Digital Investigation 29 (2019) 101–108

Contents lists available at ScienceDirect

Digital Investigation

journal homepage: www.elsevier.com/locate/diin

Cognitive and human factors in digital forensics: Problems, challenges, and the way forward

Nina Sunde [a, *], Itiel E. Dror [b]

[a] University of Oslo, Norwegian Police University College, Norway
[b] University College London, UK

ARTICLE INFO

Article history:
Received 4 February 2019
Received in revised form
24 March 2019
Accepted 26 March 2019
Available online 29 March 2019

Keywords:
Digital forensics
Digital investigation
Digital evidence
Forensic science
Cognitive bias
Human error
Human factors
Expert decision making

ABSTRACT

Digital forensics is an important and growing forensic domain. Research on miscarriages of justice and misleading evidence, as well as various inquires in the UK and the US, have highlighted human error as an issue within forensic science. This has led to increased attention to the sources of cognitive bias and potential countermeasures within many forensic disciplines. However, the area of digital forensics has yet to pay sufficient attention to this issue. The main goal of this article is to contribute to a more scientifically sound digital forensics domain by addressing the issues of cognitive bias as a source of error. In this paper we present an analysis of seven sources of cognitive and human error specifically within the digital forensics process, and discuss relevant countermeasures. We conclude that although some cognitive and bias issues are very similar across forensic domains, others are different and dependent on the specific characteristic of the domain in question, such as digital forensics. There is a need for new directions in research with regard to cognitive and human factors in digital forensics.

Introduction

With the extensive and growing use of technology in everyday life (both by law enforcement and by criminals), the importance of digital forensics and the reliance on digital evidence will continue to grow. Therefore, it is important to consider how this forensic discipline can be made as robust and reliable as possible. Digital forensics began as an ad hoc, task oriented practice performed by computer professionals with no formal processes, tools or training (Garfinkel, 2010). In recent years there has been an advance within digital forensics towards a more scientifically sound handling of the evidence by enhanced focus on quality management (Page et al., 2018), error mitigation, tool testing and verification methodologies (Casey, 2011; ENFSI, 2015; Pollitt et al., 2018; SWGDE, 2018). Digital forensics is now recognized as a discipline within forensic science by a variety of organisations (e.g., the Australian National Institute of Forensic Science (NIFS), the European Network of Forensic Science Institutes (ENFSI) and the United States National Institute of Standards and Technology (NIST).

Within the digital forensics domain the term "forensically sound" is used extensively (e.g. a search on Google Scholar on "digital evidence" + "forensically sound" returns 1610 hits). An analysis of how the term was understood in the digital forensics literature, linked the quality of which the digital forensics process was conducted to the admissibility of the digital evidence in the court of law (McKemmish, 2008). Forensically sound was defined as: "The application of a transparent digital forensics process that preserves the original meaning of the data for production in a court of law" (McKemmish, 2008, p. 10). Forensically sound reflects a threshold of minimum requirements for court presentation, inherently accepting that using only rigid and scientifically sound procedures for producing the evidence is often unachievable within the digital forensics domain. However, the quality threshold for presenting digital evidence in court should not be confused with the goal of advancing the domain towards a more *scientifically sound* domain, by adopting scientifically sound procedures for handling digital evidence which take both technological and human factors into account.

Like other forensic disciplines, and science in general, there are uncertainties, limitations, vulnerabilities, and potential for error. It is known from evaluations of wrongful conviction cases in the US, Australia, and England and Wales that from 23% to 60% of the cases

* Corresponding author.
E-mail addresses: nina.sunde@phs.no (N. Sunde), i.dror@ucl.ac.uk (I.E. Dror).

https://doi.org/10.1016/j.diin.2019.03.011

included flawed or exaggerated forensic evidence contributing to the convictions of innocent people (Dioso-Villa, 2012, 2015; Garrett, 2011; Garrett and Neufeld, 2009; National Registry of Exonerations, 2015; Smit et al., 2018).

Digital evidence is perceived as reliable and correct by many in the legal community (Van Buskirk and Liu, 2006). This should be a concern when considering the results of a recent analysis of the quality management (QM) procedures in digital forensics performed by Page et al. (2018). They compared the QM procedures between bodily fluids and DNA, fingerprint and digital forensics in the UK, and found that digital forensics operates with the least robust QM procedures.

We first distinguish between errors which arise from technology, for example errors in timestamps or loss of data (e.g. Casey, 2002; Ekfeldt, 2016; SWGDE, 2018), and those related to the lack of scientific data, research and rigor (SWGDE, 2018). In addition, and in contrast to the technological and scientific issues, there are also errors that arise from cognitive and human factors. These cross all forensic domains (as well as other expert domains). Within forensic science, there has been an increased attention towards cognitive and human factors as a source of error, with a growing body of research in this area (see below, section 2.1).

Within the digital forensics domain, there has been a movement from perceiving tools and technology as the main instruments in the digital forensics process towards greater acknowledgement of *the human* as an important instrument for examining digital evidence (e.g. ENFSI, 2015; Pollitt et al., 2018). However, as of today this issue has not been sufficiently addressed through research within this domain. When addressing human factors and human error, it is important to look into the cognitive factors that shape perception and interpretation and how they may impact the experts' decision making.

A critical element is cognitive bias, understanding where it may come from, and how to control and minimize it. Pohl (2016) explains cognitive bias as a cognitive phenomenon which reliably deviates from reality, occurs systematically and involuntarily, and is difficult or impossible to avoid by mere willpower. The main goal of this article is to contribute to a more scientifically sound digital forensics domain by addressing the issues of cognitive bias as a source of error. In this paper we examine the possible human sources of errors in the digital forensics process, and specifically the role of the human experts in making forensic decisions within this area. The research question addressed here is: When handling digital evidence through the digital forensics process; when is the digital forensics practitioner (DFP) vulnerable to cognitive bias, and what measures could be relevant and effective to mitigate bias for this specific domain?

To answer this question, we first present relevant research concerning bias and human error in forensic science. An analysis of cognitive and human error within the digital forensics process is then conducted, and some relevant countermeasures are discussed. We conclude with the need for future research and new directions with regard to cognitive and human factors in digital forensics.

Cognitive and human factors in forensic science

For decades, there was little research or attention to the human examiner in any forensic domain. This has changed, and over the past decade there has been much attention and research on bias and other cognitive factors in many forensic disciplines. However, the area of digital forensics has remained as one of the few forensic disciplines that has not sufficiently researched or given much attention to this issue. Since the digital forensics discipline shares many core forensic processes with other forensic disciplines (Pollitt et al., 2018), research from such disciplines are a good and relevant starting point for exploring the issue of cognitive and human error.

Research on bias and forensic science

For the disciplines dealing with identification, for example DNA and fingerprint analysis, there is a solid research foundation showing that contextual information can distort the forensic analyst's judgement. Dror and Hampikian (2011) tested whether DNA experts could be biased by the details of the case. After a gang rape, evidence in form of a mixture of DNA from multiple contributors was secured as evidence and analysed. This analysis required some judgement and interpretation from the forensic analyst, as is often the case in DNA mixture analyses. One of the suspects testified against other suspects as part of a plea bargain, and this was known to the two analysts who examined the DNA evidence. Both analysts concluded that the suspect in question *could not be excluded* from being a contributor to the DNA mixture. To examine the possible biasing effect of this irrelevant contextual information (the testimony against the suspect), Dror and Hampikian (2011) presented the original DNA evidence to 17 experts, but excluded the irrelevant contextual information. The vast majority of experts did not agree with the original examiners: Only 1 of the 17 agreed with the analysts who were exposed to the biasing information. Without the irrelevant biasing context 16 examiners did not conclude that the suspect could not be excluded.

Within the fingerprint domain, Dror and Charlton (2006) presented fingerprint experts with evidence they had examined before (without their knowledge), but did so with new irrelevant contextual information. Their data demonstrated that fingerprint experts' decisions could be biased by irrelevant contextual information (see also Dror et al., 2006; Dror and Rosenthal, 2008). These findings have now been replicated by Stevenage and Bennett (2017) who found that knowledge of the results of a DNA test affected the decisions in fingerprint matching tasks, as well as by Smalarz et al. (2016) who found that even criminal stereotypes can bias forensic evidence analysis.

The biasing effect of irrelevant contextual information has also been found in research within other forensic domains, such as bloodstain pattern analysis (Taylor et al., 2016), arson investigation (Bieber, 2012), forensic pathology (Oliver, 2017), and crime scene investigation (van den Eeden et al., 2016). Similar effects have also been found in a number of studies within the domain of forensic anthropology (Nakhaeizadeh et al., 2014, 2018). For a summary of much of this research across forensic domains, see (Dror, 2016).

Compatible with a huge body of research in many scientific domains, the empirical research in various forensic disciplines has predominantly shown that human examiners do not always produce consistent results, and that they are susceptible to cognitive bias. In the context of a criminal investigation, biased conclusions may lead to errors in the fair administration of justice.

Research on misleading forensic science evidence

Errors originating from several forensic disciplines have caused or contributed to wrongful convictions in several countries (for an overview, see Smit et al., 2018). This issue has been subject to several examinations, aimed at trying to establish the causes that underlie miscarriages of justice. According to The National Registry of Exonerations, false or misleading forensic evidence were found in 23% of the 1600 individual exonerations in US from 1989 to 2015 (National Registry of Exonerations, 2015). Garrett and Neufeld (2009) looked at 137 cases where innocent persons were convicted for serious crimes such as rape or murder, and later exonerated by post-conviction DNA testing. They found that in 60% of these cases forensic evidence contributed to the wrongful

conviction. In Australia, Dioso-Villa (2012, 2015) looked into 71 cases of wrongful conviction, and found that 31% of the cases involved forensic errors or misleading forensic evidence. In the UK, a study by Smit et al. (2018) found that forensic evidence (including digital evidence) was an issue in 32% of the 235 cases they looked at in which the Court of Appeal quashed a conviction. With regards to digital evidence in particular, this evidence type is only highlighted in the UK study. Smit et al. identified 4 individual cases where presentation of CCTV recordings, SIM-card, DVD content, and web content were the misleading elements (personal communication with N. Smit, 5th of April, 2018).

To summarize, research from several forensic disciplines, as well as real criminal cases, has demonstrated that experts conducting forensic examination that involve subjectivity, interpretation or opinion are susceptible to cognitive bias. Based on the research in other forensic disciplines, as well as research in many other expert domains, and understanding of human cognitive processing, there is no reason to believe that DFPs are more reliable "instruments" than other forensic experts, nor that they are immune to bias. However, the nature of the activities and processes when analysing digital evidence differs from other disciplines, and therefore we need to examine the unique situations in digital forensics with particular risk of bias, and identify ways to specifically combat them.

In the next section we examine the various sources of bias in relation to the digital forensics process as described by Flaglien (2018), with the phases of identification, collection, examination, analysis and presentation. Within the digital forensic process, both investigative and forensic processes may be carried out, and the issues raised in this paper apply to both. Experiences from casework will be used to illustrate and demonstrate the issues, including data from interviews of Norwegian DFPs and criminal detectives involved in investigations with digital evidence.

A taxonomy of sources of bias that may affect forensic decisions contextualized within the digital forensics process

The seven level taxonomy (Fig. 1) elaborates on several factors that could lead to bias, each of which entails different countermeasures (Dror, 2017). It is important to emphasize and clarify that we are talking about *cognitive bias*, and not intentional biases or errors that arise from lack of training, motivation or incompetence. The cognitive biases arise from how the brain processes information (Nickerson, 1998). Cognitive biases are sometimes mistaken as ethical issues. This is not correct, since they are the result of computational tradeoffs occurring in the brain, and not the result of intentional or conscious acts (Dror, 2014). Experts are not immune to such biases (Dror, 2011), but they are mostly unaware of them due to the *bias blind spot* (Pronin et al., 2002), which has recently been demonstrated specifically within forensic science (Kukucka et al., 2017). Below we examine each of the levels from the taxonomy in Fig. 1 in relation to activities conducted within the digital forensics process.

The term *digital evidence* covers both digital devices as well as relevant pieces of information (data) found on the devices. E.g. a collected digital device such as a laptop or a smartphone will often, by itself, be considered as evidence. During the digital forensics process, the information on the digital device is analysed, and information such as images, documents, chat conversations etc. will be documented, and constitutes separate pieces of evidence.

Cognitive architecture and the brain

The base level of the taxonomy, at the very bottom, addresses the brain, and our cognitive architecture. The human brain has limited capacity to process all the information it is presented with, and uses several strategies to cope with this. Zapf and Dror (2017) describes one such strategy, *chunking*, which means to bind individual pieces of information together, so they are represented as one within our mental representation. For example, when trying to memorize a phone number, instead of memorizing the number sequence 98348758, it could be divided in three chunks; 983-48-758, making the information easier to retain and recall. Another strategy is *selective attention*, meaning that we attend to a specific piece of information while ignoring other information. *Top down processing* refers to a whole set of cognitive processes that is conceptually driven by what is already in our brains (i.e., our knowledge, expectations, hope, fears, personality, prior experiences, etc.). The top-down processing uses context to make sense of the incoming information (Lindsay and Norman, 2013; Zapf and Dror, 2017). These and other cognitive strategies facilitate efficient information handling to a person in everyday life, and are used more as we gain experience and expertise. These mechanisms are very helpful, useful and needed (and mostly out of our control

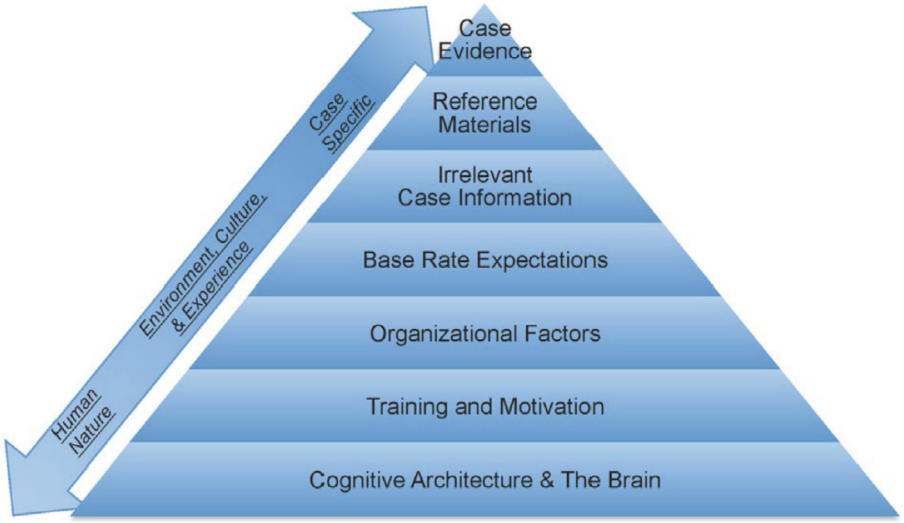

Fig. 1. Influences that might interfere with accurate observations and inferences in forensic decision making (Dror, 2017).

and awareness), but they can also degrade our impartiality and cause bias. Contrary to what many believe, in various ways an expert is actually more susceptible to bias compared to novices due to the underpinning of expertise (Dror, 2011).

The list of biases described in research literature is quite substantial. We will focus on three biases that are particularly relevant to the work of a forensic examiner. *Anchoring bias* is due to information we encounter first, making it more influential than information that is received later (Tversky and Kahneman, 1974). Once we get the first piece of information, we may fixate, escalate commitment and can have tunnel vision. *Availability bias* makes us overestimate the likelihood of an outcome based on our ability to recall similar instances (Tversky and Kahneman, 1973). In the process of making sense of what is observed, a hypothesis will often evolve (often subconsciously at first) (Kahneman, 2011), leading to *confirmation bias* (Nickerson, 1998), which is the tendency to:

a) seek information that supports the hypothesis
b) interpret new information in relation to the hypothesis
c) interpret ambiguous or neutral information as supportive to the hypothesis
d) ignore or explain away information that contradicts the hypothesis
e) give little weight to information that does not support the hypothesis

Cognitive bias is also impacted by emotions, such as confidence, frustration, sorrow and anger, personal responsibility, and concern about future consequences (Ask, 2013). When the biases influence the DFP, they might lead to several effects, which in turn might introduce errors in the digital forensics work.

At a physical crime scene, there is a constant flow of new information from the observations at the scene, which activates top-down processing, influenced by what the DFP already knows, has experienced before or expects from the situation. The mind is constantly trying to make sense of the situation, filling in the gaps where information is missing, giving meaning to ambiguous information, and so on. What the DFP observes during a search may affect the interpretation of the latter due to anchoring and other biases. During all the phases of the digital forensics process the hypothesis under consideration has profound effects on the DFPs' observations and conclusions. If a person is suspected of committing a criminal act, the confirmation bias might lead to an aggregated focus within the investigative or forensic tasks on finding information consistent with the guilt hypothesis, while a tendency to overlook or explain away information that contradicts it, such as exonerating information or information indicating mitigating circumstances.

Training and motivation

Moving up the pyramid (Fig. 1), the next three levels relate to the environment, culture and experience. These are about our personal nature, our motivations and preferences, and are developed throughout our lives. They affect how we observe, reason and make decisions (Balcetis and Dunning, 2006). Motivation is a source of erroneous decision making in criminal investigation (Ask and Granhag, 2005). Pre-existing attitudes and with whom the individual identifies might also sway observations and conclusions (Neal, 2016).

The DFP often will come across sexual crimes against children, since the evidence in such cases often is digital images, videos or online communication. The pre-existing attitudes of the DFP may affect the analysis of digital evidence. A motivation for "saving" a child by finding evidence against a perpetrator who has harmed children in the past might be a strong influential factor on their analysis.

Concerning the issue of training, digital forensics has roots in computer science, physics (electronics) and mathematical theory (ENFSI, 2015). Within these sciences and theories, the role of the cognitive processes and the human brain are usually not given much (if any) attention in the curriculum. The typical DFP will most often lack knowledge of cognitive mechanisms, biases and relevant countermeasures. This was apparent in the interviews of Norwegian DFPs with civilian backgrounds (Sunde, 2017). The interviews showed that they had no formal knowledge of bias and relevant countermeasures from their educational background or professional training. An important issue concerning human error and the digital forensics work is the lack of formal qualification and certification requirements. Page et al. (2018) found that compared to the DNA and fingerprint domains, the digital forensics discipline "is operating under arguably less rigorously defined standards, practitioner governance and evidence validation procedures" (Page et al., 2018, p. 2).

Organisational factors

The next level of the taxonomy (Fig. 1) is about organisational and cultural factors, which implicates social interactions, identification with ideologies and organizations, and communication with others. The words we use, the terminology, vocabulary and jargon can cause errors in how we interpret and understand information (Zapf and Dror, 2017). The information flow between different systems, organisations and entities in an organisational structure may also affect the human examiner and bias their work. Most concerning is how bias can cascade and snowball between different, supposedly independent, elements involved in criminal investigations (Dror, 2018).

The overall organization of the digital forensics work, as well as cooperation and communication procedures are relevant aspects to discuss in relation to bias. How digital forensics work is organised varies a lot, from organisations outside law enforcement providing digital forensics as a service, to digital forensics integrated within the law enforcement organisation. Norway and Finland are examples of the latter (Leppänen and Kankaanranta, 2017; Sunde, 2017). In Norway, the DFPs often work in close cooperation with both the investigation team and the prosecution (Sunde, 2017). The defence may hire private DFPs to do own analysis, and to challenge the relevance or reliability of the digital evidence presented by the prosecution in court. Research from field studies and experimental studies has showed that whether the forensic experts believed they were working for the prosecution or the defence affected their forensic conclusions (Murrie et al., 2009). This has been termed *adversarial allegiance*. This implies that although the DFP is obliged to be impartial and objective during investigation and trial, whether DFPs are working for the prosecution or the defence may affect how they observe and what they conclude about digital evidence, and how they present the evidence in court.

In relation to cooperation and communication procedures, integrating and combining digital forensics within the investigation team is sometimes regarded as an advantage. The experience from a project in the Oslo Police District was that close cooperation between the DFP and the tactical investigation team led to increased motivation, knowledge exchange and efficient communication in the investigation of sexual crimes against children (Hansen et al., 2017). In Finland, close cooperation was highlighted as a success factor behind a cost-efficient investigation of computer-integrity crimes (Leppänen and Kankaanranta, 2017). However, a downside of close cooperation between the DFP and the tactical investigation

team is the flow of biasing irrelevant information within the team, such as the results of DNA analyses, witness statements, eyewitness identifications or whether the suspect has confessed to the crime, which might lead to bias cascade and bias snowball (Dror, 2018).

Whether the forensic work is done within law enforcement or by independent laboratories, either way, the critical question is whether the forensic examiners had independence of mind during their work. Organizational factors and biasing influencing are not limited to law enforcement, they exist everywhere. Several people have tried to develop integrated models for digital forensics and criminal investigation (e.g. Ciardhuáin, 2004; Hunton, 2011; Sun et al., 2015) but none of these has addressed organisation structures (Leppänen and Kankaanranta, 2017) or taken the issue of bias mitigation into particular consideration.

Base rate expectations

The next level in the taxonomy is base rate expectations, which is about how past regularities bring about an expectation to what will be found in a new case. These are expectations for the current case based on past experience or from work with other cases in the past (Dror, 2017). Of course, the past cases can be very helpful, but they can also be misleading. This is a bias, because the impact has nothing to do with the current case at hand.

Within the identification phase of the digital forensics process, the decision about seizure may be regarded as a particularly vulnerable to bias. This phase involves search for, recognition and documentation of potential digital evidence (ASTM, 2015). Before the examination, the digital device is often a "black box" where information that could be relevant to the investigation is hidden. There are available tools such as hardware and software write blockers that enables the DFP to access and preview information from the device without altering the data (triage). However, this is not always possible due to missing tools, expertise in using the tools, or conditions on the search scene. Thus, the decision about seizure must often be made under some level of uncertainty, and base rate expectations might affect and bias the decision.

Since the base rate expectations arise from the examiners' former experience with other cases (Dror et al., 2015), having experience with not finding anything of interest on a particular type of device might result in a decision not to seize the device. There are several other important decisions to be made throughout the digital forensic process. In the collection phase the decision about how to collect (e.g. live or pull-the-plug) might also be influenced by expectations from former experience about what was successful or what effort did not pay off. In the examination phase the choice of processing software might be conducted due to the expectations about what will be found, and the base rate expectations could also affect how much effort the DFP dedicate to unpacking compressed, archived or decrypted files.

The base rate expectations are formed by previous casework, and they could create expectancies about conclusions and outcomes prior to the analysis, and may also affect what traces to look for during the analysis. This is in many ways what makes an expert, but it can also be misleading and bias them in the wrong direction when the specific case in front of them is different than previous cases. The base rate expectations degrade various aspects in the decision making process, e.g., the ability to look for, and find, critical information based on the case at hand. When the case at hand is comparable to previous cases the base rate expectations will not cause problems – they may even be helpful. However, when the case at hand has unexpected critical information, it may be overlooked and missed.

Irrelevant case information

The four previous levels are independent to the case at hand, but the three top levels of the taxonomy relate to the specific case under investigation. The case specific biasing factors are: irrelevant case information, reference materials and actual case evidence. Irrelevant case information, is about the contextual information concerning the particular case, but which is irrelevant for the referral question (i.e., not relevant to the experts work and expertise). One of the most potentially biasing considerations at this level involves the inferences made by others (Zapf and Dror, 2017). Irrelevant case information may cause *bias cascade* or *snowball effects* (Dror, 2018). The cascading effect is when a bias cascades from one stage to another within the process of handling the evidence, e.g., from the collection to the analysis stage. The snowball effect is when the bias increases as irrelevant information from a variety of sources (e.g., erroneous identification of suspect by witness, arrest of suspect, confession, etc.) is integrated and influences each other. As more people are influenced, they in turn influence others, and hence the bias increases, and gathers more momentum as more people are affected by it and then affect others, hence the snowball effect.

Within digital forensics the issue of irrelevant case information has several aspects, regarding where it occurs and how it could be managed. Irrelevant case information regarding digital evidence may be the formal case file, but also includes informal knowledge concerning the case from those involved in the investigation. This could be factual information, e.g. whether the suspect has confessed or results from analysis of other lines of evidence. Irrelevant case information could also be subjective information such as sharing of personal opinions about which offence hypothesis they consider the most likely, how they feel about the suspect, or their sympathy for the victim. If the case gets attention from the media, further biasing irrelevant case information may be found there. Due to insufficient quality management systems (Page et al., 2018), information irrelevant to the work of the DFP may be poorly managed. This is particularly problematic when the DFP is integrated within the criminal investigation team. When the DFP is conducting the collection on the search scene and then does the analysis, the bias cascades, and irrelevant contextual information from the case and other lines of evidence may influence the analysis.

In contrast to fingerprint and firearm analysis, handwriting examination, and many other forensic domains, digital evidence analysis often *requires* more context. For example in the fingerprint domain, the expert may conduct the identification process with limited contextual information, restricted to only how the prints were lifted and developed, from what surface the prints were collected, etc. In the digital forensics domain the situation is quite different. In the past digital storage devices had limited capacity, and it was possible to examine them from start to end. Now, due to the amount and complexity of data, this approach is often unfeasible. As reflected in an interview of a DFP, such situations require the use of different search strategies, and hence it was important to also report not only the result, but also what was searched for (Sunde, 2017). This is done to avoid misunderstandings about negative findings. For example, a search for images of sexual abuse of children with negative result, a null finding, could not exclude the possibility that such images were on the digital device but were just not found. Such strategies often require some context to be sufficiently customised for the particular information requirement, and a completely blinded examination of a hard drive would be inefficient and lengthy.

The case file might contain relevant information for a targeted analysis of the digital seizure. However, there would also be

irrelevant information in the case file. The case information available to the DFP would affect what they would look for, and how they would perceive the relevance of the potential evidence they observe. For instance multiple Norwegian DFPs said that they would normally have access to the case file, and would sometimes read the suspect interviews prior to their analysis (Sunde, 2017). Depending on the task the DFP is required to do, the information from this interview could be relevant or irrelevant, as well as biasing.

Reference material

Several of the forensic disciplines conduct tasks involving reference material. In an investigation involving firearms, the forensic analyst might be tasked to compare the breech face mark found on a cartridge at a crime scene with a test fire cartridge from a weapon found at the house of the suspect. The reference material itself could be biasing, especially when the forensic examiners works backwards, from the suspect to the evidence. In digital forensics, reference material could be used when, for example, the DFP is tasked to determine whether a particular file, such as an image, a document or a video is present on a computer. The task would be performed by calculating a checksum and searching for a match. In such occasions there would be very little need for interpretation by the DFP, and hence a low risk of a biased decision.

However, there are tasks where the reference materials could bias the interpretation of the actual evidence. The digital evidence is often used for cross-checking information retrieved through the investigation, for example from police interviews. When reading the interview, the DFP would often be exposed to more (biasing) information than what is necessary to conduct the cross-checking task. To minimize potential bias the DFP does not need the full statement, nor all the details, but some key information from the statement. The reference material constitutes a *target driven bias* where it, rather than the actual evidence, is guiding the cognitive process. For example, in DNA analyses, this occurs when the suspect's profile influences how the biological material from the crime scene – the actual evidence - is interpreted (Jeanguenat et al., 2017). This could occur in digital forensics, for example, when the DFP has an image of a child who has been kidnapped and now is suspected to be present in video. This may influence how the image of a child in the digital evidence is interpreted to fit the target image.

Case evidence

On the very top of the taxonomy is the case evidence itself. This category includes different types of evidence that was obtained through the investigation of the crime. Whether the piece of evidence could be a source of bias depends on whether it inherently includes biasing information within the evidence itself. A latent fingerprint or a firearm cartridge by itself does not convey any biasing information (e.g., a cartridge, by itself, does not say whether this was a misfire from a firearm, or if it was involved in a homicide). In contrast, a seized digital *device* might by itself be biasing, since it may have information such as nametags, stickers or other characteristics which could affect the observations and decisions of the DFP. Furthermore, the *evidence in itself*, in the form of information found through the examination of the data on the digital device, could be biasing. Examples of such could be images of sexual abuse, chat logs about the planning of a terrorist attack, or an internet search log conveying search phrases about how to get away with murder or how to spike a drink with drugs to commit a rape.

To summarize, the analysis of the seven level taxonomy in relation to digital forensics indicates that human cognition is involved to a great extent in the work of the DFP. Digital forensics has specific sources and ways that can cause bias. Conducting the digital forensics process encompasses many observations, judgements and conclusions which involve subjectivity, interpretation and opinion by the DFP. Hence, none of the phases of the digital forensics process should be considered safe against cognitive bias or other errors caused by human cognition. In the following section we will look into ways in which bias may be mitigated, based on research and examples from other forensic disciplines.

Bias mitigating countermeasures with relevance to digital forensics

Every examination and evaluation involving a human judgement is prone to bias (Zapf and Dror, 2017). However, research has identified several promising cognitive practices and operating procedures for mitigation of biasing effects, and their relevance and feasibility for the digital forensics domain should also be considered. First and foremost, to understand how observations and conclusions may be biased, we believe that relevant cognitive psychology should be included in the *training* of the DFP. To counter the bias blind spot, the training should be practical and scenario-based, allowing the DFPs themselves to understand and experience how bias could occur without them noticing it in the normal work situation within the digital forensics domain.

The DFP, similar to criminal investigations and scientific inquiry in general, should conduct their examination by testing and eliminating *multiple*, and preferably *competing* hypotheses. A similar bias mitigation strategy is used in well managed line-ups. Instead of presenting the suspect alone, the eye-witness is exposed to multiple possible offenders together with the suspect (Steblay et al., 2003; Wells et al., 1998). The hypotheses should be defined on either offence, activity or source level, (Cook et al., 1998), and constructed on the basis of relevant information from the case. A systematic collection and consideration of relevant data to test alternative hypotheses and a careful documentation of which piece of information is being evaluated, and whether it weakens or strengthens the alternative hypotheses may help to minimize confirmation bias. This is supported by the results from experiments conducted by Rassin (2018), who found that using a pen-and-paper tool for evaluation of evidence against two competing hypotheses could reduce tunnel vision.

Irrelevant case information as a source of bias may be eliminated by thorough *context management* that can control what information is given, when, to whom, and how (Dror, 2014; Dror et al., 2015; Stoel et al., 2014). This issue, if understood, can be managed through the use of case managers (Dror, 2014), who should know everything about the case, and interact with the investigation team throughout, but they would not carry out the actual forensic analysis.

Relevant case information could be biasing, and a measure to simply avoid or isolate it is not necessarily the right course of action due to the role it may play in examination and evaluation of the evidence. Hence, a detailed analysis needs to be done to determine how to handle such situations (Dror, 2012). Rather than just blinding the examiner to information, there are approaches that control when the information is given. Dror et al. (2015) developed the Linear Sequential Unmasking (LSU) procedure, aimed at mitigating bias caused by relevant case information through *exposure control*. The LSU procedure is aimed to ensure the cognitive process goes in the right direction, from the evidence to the suspect, by regulating when the examiner should be exposed to the biasing case relevant information, and provides restrictions to when and to what degree they may go backwards and change their analysis

decisions. An important principle of LSU is that information is given *when* it is needed, hence it is not about blinding to information, but rather at optimising the sequencing.

In addition to blinding and controlling when information is given, another approach is to *compartmentalise* the work, so as to avoid bias cascade and cross cognitive contamination (Dror, 2014). In digital forensics, this could be done by separating the task, so that the analysis is assigned to another DFP than the one collecting the evidence at the crime scene. This way, the DFP that was exposed to a lot of irrelevant contextual information from the crime scene is not the DFP who conducts the subsequent analysis back at the forensic laboratory.

As bias can never be totally eliminated, procedures to *uncover* cognitive or human error will also be necessary, so as to be transparent. There are few formalized quality assurance procedures within digital forensics, such as verification or peer review, but these have to be done correctly from a cognitive point of view, e.g., that verification is done blindly. Peer review is mentioned in some guidelines and standards (e.g. ENFSI, 2015; SWGDE, 2018), but the detail in how the measure should be carried out varies from just mentioning the measure with no elaboration (e.g. SWGDE, 2018), to a more detailed description on how peer review should be conducted and documented (ENFSI, 2015). The efficiency of the peer-review procedure depends partly on the ability to reduce bias with the assessor conducting the peer review.

Peer review, or case review, is typically different from verification. Verification may be done by totally repeating the analysis by a second independent examiner to see whether they arrive at the same conclusions (Ballantyne et al., 2017). Verification hinges and depends on repeatability, which is an issue across forensic domains, including fingerprinting (Dror et al., 2011; Ulery et al., 2012) and DNA (Barrio et al., 2018; Butler et al., 2018; Dror and Hampikian, 2011). According to Page et al. (2018), proper verification does not seem to be frequently used within the digital forensic discipline. To help combat bias, the verification of digital forensics decisions should be a controlled process in which blind and double-blind procedures are used whenever possible. Such procedures would require that the verifier is not informed of the initial conclusion; if possible, that the verifier does not know who the examiner was; and that the examiner does not select the verifier (Kassin et al., 2013). To maximize independence in the verification process it is also advised, if possible and warranted, to do cross-laboratory verifications of the forensic work (Koppl et al., 2008). To avoid the base rate problem, verification should not only be undertaken on positive results, but also include verification of negative results (Dror, 2014).

Bias mitigating measures are sometimes applicable, but other times they are neither practical, possible nor advisable. Regardless of the measures that are taken to minimize bias, it is important that the DFP is *transparent* about their work. They need to provide detail in the documentation and presentation in court about what context they knew, who gave it to them, when, and why. This is particularly important regarding irrelevant contextual information they were exposed to. Hence, if bias cannot be minimized, then at least there needs to be transparency to its possible presence, which in turn will make it possible for others to assess the reliability of the outcome, and to develop new ways to minimize bias.

Conclusion

Within forensic science, there has been increased attention on cognitive and human factor as a source of error, and quite an extensive research foundation from many forensic disciplines, which have shown that the forensic expert is susceptible to bias when making decisions. In this paper we have addressed the research question: When handling digital evidence through the digital forensics process; when is the DFP vulnerable to cognitive bias, and what measures could be relevant and effective to mitigate bias for this specific domain?

There is a lack of research on cognitive bias within the digital forensics domain. The analysis of the digital forensics work within the seven level taxonomy of the sources of bias showed that the cognitive processes and judgement of the DFP are susceptible to many sources of bias during the digital forensics process. Given that the digital forensics domain is operating under less rigorously defined standards, practitioner governance and validation procedures compared to other forensic disciplines (Page et al., 2018), there should be a concern in relation to the risk of human error. Research on miscarriages of justice and misleading evidence has shown that human error within forensic science is not just a theoretical issue, but a real and serious problem in the justice system.

There are several research based approaches aimed at mitigating cognitive and human error in forensic examinations such as a strict scientific approach, context management, exposure control and compartmentalization of work, and some variations of these will be applicable to digital forensics. However, there are factors that are unique to the nature of digital evidence and the work of the DFP which underpins the need for research on which situations or activities are most vulnerable in relation to cognitive bias, and which are the most effective countermeasures, specific to this domain. The Hierarchy of Expert Performance framework (Dror, 2016) could be used for unpacking and measuring forensic expert performance in relation to susceptibility to bias and reliability issues through empirical research. Further research on cognitive and human error within digital forensics would contribute to a more scientifically sound discipline.

Declaration of interests

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.

Acknowledgement

We want to thank Helene O. I. Gundhus, Johanne Y. Dahl, Fergus Toolan, and the anonymous reviewers for their valuable comments on an earlier draft of this paper. This paper was funded by a grant from the Norwegian Police University College awarded to the first author for pursing a PhD.

References

Ask, K., 2013. Bias: Fejl og faldgruber i efterforskning. In: Hald, C., Vrist Rønn, K. (Eds.), *Om at opdage: Metodiske refleksioner over Politiets Undersøgelsespraksis*. Fredriksberg: Samfundslitteratur.

Ask, K., Granhag, P.A., 2005. Motivational sources of confirmation bias in criminal investigations: the need for cognitive closure. J. Investigative Psychol. Offender Profiling 2 (1), 43–63.

ASTM, 2015. Standard Terminology for Digital and Multimedia Evidence Examination. ASTM E2916 – 13.

Balcetis, E., Dunning, D., 2006. See what you want to see: motivational influences on visual perception. J. Personal. Soc. Psychol. 91 (4), 612–625.

Ballantyne, K.N., Edmond, G., Found, B., 2017. Peer review in forensic science. Forensic Sci. Int. 277, 66–76.

Barrio, P., Crespillo, M., Luque, J., Aler, M., Baeza-Richer, C., et al., 2018. GHEP-ISFG collaborative exercise on mixture profiles (GHEP-MIX06). Reporting conclusions: results and evaluation. Forensic Sci. Int.: Genetics 35, 156–163.

Bieber, P., 2012. Measuring the impact of cognitive bias in fire investigation. In: Paper Presented at the Proceedings of the International Symposium on Fire Investigation. Science and Technology.

Butler, J.M., Kline, M.C., Coble, M.D., 2018. NIST interlaboratory studies involving DNA mixtures (MIX05 and MIX13): variation observed and lessons learned. Forensic Sci. Int.: Genetics 37, 81–94.

Casey, E., 2002. Error, uncertainty, and loss in digital evidence. International Journal

of Digital Evidence 1 (2), 1–45.
Casey, E., 2011. Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet. Elsevier, Amsterdam.
Ciardhuáin, S.Ó., 2004. An extended model of cybercrime investigations. International Journal of Digital Evidence 3 (1), 1–22.
Cook, R., Evett, I.W., Jackson, G., Jones, P., Lambert, J., 1998. A hierarchy of propositions: deciding which level to address in casework. Sci. Justice 38 (4), 231–239.
Dioso-Villa, R., 2012. Without legal obligation: compensating the wrongfully convicted in Australia. Albany Law Rev. 75 (3), 1329–1372.
Dioso-Villa, R., 2015. A repository of wrongful convictions in Australia: first steps toward estimating prevalence and causal contributing factors. Flinders Law Journal 17, 163–202.
Dror, I.E., 2011. The paradox of human expertise: why experts get it wrong. In: Kapur, N. (Ed.), The Paradoxical Brain. Cambridge University Press, Cambridge, pp. 177–188.
Dror, I.E., 2012. Letter to the Editor—combating bias: the next step in fighting cognitive and psychological contamination. J. Forensic Sci. 57 (1), 276–277.
Dror, I.E., 2014. Practical solutions to cognitive and human factor challenges in forensic science. Forensic Sci. Policy Manag.: Int. J. 4 (3–4), 105–113.
Dror, I.E., 2016. A hierarchy of expert performance. Journal of Applied Research in Memory and Cognition 5 (2), 121–127.
Dror, I.E., 2017. Human expert performance in forensic decision making: seven different sources of bias. Aust. J. Forensic Sci. 49 (5), 541–547.
Dror, I.E., 2018. Biases in forensic experts. Science 360 (6386), 243. https://doi.org/10.1126/science.aat8443.
Dror, I.E., Champod, C., Langenburg, G., Charlton, D., Hunt, H., Rosenthal, R., 2011. Cognitive issues in fingerprint analysis: inter-and intra-expert consistency and the effect of a 'target'comparison. Forensic Sci. Int. 208 (1–3), 10–17.
Dror, I.E., Charlton, D., 2006. Why experts make errors. J. Forensic Identif. 56 (4), 600–616.
Dror, I.E., Charlton, D., Péron, A.E., 2006. Contextual information renders experts vulnerable to making erroneous identifications. Forensic Sci. Int. 156 (1), 74–78.
Dror, I.E., Hampikian, G., 2011. Subjectivity and bias in forensic DNA mixture interpretation. Sci. Justice 51 (4), 204–208.
Dror, I.E., Rosenthal, R., 2008. Meta-analytically quantifying the reliability and biasability of forensic experts. J. Forensic Sci. 53 (4), 900–903.
Dror, I.E., Thompson, W.C., Meissner, C.A., Kornfield, I., Krane, D., Saks, M., Risinger, M., 2015. Letter to the editor - context management toolbox: a linear sequential unmasking (LSU) approach for minimizing cognitive bias in forensic decision making. J. Forensic Sci. 60 (4).
Ekfeldt, J., 2016. Om Informationstekniskt Bevis. Doctoral Dissertation. Juridiska Institutionen, Stockholms Universitet), Ekfeldt, J., Stockholm.
ENFSI, 2015. Best Practice Manual for the Forensic Examination of Digital Technology, ENFSI-BPM-FOT-01. Version 01 (November 2015).
Flaglien, A.O., 2018. The digital forensics process. In: Årnes, A. (Ed.), Digital Forensics. Wiley, Hoboken.
Garfinkel, S.L., 2010. Digital forensics research: the next 10 years. Digit. Invest. 7, 64–73.
Garrett, B.L., 2011. Convicting the Innocent. Harvard University Press.
Garrett, B.L., Neufeld, P.J., 2009. Invalid Forensic Science Testimony and Wrongful Convictions. Virginia Law Review, pp. 1–97.
Hansen, H.A., Andersen, S., Axelsson, S., Hopland, S., 2017. Case study: a new method for investigating crimes against children. In: Paper Presented at the Proceedings of the Conference on Digital Forensics, Security and Law, Maidens.
Hunton, P., 2011. A rigorous approach to formalising the technical investigation stages of cybercrime and criminality within a UK law enforcement environment. Digit. Invest. 7 (3–4), 105–113.
Jeanguenat, A.M., Bruce Budowle, B., Dror, I.E., 2017. Strengthening forensic DNA decision making through a better understanding of the influence of cognitive bias. Sci. Justice 57 (6), 415–420.
Kahneman, D., 2011. Thinking, Fast and Slow. Macmillan.
Kassin, S.M., Dror, I.E., Kukucka, J., 2013. The forensic confirmation bias: problems, perspectives, and proposed solutions. Journal of Applied Research in Memory and Cognition 2 (1), 42–52.
Koppl, R.G., Kurzban, R., Kobilinsky, L., 2008. Epistemics for forensics. Episteme 5 (2), 141–159.
Kukucka, J., Kassin, S.M., Zapf, P.A., Dror, I.E., 2017. Cognitive bias and blindness: a global survey of forensic science examiners. Journal of Applied Research in Memory and Cognition 6 (4), 452–459.
Leppänen, A., Kankaanranta, T., 2017. Cybercrime investigation in Finland. J. Scand. Stud. Criminol. Crime Prev. 18 (2), 157–175.
Lindsay, P.H., Norman, D.A., 2013. Human Information Processing: an Introduction to Psychology. Academic press.
McKemmish, R., 2008. When is digital evidence forensically sound?. In: IFIP International Conference on Digital Forensics. Springer, Boston, MA, pp. 3–15.
Murrie, D.C., Boccaccini, M.T., Turner, D.B., Meeks, M., Woods, C., Tussey, C., 2009. Rater (dis) agreement on risk assessment measures in sexually violent predator proceedings: evidence of adversarial allegiance in forensic evaluation? Psychol. Publ. Pol. Law 15 (1), 19–53.
Nakhaeizadeh, S., Dror, I.E., Morgan, R.M., 2014. Cognitive bias in forensic anthropology: visual assessment of skeletal remains is susceptible to confirmation bias. Sci. Justice 54 (3), 208–214.
Nakhaeizadeh, S., Morgan, R.M., Rando, C., Dror, I.E., 2018. Cascading bias of initial exposure to information at the crime scene to the subsequent evaluation of skeletal remains. J. Forensic Sci. 63 (2), 403–411.
National Registry of Exonerations, 2015. The First 1600 Exonerations downloaded. https://www.law.umich.edu/special/exoneration/Documents/1600_Exonerations.pdf. (Accessed 20 March 2019).
Neal, T.M., 2016. Are forensic experts already biased before adversarial legal parties hire them? PLoS One 11 (4) e0154434. https://doi.org/10.1371/journal.pone.0154434.
Nickerson, R.S., 1998. Confirmation bias: a ubiquitous phenomenon in many guises. Rev. Gen. Psychol. 2 (2), 175–220.
Oliver, W.R., 2017. Effect of history and context on forensic pathologist interpretation of photographs of patterned injury of the skin. J. Forensic Sci. 62 (6), 1500–1505.
Page, H., Horsman, G., Sarna, A., Foster, J., 2018. A review of quality procedures in the UK forensic sciences: what can the field of digital forensics learn? Sci. Justice 59 (1), 83–92.
Pohl, R.F., 2016. Cognitive Illusions: Intriguing Phenomena in Judgement, Thinking and Memory. Psychology Press.
Pollitt, M., Casey, E., Jaquet-Chiffelle, D., Gladyshev, P., 2018. A Framework for Harmonizing Forensic Science Practices and Digital/Multimedia Evidence. The Organization of Scientific Area Committees for Forensic Science (OSAC), USA.
Pronin, E., Lin, D.Y., Ross, L., 2002. The bias blind spot: perceptions of bias in self versus others. Pers. Soc. Psychol. Bull. 28 (3), 369–381.
Rassin, E., 2018. Reducing tunnel vision with a pen-and-paper tool for the weighting of criminal evidence. J. Investigative Psychol. Offender Profiling 15 (2), 227–233.
Smalarz, L., Madon, S., Yang, Y., Guyll, M., Buck, S., 2016. The perfect match: do criminal stereotypes bias forensic evidence analysis? Law Hum. Behav. 40 (4), 420–429. https://doi.org/10.1037/lhb0000190.
Smit, N.M., Morgan, R.M., Lagnado, D.A., 2018. A systematic analysis of misleading evidence in unsafe rulings in England and Wales. Sci. Justice 58 (2), 128–137.
Steblay, N., Dysart, J., Fulero, S., Lindsay, R., 2003. Eyewitness accuracy rates in police showup and lineup presentations: a meta-analytic comparison. Law Hum. Behav. 27 (5), 523–540.
Stevenage, S.V., Bennett, A., 2017. A biased opinion: demonstration of cognitive bias on a fingerprint matching task through knowledge of DNA test results. Forensic Sci. Int. 276, 93–106.
Stoel, R.D., Berger, C., Kerkhoff, W., Mattijssen, E., Dror, I.E., Hickman, M., Strom, K., 2014. Minimizing contextual bias in forensic casework. Forensic Science and the Administration of Justice: Critical Issues and Directions 67, 67–86.
Sun, J.-R., Shih, M.-L., Hwang, M.-S., 2015. A survey of digital evidences forensic and cybercrime investigation procedure. Int. J. Netw. Secur. 17 (5), 497–509.
Sunde, N., 2017. Non-technical Sources of Errors when Handling Digital Evidence within a Criminal Investigation. Master's Thesis. Faculty of Technology and Electrical Engineering, Norwegian University of Science and Technology, Sunde, N. Gjøvik).
SWGDE, 2018. Establishing Confidence in Digital Forensic Results by Error Mitigating Analysis. Version: 2.0 (20. Nov, 2018).
Taylor, M.C., Laber, T.L., Kish, P.E., Owens, G., Osborne, N.K., 2016. The reliability of pattern classification in bloodstain pattern analysis, Part 1: bloodstain patterns on rigid non-absorbent surfaces. J. Forensic Sci. 61 (4), 922–927.
Tversky, A., Kahneman, D., 1973. Availability: a heuristic for judging frequency and probability. Cogn. Psychol. 5 (2), 207–232.
Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. Science 185 (4157), 1124–1131.
Ulery, B.T., Hicklin, R.A., Buscaglia, J., Roberts, M.A., 2012. Repeatability and reproducibility of decisions by latent fingerprint examiners. PLoS One 7 (3), 1–12.
Van Buskirk, E., Liu, V.T., 2006. Digital evidence: challenging the presumption of reliability. J. Digit. Forensic Pract. 1 (1), 19–26.
van den Eeden, C.A., de Poot, C.J., Van Koppen, P.J., 2016. Forensic expectations: investigating a crime scene with prior information. Sci. Justice 56 (6), 475–481.
Wells, G.L., Small, M., Penrod, S., Malpass, R.S., Fulero, S.M., Brimacombe, C., 1998. Eyewitness identification procedures: recommendations for lineups and photospreads. Law Hum. Behav. 22 (6), 603–647. https://doi.org/10.1023/A:1025750605807.
Zapf, P.A., Dror, I.E., 2017. Understanding and mitigating bias in forensic evaluation: lessons from forensic science. Int. J. Forensic Ment. Health 16 (3), 227–238.