manipulated using many pieces of information in order to simulate real life criminal cases as far as possible. In addition, the combined context, including the choice of the photo for each condition, has the potential to strengthen the experimental manipulation, thus addressing methodological concerns on ineffective manipulation designs in forensic decision-making research [75]. This is, however, associated with the limitation that it cannot be concluded which specific piece of information had the observed effect, but simply that the context in a wider sense had such an impact. Furthermore, we decided to only show the photo of the wrist in the suicide condition and only the photo of the neck in the homicide condition to simulate how investigators’ choices to communicate some but not other information to CSIs can be a biasing factors, especially when combined with other information consistent with the preferred hypothesis.

Furthermore, as illustrated in Heat map 2, participants in the suicide condition with unlimited resources appear to have collected more blood from the handle on the inside of the toilet cubicle door, compared to participants in the homicide condition with unlimited resources. The door handle ought to be a relevant place for collection, even if it is likely, in a public toilet, that DNA from multiple individuals will be present. It is unclear why participants in the homicide condition would not to the same extent perceive of this as a relevant collection spot. One possible interpretation is that they presumed they would find blood from the perpetrator elsewhere, closer to the victim, and that DNA from the handle would not be sufficient for proving guilt.

Interestingly, however, when participants had limited resources, the differences between the homicide and suicide conditions were smaller (as also more participants in the homicide condition collected blood more often from the handle, compared with the unlimited condition). Possibly, the limitation made them motivated to secure DNA from the handle, should it not be possible to identify closer to the victim. Furthermore, as illustrated in Heat map 3, participants in the homicide condition appear to have been somewhat more likely to collect blood from what appears to be a smear stain right above the toilet. The difference appears more clearly in the unlimited than the limited condition, although in all cases it appears to be a relatively small difference. If there has been struggle between two individuals within the toilet cubicle, it is possible that one or both individuals have deposited blood in a smear stain. If blood from someone other than the victim is found within this stain it can therefore both be informative of identity and activity. This can be compared to any potential DNA left on the handle, which is more indicative of identity, and not necessarily that the individual has been involved in the death of the man. Seen together, Heat map 2 and 3 therefore seems to indicate that participants in the homicide condition may have been more focused on tying a perpetrator to not only having been inside the cubicle but also having, potentially, committed a crime against the victim.

There are also some relevant differences appearing in Heat map 4. There appears to have been more variation and spread in the collection made by participants in the homicide condition compared to in the suicide condition, particularly with unlimited resources. Possibly, this is an indication that participants in the homicide condition, to a larger extent than those in the suicide condition, wanted to ensure they did not leave behind any blood from a perpetrator at the scene.

Based on visual inspection, the heat maps together seem to suggest that participants in the respective conditions were more likely to collect blood from where it was expected to be found, given that the hypothesis they had been provided with was true. This tendency appeared more

clearly with unlimited resources. However, two things should be noted in this regard. Firstly, in a few places (see e.g. [Heat map 1 and 3](#) in Experiment 1), participants in both the homicide and the suicide conditions did collect blood from the same location although to different extents. Apparently, some traces were important for participants in both conditions, while some were chosen more often in one of the conditions. Secondly, it should be noted that since the location of the markups have been recorded using heatmaps we have not been able to test for significant differences in meaningful ways and future research still needs to evaluate the nature of this difference. For example, also participants in the suicide condition appear to have collected blood traces to see if another individual could be linked to the scene. One possible interpretation is that participants in the suicide condition wanted to examine the possibility primarily to exclude the possibility of a homicide while participants in the homicide condition wanted to identify a perpetrator.

It is also interesting to note that there are some indications in the Heat Maps (see e.g. [Heat map 3](#)), that CSIs, with limited resources, do not simply collect fewer traces from the same spots as with unlimited resources but they in fact change the spots to. This can be interpreted as an issue with Within Expert Biasability, a term also stemming from the Hierarchy Expert Performance (HEP) [12]. This means that one and the same expert, who conducts the same task at two different points in time will vary in their decisions given that the context changes.

Apart from the location of the markups, there are also some relevant differences in the number of markups made by the participants in the different conditions. For example, in the homicide condition with unlimited resources, participants collected an average of 13.06 samples. This is almost twice as many as when participants were aware of limited resources in the homicide condition where the average was 6.71 samples. The resource limitation did have a significant impact ( $p = .002$ ,  $\eta_p^2 = 0.246$ ). The context (Homicide v. Suicide) also seem to matter in this regard as, for example, participants in suicide condition with unlimited resources collected an average of 8.75 samples, to be compared to the corresponding average in the homicide condition (13.06), although the context did not have a significant impact in this regard ( $p = .071$ ,  $\eta_p^2 = 0.090$ ). The largest difference in averages was between the homicide condition with unlimited resources (13.06) and the suicide condition with awareness of limited resources (5.50), although there was no significant interaction effect ( $p = .283$ ,  $\eta_p^2 = 0.033$ ). However, as illustrated by [Fig. 1](#), the significant main effect of Resources appear to have been driven differences in the homicide condition.

Similarly, participants in the homicide condition were more likely to state that their first impression was that a homicide had taken place. Specifically, 52.90 % of these participants stated their impression was that a homicide had taken place, and they were, on average 69.89 % confident. This is somewhat higher than in the suicide condition in which 30.00 % of participants stated “homicide” as their first impression and they were on average 55.00 % confident. Interestingly, among the participants in the suicide conditions, participants also appeared to have been more likely to state “other” (55.00 %) than in the homicide condition (29.40 %). Potentially, this is an indication that participants in the suicide condition were more open to alternative courses of events and/or more uncertain about what had occurred. For example, in the suicide condition participants who answered “other” explained that the death may have happened due to medical conditions or illnesses, which was not mentioned in the homicide condition.

### 8.2. Between expert reliability

Even among expert participants who made their decisions in identical conditions, there was some relevant individual variation. This is clearly illustrated in [Heat map 5](#). Both Participant A and B were in the homicide conditions but when they collected blood with unlimited resources, Participant A decided to secure a total of 15 traces while Participant B secured 38 traces. The location from which they secured blood also seem to be different, for example because Participant A

collected blood only from the main blood stain below the toilet (2nd picture from the left) while Participant B also collected from the blood pattern next to it. While it is unknown whether the collection of either participant would have been better in identifying a potential perpetrator, it seems that such differences may have important implications. Also, when participants explained why they decided to secure blood from a certain location there was relevant variation. For example, among participants in the homicide condition, 7 out of 18 (38.89 %) decided on the basis of different blood patterns. A few others, 5 out of 18 (27.78 %) wanted to do as many markups as possible from different parts of the room, while yet others (the same number) decided to collect based on what blood could be informative of activity. One participant (5.56 %) decided to collect blood only from small patterns in order to avoid mixed DNA. Similarly, different CSIs seemed differently prepared to state a first impression. For example, while some stated “Homicide” and were around 70 % confident, others would say that it was not for them to decide, or that it was too early to decide. It is possible that this reflects differences in metacognitive judgments or subjective decision thresholds, or tolerance to ambiguity [76,77]. In other words, different CSIs vary in when, and whether, they are comfortable in stating their opinion in the case. It is unknown whether and what effect such differences would have in real life criminal cases. However, possibly, individuals who are more likely to say “Homicide” already from the outset, would be more susceptible to biases such as confirmation bias. This bias, however, also operates on a subconscious level and is not necessarily connected to what a CSI decides to declare or not declare about their beliefs to someone else. On a similar note, there appears to have some differences in the total time taken to complete the task. Specifically, CSIs of higher age and with more years of experience seem to have taken longer to complete the task. However, age and years of experience were not correlated with, for example, the impression of the scene. What exact consequences such individual differences have in real life criminal cases is unknown, but it does underline that individual differences as a source of lacking Between Expert Reliability [78] deserves more attention in future research.

### 8.3. Experts vs. Novices

When comparing the expert and novice data some relevant differences and similarities appear.

While experts’ collection of blood traces differed in some ways depending on what condition they were in (Resources, Context) – see [Heat maps 1-4](#), novices collection show no apparent differences in this regard – see [Heat maps 6-9](#). Yet, for both experts and novices, it seems the Context manipulation was relevant for their first impressions of the scene, although to different degrees. For example, among novices in the homicide condition, 97.60 % of participants stated that their impressions of the scene was homicide and they were, on average, 80.43 % confident. Among the experts in the homicide condition, 52.90 % answered that their first impression was homicide and they were, on average, 69.89 % confident. Specifically, novices in the homicide condition were 35.56 times more likely than experts in the same condition to state “homicide” as their first impression. In addition, none of the novice participants thought of “other” answers while this was fairly common among the experts, especially those in the suicide condition. This indicates that novices were more likely than experts to jump to conclusions about what might have happened.

In the explanations provided for the location of their markups, most novices decided to secure blood from places where there was a lot of blood (41.46 %), where they expected to find blood from another person than the victim (36.59 %) or they tried to spread the collection all over the scene (20.51 %). Among experts, some of the most commonly mentioned explanations was that they decided to secure blood traces that could say something about what happened (activity) and who was there, e.g. from the handle or broken glass or areas where the blood was swiped or smeared (42.11 % in the Suicide condition). Furthermore,

many experts decided on the basis of different blood patterns (38.89 % in the Homicide condition) or wanted to do as many markups as possible from different parts of the room (27.78 % in the Homicide condition).

As such, there were some similarities in the reasons provided among novices and experts. It is possible that the reasons for securing blood from one or the other location could be further examined to provide CSIs with a more scientifically sound basis for their decision making. These findings in fact suggest that the motivation for collecting blood from one or the other location is not necessarily that different from that provided by novices. Note however, that while the explanations provided for collections were similar, experts and novices seem to have varied in how they actually collected traces, based on the heat maps.

Overall, the findings in this study pertaining to Between Expert Biasability and Reliability are consistent with but also expand on previous research. For example, it is in line with previous findings on how collection of other crime-relevant traces than blood can be biased by contextual information [52]. Similarly, it reinforces the importance of the human factor when it comes to blood traces but compared to previous research it focuses less on blood stain pattern analysis [54] and more on the actual collection of blood, which can be informative of both activity and identity.

## 9. Conclusion

The purpose of this research was to examine Between CSI Biasability and Reliability in the selection of blood traces from a crime scene. **Experiment 1** with CSIs operating in Sweden, Norway and the UK provided data on how experienced CSIs collect blood from a scene following different resources and different contexts. Then, as a comparison, **Experiment 2** used the same experimental design with novices to obtain baseline data for how individuals without the relevant expertise would behave.

The results points to potential issues with both biasability and reliability. As such, it appears that the decision making of CSIs could benefit from a more sound scientific basis. It is unknown whether one blood collection strategy generally would increase the chances of, for example, identifying a perpetrator, at a high volume blood scene. To gain important insight into this question different strategies need to be empirically evaluated, thereby creating an evidence-base for the decisions made at such a scene. Possible strategies could be investigated using realistic controlled experiments, using virtual or in-person mock crime scenes, as opposed to photographs as in the current study. Virtual/in-person mock scenes can enable CSIs to better imagine how things might have happened and where traces can be left. For example, this approach can be used to test the potential of different trace collection strategies like random collection vs. structured strategies like collecting in circles or lines vs. strategies that focus on the different blood stain patterns seen on the scene. To test and compare the outcomes of such strategies would allow conclusions as to whether one works better than another and thereby create an evidence-base for the decisions made by CSIs at crime scenes. The importance of sound collection strategies can hardly be overstated, considering, for example, that the collection dictates what will, and what will not, be analyzed forensically. The fact that such scientific basis is lacking can be an explanation as to the similarities in the reasoning between novices and experts in this regard. Possibly, the collection of blood happens more on the basis of the individual's common sense instead of scientifically established methodology. To avoid that important traces are left at the scene, it is essential that researchers and CSIs in collaboration examine scientifically valid methodologies for trace collection. Since both novices and experts appear to have been impacted by the Resource limitation and the irrelevant contextual information, the importance of contextual information management (CIM) models deserves to be mentioned. In other words, it is essential to make informed decisions about what CSIs should know when they enter a scene, and to also document and disclose of these decisions [79]. This assessment needs to be made by a context manager

on a case to case basis [80]. Importantly, a context manager's job is not to completely shield CSIs from contextual information but rather to make informed decisions about what should be disclosed to CSIs and when, and this needs to be based on research. Today, recommendations are starting to appear aiding context managers' assessments of what is or is not relevant (for an overview see [81]).

On top of this, the importance of individual differences in CSIs and the relevance for their decision making need to be further explored. This experiment provides no explanation of these differences since age, biological sex, years of experience, type of training or the country in which the CSIs operate were not significantly related with the dependent measures. Note that more senior CSIs (more experience, older) seemed to take slightly longer to complete the task, but their decision making was not significantly different from the less experienced and younger CSIs. Individual differences are not necessarily inherently a threat to the administration of justice. In fact, such differences can be beneficial if the dissent is effectively incorporated into decision making processes. If blood is only collected by one CSI, it appears the chances to identify a perpetrator, and outcome in a more general sense, may vary in important ways. It is possible that trace collection can be discussed among multiple CSIs, for example on the basis of photos, before the actual trace collection is conducted. Hence, this research has identified a few important lines for future research to address. Until such research is available, the awareness of contextual effects, resource limitations and individual differences may make CSIs and police and/or prosecution authorities rethink methodologies that they currently use, preferably in collaboration with researchers.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

The authors extend their gratitude to the participants in this study.

## Funding sources

This research was supported by Ragnar Söderberg's Foundation and the Swedish Research Council.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scijus.2023.01.005>.

## References

- [1] R.C. Wayment, The Role of the Civilian Scenes of Grime Officer, *J. Forensic Sci. Soc.* (1982), [https://doi.org/10.1016/S0015-7368\(82\)71523-3](https://doi.org/10.1016/S0015-7368(82)71523-3).
- [2] A. Ludwig, J. Fraser, R. Williams, Crime Scene Examiners and Volume Crime Investigations: An Empirical Study of Perception and Practice., *Foren. Sci. Policy Manag. An Int. J.* (2012), <https://doi.org/10.1080/19409044.2012.728680>.
- [3] D. Wyatt, Practising crime scene investigation: Trace and contamination in routine work, *Polic. Soc.* (2014), <https://doi.org/10.1080/10439463.2013.868460>.
- [4] C. Kruse, Swedish crime scene technicians: facilitations, epistemic frictions and professionalization from the outside, *Nord. J. Criminol.* (2020), <https://doi.org/10.1080/2578983X.2019.1627808>.
- [5] M. Lidén, *Confirmation Bias in Criminal Cases* (2021).
- [6] C. Roux, B. Talbot-Wright, J. Robertson, F. Crispino, O. Ribaux, The end of the (forensic science) world as we know it? The example of trace evidence, *Philos. Trans. R. Soc. B Biol. Sci.* (2015), <https://doi.org/10.1098/rstb.2014.0260>.
- [7] C. Roux, R. Bucht, F. Crispino, P. De Forest, C. Lennard, P. Margot, M.D. Miranda, N. NicDauid, O. Ribaux, A. Ross, S. Wills, The Sidney Declaration - Revisiting the Essence of Forensic Science through its Fundamental Principles, *Forensic Sci. Int.* 332 (2022).
- [8] R.M. Arthur, S.L. Cockerton, K.G. de Bruin, M.C. Taylor, A novel, element-based approach for the objective classification of bloodstain patterns, *Forensic Sci. Int.* (2015), <https://doi.org/10.1016/j.forsciint.2015.08.028>.



- [9] E. Bernstein, Science in Bloodstain Pattern Analysis, *Int. Assoc. Bloodstain Pattern Anal. News*. (2005).
- [10] R.M. Arthur, P.J. Humberg, J. Hoogenboom, M. Baiker, M.C. Taylor, K.G. de Bruin, An image-processing methodology for extracting bloodstain pattern features, *Forensic Sci. Int.* (2017), <https://doi.org/10.1016/j.forsciint.2017.05.022>.
- [11] W.A. Wagenaar, Anchored narratives: A theory of judicial reasoning, and its consequences, in: *Psychol. Law, Crim. Justice Int. Dev. Res. Pract.*, 1995.
- [12] I.E. Dror, A Hierarchy of Expert Performance, *J. Appl. Res. Mem. Cogn.* (2016), <https://doi.org/10.1016/j.jarmac.2016.03.001>.
- [13] M. De Gruitjer, C.J. De Poot, H. Elffers, The Influence of New Technologies on the Visual Attention of CSIs performing a Crime Scene Investigation, *J. Forensic Sci.* 61 (2016) 43–51.
- [14] L.A. Guarnera, D.C. Murrie, M.T. Boccaccini, Why do forensic experts disagree? Sources of unreliability and bias in forensic psychology evaluations, *Transl. Issues Psychol. Sci.* (2017), <https://doi.org/10.1037/tps0000114>.
- [15] R.D. Start, C.M. Milroy, M.A. Green, Suicide by self-stabbing, *Forensic Sci. Int.* (1992), [https://doi.org/10.1016/0379-0738\(92\)90151-L](https://doi.org/10.1016/0379-0738(92)90151-L).
- [16] T. Konopka, F. Bolechala, M. Strona, Chest stab wound comparison in suicidal and homicidal cases, *Arch. Med. Sadowej i Kryminol.* (2003).
- [17] F. Abdullah, A. Nuernberg, R. Rabinovic, Self-inflicted abdominal stab wounds, *Injury*. (2003), [https://doi.org/10.1016/S0020-1383\(02\)00084-0](https://doi.org/10.1016/S0020-1383(02)00084-0).
- [18] B. Karger, J. Niemeyer, B. Brinkmann, Suicides by sharp force: Typical and atypical features, *Int. J. Legal Med.* (2000), <https://doi.org/10.1007/s004149900093>.
- [19] J. Schnider, M.J. Thali, S. Ross, L. Oesterhelweg, D. Spendlove, S.A. Bolliger, Injuries due to sharp trauma detected by post-mortem multislice computed tomography (MSCT): A feasibility study, *Leg. Med.* (2009), <https://doi.org/10.1016/j.legalmed.2008.07.001>.
- [20] V.D. Singh, S.L. Lathrop, Youth suicide in New Mexico: A 26-year retrospective review, *J. Forensic Sci.* (2008), <https://doi.org/10.1111/j.1556-4029.2008.00721.x>.
- [21] G.M. Mazzolo, L. Desinan, Sharp force fatalities: Suicide, homicide or accident? A series of 21 cases, *Forensic Sci. Int.* (2005), <https://doi.org/10.1016/j.forsciint.2004.09.097>.
- [22] P. Vanezis, I.E. West, Tentative injuries in self stabbing, *Forensic Sci. Int.* (1983), [https://doi.org/10.1016/0379-0738\(83\)90091-9](https://doi.org/10.1016/0379-0738(83)90091-9).
- [23] P. Betz, E. Tutsch-Bauer, W. Eisenmenger, “Tentative” injuries in a homicide, *Am. J. Forensic Med. Pathol.* (1995), <https://doi.org/10.1097/0000433-199509000-00012>.
- [24] T. Karlsson, Multivariate analysis (‘Forensimetrics’) - A new tool in forensic medicine. Differentiation between firearm-related homicides and suicides, *Forensic Sci. Int.* (1999), [https://doi.org/10.1016/S0379-0738\(99\)00017-1](https://doi.org/10.1016/S0379-0738(99)00017-1).
- [25] K. Ormstad, T. Karlsson, L. Enkler, J. Rajs, Patterns in Sharp Force Fatalities—A Comprehensive Forensic Medical Study, *J. Forensic Sci.* (1986), <https://doi.org/10.1520/jfs12284j>.
- [26] M. Olajossy, A. Wysocka, M. Chuchra, Study of an atypical suicidal attempt, *Arch Med Sad Krym.* (1985).
- [27] S. Fukube, T. Hayashi, Y. Ishida, H. Kamon, M. Kawaguchi, A. Kimura, T. Kondo, Retrospective study on suicidal cases by sharp force injuries, *J. Forensic Leg. Med.* (2008), <https://doi.org/10.1016/j.jflm.2007.08.006>.
- [28] A. Marx, Die Bedeutung von Verletzungsbefunden für die Frage “Selbstmord oder Mord”, *Dtsch Z Ges Gerichtl Med* 2 412–421 (1923) (n.d.).
- [29] B. Karger, B. Vennemann, Suicide by more than 90 stab wounds including perforation of the skull, *Int. J. Legal Med.* (2001), <https://doi.org/10.1007/s004140100244>.
- [30] K. Lieske, K. Puschel, E. Bussmann, SUIZID DURCH 120 BRUSTSTICHVERLETZUNGEN? *Arch. Kriminol.* 1987.
- [31] J.F. Fekete, A.D. Fox, Successful Suicide by Self-Inflicted Multiple Stab Wounds of the Skull, Abdomen, and Chest, *J. Forensic Sci.* (1980), <https://doi.org/10.1520/jfs11267j>.
- [32] C. Rouge-Maillart, A. Tracqui, A. Le Bouil, A. Turcant, M. Penneau, Exceptional suicide by sharp force during mefloquine therapy: a case of drug induced psychosis?, (n.d.).
- [33] T. Kondo, T. Ohshima, Retrospective investigation of medico-legal autopsy cases involving mentally handicapped individuals, *Japanese J. Leg. Med.* (1995).
- [34] T.M.S. Neal, T. Grisso, Assessment Practices and Expert Judgment Methods in Forensic Psychology and Psychiatry: An International Snapshot, *Crim. Justice Behav.* (2014), <https://doi.org/10.1177/0093854814548449>.
- [35] R.M. Gardner, Defining a methodology for bloodstain pattern analysis, *J. Forensic Identif.* (2006).
- [36] J. Saviano, Articulating a concise scientific methodology for bloodstain pattern analysis, *J. Forensic Identif.* (2005).
- [37] M. Almazrouei, I.E. Dror, R. Morgan, Organizational and human factors affecting forensic decision-making, *J. Forensic Sci.* 1968–1977 (2020).
- [38] M.A. Almazrouei, R. Morgan, I.E. Dror, Stress and support in the workplace: The perspective of forensic examiners, *Forensic Sci. Int. Mind Law.* (2021).
- [39] A.K. Miller, K.A. Rufino, M.T. Boccaccini, R.L. Jackson, D.C. Murrie, On individual differences in person perception: Raters’ personality traits relate to their psychopathy checklist-revised scoring tendencies, *Assessment*. (2011), <https://doi.org/10.1177/1073191111402460>.
- [40] R.J. Homant, D.B. Kennedy, Subjective factors in the judgment of insanity, *Crim. Justice Behav.* (1987), <https://doi.org/10.1177/0093854887014001005>.
- [41] T.M.S. Neal, S.L. Brodsky, Forensic psychologists’ perceptions of bias and potential correction strategies in forensic mental health evaluations, *Psychol. Public Policy, Law.* (2016), <https://doi.org/10.1037/law0000077>.
- [42] A.M. Palker-Corell, Mental health professionals’ decision-making in competence for execution evaluations, *Diss. Abstr. Int. Sect. B Sci, Eng.* 2008.
- [43] M. Lidén, I.E. Dror, Expert Reliability in Legal Proceedings: Eeny, Meeny, Miny, Moe, With Which Expert Should We Go?, *Sci. Justice.* (n.d.).
- [44] D.C. Murrie, J.I. Warren, Clinician variation in rates of legal sanity opinions: Implications for self-monitoring, *Prof. Psychol. Res. Pract.* (2005), <https://doi.org/10.1037/0735-7028.36.5.519>.
- [45] W. Neil Gowensmith, D.C. Murrie, M.T. Boccaccini, Field reliability of competence to stand trial opinions: How often do evaluators agree, and what do judges decide when evaluators disagree? *Law Hum. Behav.* (2012) <https://doi.org/10.1037/h0093958>.
- [46] W.N. Gowensmith, D.C. Murrie, M.T. Boccaccini, How reliable are forensic evaluations of legal sanity? *Law Hum. Behav.* (2013) <https://doi.org/10.1037/lhb0000001>.
- [47] A. Sandras, C. Guilbeau-Frugier, F. Savall, N. Telmon, C. Capuani, Sharp bone trauma diagnosis: a validation study using epifluorescence microscopy, *Int. J. Legal Med.* (2019), <https://doi.org/10.1007/s00414-018-1944-z>.
- [48] N.K.P. Osborne, M.C. Taylor, M. Healey, R. Zajac, Bloodstain pattern classification: Accuracy, effect of contextual information and the role of analyst characteristics, *Sci. Justice.* (2016), <https://doi.org/10.1016/j.scijus.2015.12.005>.
- [49] S.G. Scott, R.A. Bruce, Decision-Making Style: The Development and Assessment of a New Measure, *Educ. Psychol. Meas.* (1995), <https://doi.org/10.1177/0013164495055005017>.
- [50] Supreme Court of Florida, Clemente ajvier Aguirre-Jarquín v. State of Florida, (2014).
- [51] I.S. Court, David R. Camm v. State of Indiana, (2009).
- [52] C.A.J. van den Eeden, C.J. de Poot, P.J. van Koppen, Forensic expectations: Investigating a crime scene with prior information, *Sci. Justice.* (2016), <https://doi.org/10.1016/j.scijus.2016.08.003>.
- [53] M.C. Taylor, T.L. Laber, P.E. Kish, G. Owens, N.K.P. Osborne, The Reliability of Pattern Classification in Bloodstain Pattern Analysis, Part 1: Bloodstain Patterns on Rigid Non-absorbent Surfaces, *J. Forensic Sci.* (2016), <https://doi.org/10.1111/1556-4029.13091>.
- [54] M.C. Taylor, T.L. Laber, P.E. Kish, G. Owens, N.K.P. Osborne, The Reliability of Pattern Classification in Bloodstain Pattern Analysis—PART 2: Bloodstain Patterns on Fabric Surfaces, *J. Forensic Sci.* (2016), <https://doi.org/10.1111/1556-4029.13191>.
- [55] N.K.P. Osborne, M.C. Taylor, R. Zajac, Exploring the role of contextual information in bloodstain pattern analysis: A qualitative approach, *Forensic Sci. Int.* (2016), <https://doi.org/10.1016/j.forsciint.2015.12.039>.
- [56] R.S. Nickerson, Confirmation bias: A ubiquitous phenomenon in many guises, *Rev. Gen. Psychol.* (1998), <https://doi.org/10.1037/1089-2680.2.2.175>.
- [57] S. Andrews, Lexical Expertise and Reading Skill, *Psychol. Learn. Motiv. - Adv. Res. Theory.* (2008), [https://doi.org/10.1016/S0079-7421\(08\)00007-8](https://doi.org/10.1016/S0079-7421(08)00007-8).
- [58] B.L. Bartels, Top-Down and Bottom-Up Models of Judicial Reasoning, *Psychol. Judic. Decis. Mak.* (2010), <https://doi.org/10.1093/acprof:oso/9780195367584.003.0003>.
- [59] Y.-W. Ha, S.J. Hoch, Ambiguity, Processing Strategy, and Advertising-Evidence Interactions, *J. Consum. Res.* (1989), <https://doi.org/10.1086/209221>.
- [60] N.K.P. Osborne, R. Zajac, An Imperfect Match? Crime-related Context Influences Fingerprint Decisions, *Appl. Cogn. Psychol.* (2016), <https://doi.org/10.1002/acp.3180>.
- [61] O. Ribaux, A. Baylon, E. Lock, O. Delémont, C. Roux, C. Zingg, P. Margot, Intelligence-led crime scene processing. Part II: Intelligence and crime scene examination, *Forensic Sci. Int.* (2010), <https://doi.org/10.1016/j.forsciint.2010.03.011>.
- [62] A. Baylon, L’Utilisation du renseignement forensique pour guider les décisions liées à l’investigation de scène de crime, *Rev Suisse Criminol.* (2012) 40–46.
- [63] T. Resnikoff, O. Ribaux, A. Baylon, M. Jendly, Q. Rossy, The polymorphism of crime scene investigation: An exploratory analysis of the influence of crime and forensic intelligence on decisions made by crime scene examiners, *Forensic Sci. Int.* (2015), <https://doi.org/10.1016/j.forsciint.2015.10.022>.
- [64] M. Lidén, Confirmation bias in criminal cases, 2018.
- [65] A. Jamieson, A rational approach to the principles and practice of crime scene investigation: I. Principles, *Sci. Justice - J. Forensic, Sci. Soc.* (2004), [https://doi.org/10.1016/S1355-0306\(04\)71678-0](https://doi.org/10.1016/S1355-0306(04)71678-0).
- [66] I. Dror, The ambition to be scientific: Human expert performance and objectivity, *Sci. Justice.* (2013), <https://doi.org/10.1016/j.scijus.2013.03.002>.
- [67] I. Dror, Letter to the Editor-Combating Bias: The Next Step in Fighting Cognitive and Psychological Contamination, *J. Forensic Sci.* (2012), <https://doi.org/10.1111/j.1556-4029.2011.01940.x>.
- [68] J.E. Laurin, Remapping the path forward: Toward a systemic view of forensic science reform and oversight, *Tex. Law Rev.* 2013.
- [69] R. Williams, P. Johnson, Trace biometrics and criminal investigations, *Handb. Crim. Investig.* (2015), <https://doi.org/10.4324/9780203118177.ch14>.
- [70] N. Tilley, M. Townsley, Forensic science in UK policing: Strategies, tactics and effectiveness, in: *Handb. Forensic Sci.*, 2009.
- [71] T. Bevel, R.M. Gardner, Bloodstain Pattern Analysis With an Introduction to Crime Scene Reconstruction, in: 2008.
- [72] A. Field, *Discovering Statistics using IBM SPSS Statistics*, Sage (2012).
- [73] J. Correll, C. Mellinger, E. Pedersen, Flexible approaches for estimating partial eta squared in mixed-effects models with crossed random factors, *Behav. Res. Methods.* 54 (2022) 1626–1642.
- [74] J. Cohen, *Statistical power analysis for the behavioural sciences*, Academic Press, 1988.
- [75] M. Saks, Will the introduction of an emotional context affect fingerprint analysis and decision-making? *Forensic Sci. Int.* 191 (2009) 36–39.

- [76] I. Levy, J. Snell, A. Nelson, A. Rustichini, W. Glimcher, P. Neural Representation of Subjective Value under Risk and Ambiguity, *J. Neurophysiol.* 103 (2010).
- [77] S. Raptis, J. Chen, F. Saposnik, R. Pelyavskyy, A. Liuni, G. Saposnik, Aversion to Ambiguity and Willingness to Take Risks Affect Therapeutic Decisions in Managing Atrial Fibrillation for Stroke Prevention, *Patient Prefer, Adherence.* (2017) 1533–1539.
- [78] I.E. Dror, Cognitive and Human Factors in Expert Decision Making: Six Fallacies and the Eight Sources of Bias, *Anal. Chem.* (2020).
- [79] M. Almazrouei, I.E. Dror, R. Morgan, The Forensic Disclosure Model: What Should be Disclosed to, and by Forensic Experts? *Int. J. Law, Crime Justice.* 59 (2019), 100330 <https://doi.org/10.1016/j.ijlcrj.2019.05.003>.
- [80] N.K. Osborne, M. Taylor, Contextual Information Management: An Example of Independent-Checking in the Review of Laboratory-based Bloodstain Pattern Analysis, *Sci. Justice.* (2018) 226–231.
- [81] M. Lidén, *Confirmation Bias in Criminal Cases*, Oxford University Press, n.d.