Journal of Universal Computer Science, vol. 28, no. 8 (2022), 827-853
submitted: 21/4/2022, accepted: 28/7/2022, appeared: 28/8/2022 CC BY-ND 4.0

Exploratory study of a system to reduce information overload and tunnel vision in homicide investigations

Marloes Vredenborg
(Utrecht University, Utrecht, Netherlands
https://orcid.org/0000-0001-5724-6851, m.t.r.vredenborg@uu.nl)

Daan Sutmuller
(Dutch Police, Netherlands
https://orcid.org/0000-0002-3489-2866, daan.sutmuller@politie.nl)

Mariëlle den Hengst-Bruggeling
(Dutch Police, Netherlands
Netherlands Institute for the Study of Crime and Law Enforcement, Amsterdam, Netherlands
marielle.den.hengst@politie.nl; mdenhengst@nscr.nl)

Judith Masthoff
(Utrecht University, Utrecht, Netherlands
https://orcid.org/0000-0001-8099-0515, j.f.m.masthoff@uu.nl)

Abstract: In homicide investigations, the growing availability of data results in an increasing amount of information and Persons of Interest (PoIs) that can be collected and incorporated during an investigation. This might result in information overload and increased tunnel vision during a homicide investigation. In this paper, we designed a system to support homicide investigations in such a way that it reduces information overload and tunnel vision. For evaluation purposes, we built a prototype that was filled with a fictional homicide investigation. A user study indicated that criminal investigators experienced a significantly low level of information overload and tunnel vision using the prototype. Moreover, the results showed acceptable usability and verbal statements indicated a largely positive attitude towards the prototype. This research clearly shows the opportunity to use interface design artefacts to support the prevention of information overload and tunnel vision.

Keywords: Tunnel Vision, Information Overload, Experiment, Homicide Investigation, Police
Categories: H.5
DOI: 10.3897/jucs.85648

1 Introduction

We have witnessed an exponential growth in the amount of digital data over the past years and this growth will continue for years to come [Holst, 2020]. In homicide investigations, the growing availability of data results in an increasing amount of information and Persons of Interest (PoIs) that can be collected and incorporated during an investigation. Eventually, this may support the incorporation of the perpetrator into the investigation and resolution of the case [Sutmuller, den Hengst, Baros, and van Gelder, 2020].

However, as the amount of information in an investigation increases, the navigation and understanding of information that may prove crucial becomes challenging [Zamanirad, Benatallah, Barukh, Rodriguez, and Nouri, 2020]. The increasing amount of information, in addition to the complex nature of homicide investigations, leads to investigators experiencing a state of information overload [Brookman, Maguire, and Maguire, 2019]. This is problematic as it has been shown that information overload can result in poorer decision making, reduced productivity, difficulties in sense-making and reasoning, poorer memory recall, or experiences of frustration, tiredness, stress, confusion and/or anxiety [Chewning and Harrell, 1990; Schick, Gordon, and Haka, 1990; Farhoomand and Drury, 2002; Benselin and Ragsdell, 2016; Jones and Kelly, 2018]. Moreover, the combination of these factors makes homicide investigators susceptible to cognitive biases, which may lead to investigators encountering the well-known tunnel vision phenomenon [Tversky and Kahneman, 1974; Gruijter and Poot, 2019; Cao, 2008; Snook and Cullen, 2008; Goette, Han and Leung, 2019]. Tunnel vision could be problematic as it is connected to several investigative failures [Bronkhorst, 2014; Findley and Scott, 2006].

In the current practice of homicide investigations various methodologies for managing information and guiding investigators to an evidence-based focus exist [e.g., Heuer, 1999; ACPO, 2006; Hanshew, 2010; Sutmuller et al., 2020]. Little is known about the extent to which these methodologies take homicide investigators' information overload and tunnel vision into account. Furthermore, little is known about software-based tools to support these methodologies. Although technology usage could contribute to information overload, it has also been suggested that the use of technology, and a design that minimizes users' cognitive load, could free up information processing capacity [Eppler and Mengis, 2004; Oviatt, 2006]. Therefore, this research studies the impact of a system on information overload and tunnel vision of criminal investigators.

The remainder of this paper is structured as follows. Section 2 provides a background on homicide investigations and the key concepts of this paper, information overload and tunnel vision. Section 3 describes the system designed to support homicide investigations. The method to evaluate the designed system on the effect on information overload and tunnel vision is described in Section 4. Section 5 presents the results of this method. Section 6 provides a discussion of the results. Conclusions of the research are described in Section 7.

2 Background

2.1 Homicide investigations

Complex cases, such as homicide investigations, are characterized by substantial amounts of information in various forms and collected from various sources, such as physical evidence (e.g., fingerprints or DNA), digital evidence (e.g., phone records, camera footage or data from seized devices), or witness statements [Epskamp-Dudink, 2016; van Wijk, van Leiden and Hardeman, 2017; Liedenbaum, Poot, Vergouw and Kouwenberg, 2015]. It is essential that all available information is being collected and stored as quickly as possible. If not, the possibility exists that valuable information gets lost over time (e.g., physical evidence such as fingerprints can vanish or the passage of time can alter the memories of witnesses) [van Wijk et al., 2017; Salet, 2015]. Moreover, the collected information is often fragmented, unstructured, ambiguous, encoded in a variety of formats, and a considerable amount turns out to be useless throughout the investigation [van Wijk et al., 2017; Zamanirad et al., 2020; Innes, 2003; Rossmo, 2016]. Therefore, it is important that evidence is collected independently and approached broadly, without making any assumption, such that it can be interpreted objectively [Baber, 2006; Epskamp-Dudink, 2016].

Investigators in a homicide investigation are engaging in a series of complex cognitive tasks. To illustrate, investigators are collecting a wide range of information of which they are trying to make sense [Baber, 2006; Innes, 2003], they reason about the information through deductive, inductive, and abductive reasoning such that insights and connections can be created [Rossmo, 2016; Fahsing and Ask, 2017], and they must make decisions about which actions to take next [Ask and Granhag, 2005; Innes, 2003; Baber, 2006]. These are complex in a way that they (1) require high-order cognitive processes and (2) are executed under complex conditions [Stelfox, 2011; Hawk and Dabney, 2014]. I.e., criminal investigators work under substantial time and societal pressure, their decision can have major consequences, they work in a police culture that promotes decisiveness and the cognitive activities are performed by interaction with more than one agent [Ask and Granhag, 2005; Baber, 2006; Knauff and Wolf, 2010; Innes, 2003]. Victims, media, society and police chiefs demanding case clearance and more cases waiting to be investigated increase the pressure on investigators.

2.2 Methodology to support homicide investigations

Generally, the management of a homicide investigation is separated into three distinct phases [ACPO, 2006; Sutmuller et al., 2020]. The first phase is known as the collection phase. This includes the collection of PoIs and pieces of evidence. The second phase is known as the prioritization phase. This includes the management of data, and the creation of information and knowledge through analysing the data with the purpose to prioritize the PoIs. Finally, the elimination phase considers creating a case file that proves the guilt of the person charged and eliminating all alternatives.

To support criminal investigators in the process of homicide investigations, several methodologies exist, such as Case-Specific Element Library (C-SEL) [Sutmuller et al., 2020], Person of Interest Priority Assessment Tool (POIPAT) [Wilson, 2012], Trace Investigate Evaluate (TIE) [ACPO, 2006], Rasterfahndung [Hoppmann, 2013], and Analysis of Competing Hypotheses (ACH) [Heuer, 1999]. Prior research [Sutmuller, 2021] has shown that C-SEL has the most (partial) attributes that aim to prevent the occurrence of tunnel vision in criminal investigation, followed closely by ACH. Therefore, we use C-SEL in this research.

[Sutmuller et al., 2020] developed a Case-Specific Element Library (C-SEL) to incorporate and prioritize PoIs in homicide investigations. This methodology consists of 24 elements concerning the opportunity, motive and means of PoIs. These elements can be used to accommodate pieces of evidence that are available in a particular investigation. The focus on elements supports the investigators in focussing on what happened instead of on persons who might have done it. PoIs are incorporated into the investigation based on group level elements and further prioritized using elements on an individual level. Including multiple PoIs prevents tunnel vision. For all elements, an initial score was obtained using expert judgement. This initial score can be adjusted with underlying factors based on the relevance and credibility of evidence in that specific case. The score on all elements is displayed in a PoI by element matrix showing the ranking of PoIs.

2.3 Information overload

[Galbraith, 1974] defined information overload (IO) as a state in which the information-processing requirements of a task exceed the information-processing capacity available to an individual. [Eppler and Mengis, 2004] stated that the term IO is often used to describe a situation in which individuals are 'receiving too much information', or in terms of [Bawden and Robinson, 2009]: information becomes an obstacle, even though the information might be relevant and useful to the task at hand. Regardless of how the term is defined, it may be tempting and logical to assume that IO is simply caused by too much information. Yet, many studies proved that IO is not only caused by too much information. Instead, other important contributors to IO should be taken into consideration. [Eppler and Mengis, 2004] suggested a conceptual framework in which five factors influence the two fundamental variables of IO: the information processing capacity and the information processing requirements. These five factors include: information characteristics (e.g., quantity, quality, ambiguity), person receiving, processing, or communicating the information (e.g., information processing capacity, motivation, personal traits), tasks or processes that need to be completed (e.g., complexity, time pressure, interdisciplinarity), organizational design (e.g., collaborative work, group heterogeneity) and information technology that is used (e.g., push vs. pull systems).

Countermeasures of information overload. Various strategies for managing IO exist. The most widely adopted approach for managing IO is known as information filtering [Hanani, Shapira, and Shoval, 2001; Savolainen, 2007]. This is concerned with weeding out information that is presumed to be irrelevant at the time and supplying relevant information to the user's attention [Savolainen, 2007]. Information withdrawal is a more personal and affectively oriented coping strategy. It defines the situation in which the number of information sources is being kept to a minimum, such that excessive information supply is not possible [Savolainen, 2007]. For instance, turning notifications off on your phone for a certain period of time. Another strategy used is chunking, where information is considered in smaller chunks at intervals, minimizing the demands on working memory [Benyon, 2014]. Moreover, queuing is a strategy in which some information is being delayed to a less busy time [Miller, 1992 as cited in Bawden and Robinson (2020)]. Visualization of information has also been suggested to manage IO [Meyer, 1998]. Visualization allows the user to examine large amounts of information. It supports the user in keeping an overview of the whole while pursuing details and keeping track of various things (by using the display as external working memory).

Several methods applied in interface design exist that aim to free up mental resources. First, externalizing aims to reduce memory load by incorporating information on the interface such that problem-solving can be structured and guided. An example of externalization of information is the use of greyed-out menu items that allows us to only execute actions that are possible at that moment [van Nimwegen, van Oostendorp, Burgos and Koper, 2006]. Second, recognition rather than recall is an often-mentioned method to minimize the user's cognitive load [Preece, Sharp, and Rogers, 2015]. Therefore, users should not have to remember information from one part of the system to another, and instructions should be visible or easily retrievable whenever appropriate [Nielsen, 2005]. Closely related to this is the usage of gestalt principles. Gestalt principles describe how humans perceive elements based on the precognitive determination of what those elements mean or do based on size, shape, position, and other elements [Evans, 2017]. It includes five main principles of perception [Koffka, 2013]: law of proximity (objects appearing close to each other tend to be perceived together), law of similarity (objects with similar features are considered to belong to the same class), law of closure (objects that are positioned close to each other are perceived as a whole), law of continuity (objects that include disconnected elements are often perceived as a continuous whole) and law of symmetry (objects that are symmetrical to each other are perceived to be one group). Applying these laws to interface design could reduce complexity, and hence cognitive load [Knight, 2019].

2.4 Tunnel vision

It has been argued that our reasoning and decision-making capabilities are constrained due to the limitation of our human cognition [Kahneman and Tversky, 1973; Baddeley, 1992]. Therefore, to make reasonable decisions, with minimal cognitive effort, human cognition uses so-called cognitive heuristics, simple mental strategies used by individuals to deal with complex tasks and uncertain situations [Tversky and Kahneman, 1974; Lau and Redlawsk, 2001]. For example, heuristics allow us to focus on information that is relevant and ignore irrelevant information when confronted with IO [Gruijter and Poot, 2019]. Cognitive heuristics can be very efficient and effective, but they might lead to severe and systemic errors, known as cognitive biases [Tversky and Kahneman, 1974].

In criminal investigations, tunnel vision is often linked to the use of these heuristics and the occurrence of cognitive biases. [Findley and Scott, 2006] describe tunnel vision in criminal investigations as: "That "Compendium of common heuristics and logical fallacies" to which we are all susceptible, that lead actors in the criminal justice system to "focus on a suspect, select and filter the evidence that will 'build a case' for conviction, while ignoring or suppressing evidence that points away from guilt."".

Thus, tunnel vision is actually a quite common and logical cognitive strategy to deal with the complex factors that characterise a criminal investigation [Snook and Cullen, 2008]. It can be very useful to create a tunnelled vision at some points in investigations. It helps investigators to have some focus and understand which scenarios are most plausible [Liedenbaum et al., 2015; Groenendaal and Helsloot, 2014; Snook and Cullen, 2008]. Yet, it becomes problematic when investigators are no longer capable of stepping out of their tunnelled vision. Then they (unwittingly) have a rigid focus on one single hypothesis and select and filter information such that it will support their hypothesis, while overlooking or suppressing contradicting information [Salet, 2015; Findley and Scott, 2006; Liedenbaum et al., 2015]. In doing so, they are consistently working towards a certain suspect and do not consider or eliminate suspects that should be investigated [Findley and Scott, 2006; Liedenbaum et al., 2015).

In literature on criminal investigations, the primary cognitive bias to which tunnel vision is connected is a type of selective thinking: confirmation bias [Helsloot and Groenendaal, 2012; Findley and Scott, 2006; Salet, 2015; Rossmo, 2016]. This is a well-known human tendency to search for information that is consistent with our beliefs, expectations, or hypotheses, rather than searching for refuting information [Ask and Granhag, 2005]. It is not only limited to the search for confirmatory information. Instead, individuals also tend to recall confirmatory information better and allocate more weight to confirmatory evidence than to disconfirmatory evidence. So confirmation bias not only operates through selective information search, but also through biased information interpretation [Ask and Granhag, 2005].

Countermeasures of tunnel vision: Preventing tunnel vision is a topic researched exhaustively in the context of criminal investigations. Besides a number of policy-related attributes, like the development of the skills of actors within the system [Findley and Scott, 2006; Groenendaal and Helsloot, 2015; Posthumus, 2005] and the recalibration of rules and procedures [Findley and Scott, 2006], seven attributes were found that aim to prevent the occurrence of the confirmation bias.

Three attributes were found that aim to prevent selective information search. Firstly, the focus of the investigation is important. Suspect-driven investigations, pursuing one suspect while ignoring other potential candidates can facilitate tunnel vision [van Koppen, Hessing, Merkelbach, and Crombach, 2002]. To overcome this problem, one should focus on what happened instead of on a single suspect [Epskamp-Dudink and Winter, 2020]. Proper reconstruction based on the victim and crime scene prevents mistakes in decision making during criminal investigation and helps to recognize confirmation bias at an early stage [Epskamp-Dudink and Winter, 2020]. Secondly, the use of alternative hypotheses is important to overcome selective information search [Posthumus, 2005]. Decision-makers should try to keep in mind for as long as possible that the suspect may be innocent after all and that the incident took place in an alternative manner [Rassin, 2018]. Thirdly, the creation of transparency by recording all investigated leads and uninvestigated leads and areas [Findley and Scott, 2006; Posthumus, 2005] can assist with the prevention of selective information search.

Four attributes were found that aim to assist decision-makers with the prevention of biased interpretation of information. The creation of an overview using matrices [Hallihan, Cheong, and Shu, 2012] and forcing to consider all evidence for multiple hypotheses seem to have a mitigating effect on confirmation bias [Rassin, 2018]. [Posthumus, 2005] recommends using critical review to create an optimal environment in which external criminal investigation experts critically review the process of investigation. Critical review focuses on questioning the choices made, posing what-if questions [Salet and Terpstra, 2014]. The use of critical review stimulates a devil's advocate approach and prevents biased interpretation of information. One difficulty with the comparison of PoIs based on the rating of evidence is the absence of a prior probability of guilt. Some PoIs could be more interesting than others at the start of an investigation. Updating the a-priori probability of a hypothesis by the likelihood ratio of a piece of evidence results in an updated posterior probability of guilt.

3 Methodology and interface to reduce IO and tunnel vision

3.1 Interface

Based on the comparison of methodologies [Sutmuller, 2021] we have selected C-SEL as the methodology to build a supporting interface. Several human-centred steps were taken to inform the interface design. First, existing documentation and interviews with the primary users were used to specify the context of use and requirements. Afterwards, low-fidelity design solutions were created and evaluated with experts using a brainstorm session and cognitive walkthroughs. Eventually, a high-fidelity prototype was created. Several pages of C-SEL were designed with which users can interact. The language in the prototype was Dutch, as this was the main language of the primary users. The design of specific pages will be described shortly. Not all pages are shown due to limited space but are available on request. All information shown in the pages is fictive and does not correspond to a real-life situation.

Home Page. The high-fidelity prototype starts with a home page, where users can start a new case or open an existing case. After creating or selecting a case, users are directed to this case ('Zaak') and presented with the matrix overview.

Matrix Page. This presents an overview of all available PoIs, their scores for each element and their total score in the given case (Fig. 1). Elements are represented using icons. Several interactions are possible in the matrix: (1) hovering over an icon provides a description of that element; (2) clicking on a PoI's name, or on their total score provides more information about that PoI (Fig. 2); (3) clicking on a specific score shows all evidence registered for that PoI for that specific element. When evidence is not yet evaluated, users receive a warning and are allowed to evaluate this evidence immediately (Fig. 3). In the matrix, emphasis is placed on unevaluated evidence: a red colour indicates that none of the evidence available for that element has been evaluated, whereas orange indicates that multiple pieces of evidence are available for an element, yet these are only partially evaluated.

Evidence Overview Page. On the Evidence page ('Tactische aanwijzingen'), an overview of all evidence available in the current case is shown (Fig. 4). Clicking on an evidence item shows all information available about that evidence. It is also possible to select evidence, or click on the three menu dots, to edit or delete evidence. Additionally, the left side of the page provides various filter options. Users can add new evidence by selecting the button on the top right ('+ Toevoegen').

Adding Evidence Pages. During the cognitive walkthrough, it became apparent that participants prefer a road map over a progress bar. Therefore, a road map is provided on the left, allowing users to keep track of their progress when adding evidence. On Page 1 of adding evidence, users have to record the date ('datum') of securing the evidence, type of evidence ('soort'), element level, element type and a description of the evidence ('omschrijving'). The element types shown are based on the selected element level. E.g., if users select 'individual' they will only be presented with element types on an individual level. Moreover, if users select tangible, elements that are not tangible will be greyed-out. On Page 2, they have to rate underlying factors. Based on the cognitive walkthrough results, radio buttons were used instead of dropdown lists (Fig. 5). Users are asked to add a substantiation to their ratings. Yet,

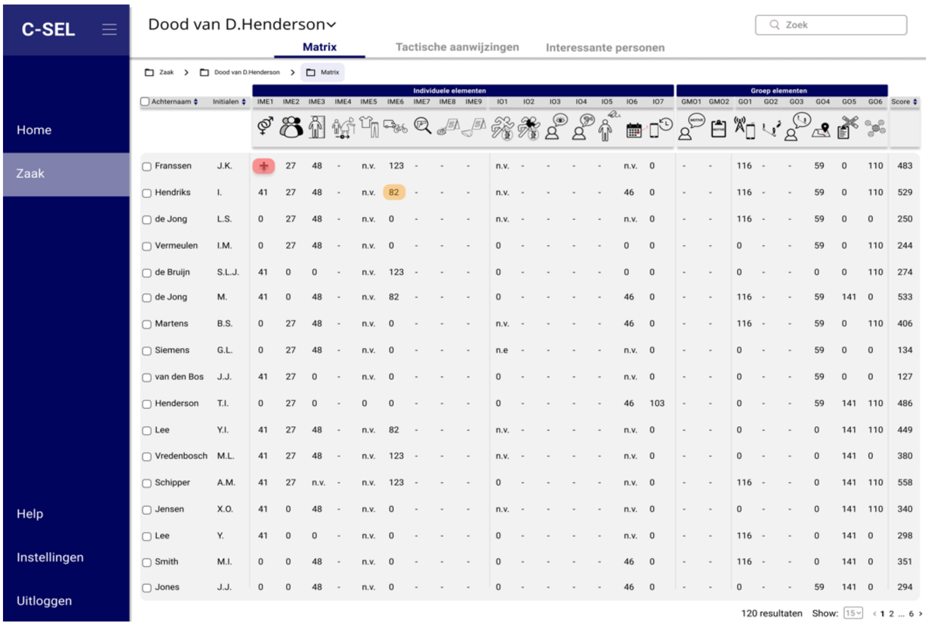

Figure 1: Matrix page C-SEL

Figure 2: Interaction when clicking on a person of interest in C-SEL

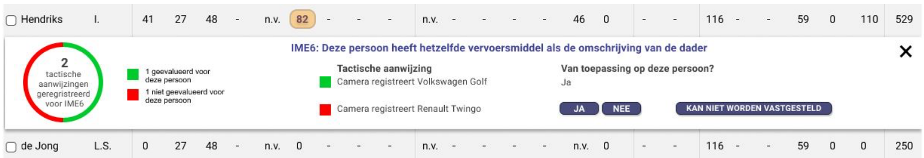

Figure 3: Interaction when clicking on a score in C-SEL

this is optional as there are cases in which it is redundant and it requires additional time which is not always available. On Page 3, they are asked if new PoIs are connected to the evidence. If yes, they can add new PoIs by adding their name, initials and ID-number. On Page 4, they are provided with a summary of the new evidence. This allows them to see if any mistakes were made, as well as the score calculated for the evidence. After confirmation, the process of adding evidence is finished.

PoIs Page. The PoI page ('Interessante personen') shows an overview of all PoIs in the current case. When clicking on a PoI, users receive all information available about that PoI, similar to when clicking on a PoI on the Matrix page.

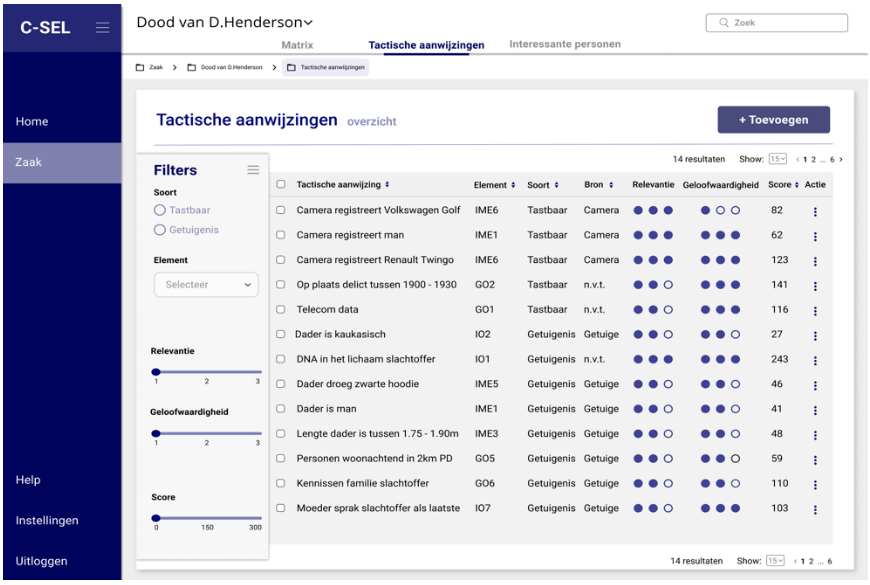

Figure 4: Evidence overview page C-SEL

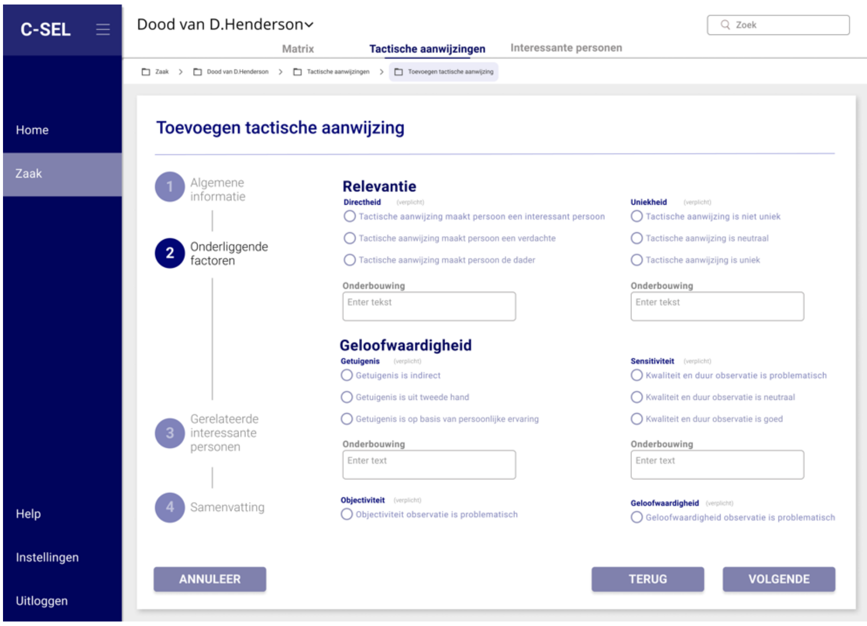

Figure 5: Add new evidence Page 2

3.2 Design decisions regarding information overload

Based on insights from the literature regarding IO, several design decisions were made: (1) Information was chunked in the interface design to minimize the demands of users' working memory in various ways. First, elements were re-arranged and grouped to make them more cohesive. Secondly, only scores were shown in the matrix at first. If users clicked on a score, the menu would expand with a summary of information. Then, when they desire to see all results, it is necessary to click on details. This ensures that details are only shown to the user when necessary. Thirdly, when adding evidence, underlying factors were grouped to help users understand which underlying factors were related. (2) Icons were created from scratch using a brainstorming session, to decrease the demands on users' cognitive load. The icons allowed users to quickly recognize what a column exemplifies. Moreover, there was a hovering option for the icons to support the users in recall. (3) Greyed-out menu-items were used where possible to allow users to only execute actions possible at that time, aiming to reduce users' memory load. (4) Universal designs were used for sorting, filtering, menu options and buttons, making it easier for users to recognize which actions were available to them. (5) Users were allowed to sort information in various screens, supporting users with supplying information most relevant at that point in time. (6) Users were allowed to filter evidence on certain properties, supporting users with supplying information most relevant at that point in time. (7) Horizontal scrolling was prevented in the matrix to ensure that users would not need to remember what scores belong to which person.

3.3 Design decisions regarding tunnel vision

Based on insights from the literature on TV, additional design decisions were made. As the TV prevention methods are mainly policy and practise based, we aimed to translate them to the area of HCI and support them using interface design artefacts: (1) The red and orange colours in the matrix overview aimed to provoke users in evaluating all evidence for all PoIs. In this way, users can quickly understand what evidence has been evaluated and what not. (2) When not all evidence is evaluated for all PoIs, the user is shown a warning to make them aware about the risk of tunnel vision when not all PoIs are evaluated against all evidence. (3) When adding new evidence, users are first asked to enter all information regarding the evidence and only then regarding PoIs. In this way, they are provoked to first focus on the evidence and then on the related PoIs. (4) When users add new evidence and are required to determine underlying factors, they are asked to add substantiation to their rating. In this way, they are provoked to actively and critically think about their decision. (5) It was decided to not give colours to PoIs with a higher score, as done in similar tools, as this could focus the users on specific scores without considering the information in the matrix. (6) The matrix page was designed such that it would be easy to consider all elements for all PoIs. This was done by creating a clean overview, with enough spacing. Yet, it was ensured that all evidence available on a PoI is directly visible.

Besides design decisions regarding information overload and tunnel vision, usability heuristics as described by [Nielsen, 2005] were taken into account.

4 Study Design

To evaluate the prototype we conducted user-based tests between June-July 2021. The evaluation aimed to understand the degree to which the target users could use the prototype successfully and measured perceived information overload, tunnel vision, and usability when interacting with the prototype. Moreover, users' opinions about the prototype were obtained. Beforehand, a pilot study was conducted to assess the evaluation and find opportunities to increase the quality and efficiency of the evaluation. User paths, interactions in the prototype, experimental procedure, task descriptions and surveys were revised based on participants' feedback.

4.1 Participants

Participants were recruited via a 'Community of Intelligence' newsletter and in direct (online) conversations. Snowballing was used to recruit additional participants. Participants were required to work in homicide investigation teams at the Dutch police, preferably as tactical analyst, or regularly work with scenarios and hypotheses. Eventually, eight participants (2 male, 6 female) were recruited. The mean age of participants was 43.6 (SD=6.6) and their work experience at the police ranged from 12 to 36 years (M=19.2, SD=9.5). Five worked as tactical analyst, one as tactical analyst coordinator, one as operational analyst and one as information coordinator.

4.2 Materials

The study was conducted remotely through Microsoft Teams due to the pandemic measures. The prototype was created using Figma and included the pages described in Section 3.2 [Figma Inc., 2016]. It was evaluated using remote testing platform Maze [Maze, 2021], as it permits synchronisation with Figma, allows creating expected user paths and surveys, and registers interaction. For the purpose of this evaluation, the prototype had to be based on a realistic scenario and have meaningful content. Given the disturbing content of the realistic homicide scenario, we decided to not include the scenario in the paper. The scenario can be obtained from the first author on request. Participants were informed that the scenario and contents were fictional and did not originate from a real homicide investigation. Realistic tasks were created for the scenario based on the main operations the tool intends to support (see Table 1). All tasks had one or more correct action paths. After the first seven tasks, new information was added to the scenario, namely that a new witness statement had been captured in the investigation. Task T8 considered this new information.

Task (T)	**Description**
T1	Open the case of Denise Henderson.
T2	Obtain all details about the person 'Hendriks, I.'.
T3	Understand what 'GO1' stands for and what evidence has been found for GO1 for the person 'de Jong, L.S.'.
T4	Obtain an overview of only tangible evidence in the investigation.
T5	Delete the evidence 'Camera registers Volkswagen Golf.'.
T6	Obtain an overview of evidence available for the person with the highest score at this moment in the investigation.
T7	Find the non-evaluated evidences and evaluate them.
T8	Add the age as described by the witness as new evidence.

Table 1: List of tasks during evaluation

4.3 Procedure

At the time of the evaluation, a connection over Microsoft Teams was established between the participant and researcher. Participants were informed about the research study including the main objective of C-SEL, the prototype, the scenario and the various tasks. Participants provided informed consent. If participants consented to recording, the researcher started the recording. Participants shared their screen, opened the link to Maze and read the introduction and scenario. Participants were informed that they did not need to remember the scenario in detail, as it would be available at all times. Thereafter, participants would start with the first task. Participants were instructed to think aloud when executing the tasks. After each task, they filled out a survey measuring perceived workload. On finishing all tasks, they filled out a post-evaluation survey measuring perceived usability, information overload and tunnel vision. Finally, they were asked verbally about their experience of the evaluation and if they had any comments or suggestions for improvement.

An alternative to the procedure had to be created as the pilot study revealed that when participants opened Maze in the secured Police environment, it had difficulty loading the prototype pages. Hence, it was necessary to perform the evaluation outside of the secured environment. Yet, some were uncomfortable with, or unable of, opening Teams and/or Maze outside the secured environment. In those three cases, the researcher opened Maze and shared her screen. Participants were then instructed to verbally explain what their actions would be as if they were in control of the mouse (e.g. where would they move their mouse, hover, click). The researcher did not take any action until explicitly instructed. The rest of the procedure remained similar.

4.4 Measurements

After each task, participants' perceived workload was measured using NASA-RTLX [Hart, 2006] translated into Dutch. NASA-RTLX is a simplified version of NASA-TLX, yet as least as sensitive and still with a high experimental validity [Hart, 2006; Byers, Bittner, and Hill, 1989]. It consists of six dimensions (mental demand, physical demand, temporal demand, performance, effort, frustration). Participants' overall perceived workload is the unweighted mean of the six subscales. Each dimension has corresponding items and measures load on a scale from 0-100. According to [Prabaswari, Basumerda, and Utomo, 2019], the score can be interpreted as low (0-10), medium (10-29), somewhat high (30-49), high (50-79), or very high (80-100).

Perceived usability, information overload, and tunnel vision were measured through a post-evaluation survey. Perceived usability was measured using the ten-item System Usability Scale (SUS) derived from [Brooke, 1996] and translated into Dutch. Items are rated on a 5-point Likert scale rating from 1 (I strongly disagree) to 5 (I strongly agree). Total usability score is calculated by subtracting 1 from the item score of the positive statements (X-1) and subtracting the score from the items score of the negative statements from 5 (5-X). The sum of these scores is then multiplied by 2.5. The SUS score ranges between 0-100, where a score $\geq$70 is considered as acceptable usability, $\geq$85 as excellent, whilst a score of $\leq$50 is considered poor or unacceptable usability [Bangor, Kortum, and Miller, 2009].

Perceived information overload was measured using the survey from [Chen, Shang, and Kao, 2009] adjusted to the context of this research and translated to Dutch. The survey consisted of seven items rated on a 5-point Likert scale rating from 1 (I strongly disagree) to 5 (I strongly agree). Items IO1 and IO5 measured perception of the 'adequacy of the information', IO2, IO3 and IO4 whether participants experienced 'too much information' and IO6 and IO7 perception of the 'quality of information'.

Perceived tunnel vision was measured using a self-constructed subjective scale as no existing scale existed to the best of our knowledge. The scale consisted of eight items rated on a 7-point Likert scale rating from 1 (I strongly disagree) to 7 (I strongly agree). The items were based on existing literature regarding tunnel vision. Items TV1 and TV5 measured participants' perceived 'type of focus', as defined by [Epskamp-Dudink and Winter, 2020]. TV2 and TV6 measured whether participants perceived to take 'all evidence in consideration', as defined by [Rassin, 2018]. TV3 and TV7 measured whether participants perceived to take 'multiple persons of interest in consideration', as defined by [Rassin, 2018] and [Posthumus, 2005]. Lastly, TV4 and TV8 measured perceived 'critical reflection in the investigation', as defined by [Posthumus, 2005]. The questions are shown in Appendix.

An additional open question was posed between tasks 6 and 7 and tasks 7 and 8. This question showed an image of the matrix at that moment in the investigation. Before Task 7 there would be evidence that had not yet been evaluated for certain PoIs. Whereas after Task 7 this evidence would have been evaluated. Thus, the first image showed a matrix with red and orange cells indicating unevaluated evidence, whereas in the second image all evidence would have been evaluated. The question asked was: "Seeing the matrix at this point in the investigation, what person or persons do you find most interesting?". This question was added to gain insight into participants' reasoning and get an indication of where participants focused on.

4.5 Analysis

Statical analysis was carried out using IBM SPSS (v27). To assess the internal consistency of each survey Cronbach's alpha was calculated after inversing negative items. An alpha above .70 is considered acceptable for research purposes [Hair, Anderson, Tatham, and Black, 1998].

Perceived workload for each task was calculated by computing the total mean of the six subscales for each participant for each task. Thereafter, the mean workload for each task was calculated. Boxplots and Shapiro-Wilk test indicated that the assumptions of normality for the NASA-RTLX were violated. Therefore, a non-parametric Friedman two-way ANOVA was used to compare the mean of the perceived workload across the tasks. The independent variable was task and the dependent variable was the perceived workload.

To validate the constructs of the IO and TV surveys, given these were newly adapted and/or created, we would have liked to conduct a factor analysis. Yet, this was unfeasible due to the limited sample size in relation to the number of variables. Therefore, to still be able to assess to some extent convergent and discriminant validity of the surveys, Spearman's rho's correlations were calculated between all items for each survey. Boxplots and Shapiro-Wilk test indicated that the assumptions of normality for the IO and TV surveys were violated. Therefore, a one-sample Wilcoxon Signed Rank test was used to compare the median score of the responses against the median of the scale. Before doing so, positive questions were inversed.

All recordings were analysed and relevant verbal comments regarding tasks, the open questions and general comments were transcribed. Afterward, all comments were organized in a structured manner, colour coded and translated into English.

5 Results

This section describes the results of the evaluation. Each session lasted approximately between 45 and 60 minutes.

5.1 Perceived Workload

Cronbach's alpha for the six-item NASA-RTLX was .84, thus considered adequate for research purposes. Table 2 indicates that the perceived workload of T1 to T6 can be considered 'medium' and that of T7 and T8 can be considered 'somewhat high'.

Task (T)	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**	**T8**
Avg. workload	19.37	19.38	23.96	22.50	17.71	26.87	33.33	36.67

Table 2: Perceived workload of individual tasks

A Friedman two-way ANOVA indicated that perceived workload varied significantly across the tasks $X^2=21.98$ (corrected for ties), df=7, p=.003. Follow-up pairwise comparisons with the Wilcoxon Signed Rank test and a Bonferroni adjusted α of .017 indicated that the workload of T5 (Mean rank=2.81) was perceived significantly lower than that of T7 (Mean rank=6.00), T = 36, z=-2.52, p=.012. The workload of T5 'Delete the evidence (Mean rank=2.81) was also perceived significantly lower than that of T8 (Mean rank=7.19), T = 36, z=-2.53, p=.011.

5.2 Perceived Usability

Cronbach's alpha for the ten-item SUS survey was .88, thus considered adequate for research purposes. Total mean SUS score was 78.1 (SD = 13.5), indicating an acceptable usability score. Fig. 6 shows the percent distributions of the participants' answers on the ten SUS items, negative items have been inversed. Fig. 6 indicates that participants responded positively to most items. Yet, one participant responded that the "tool was cumbersome to use".

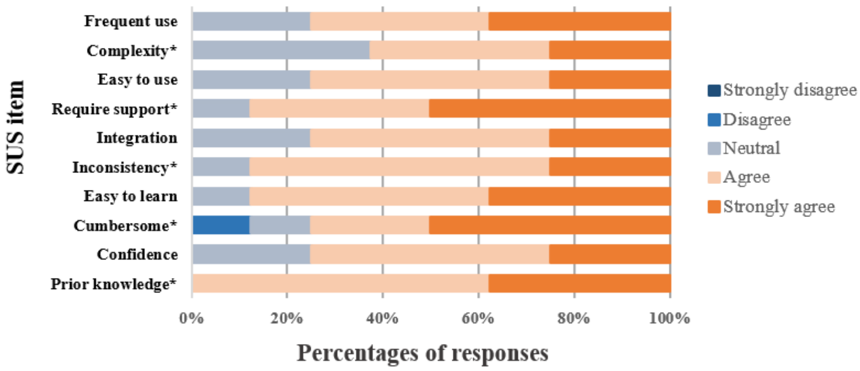

*Figure 6. Percent distribution of item responses on the SUS items. *=item inversed*

		Adequacy of information		Too much information			Quality of information	
		IO1	**IO5**	**IO2**	**IO3**	**IO4**	**IO6**	**IO7**
Adequacy of information	**IO1**	1.00						
	IO5	**-.72***	1.00					
Too much information	**IO2**	-.35	.17	1.00				
	IO3	.00	.20	-.35	1.00			
	IO4	-.08	.02	**.78***	**-.80***	1.00		
Quality of information	**IO6**	.13	-.22	-.62	.56	**-.77***	1.00	
	IO7	-.36	.54	.26	**.73***	-.25	.00	1.00

Table 3: Spearman's Rho correlation of IO items. is significant at .05 level (2-tailed)*

5.3 Perceived Information Overload

Cronbach's alpha for the seven-item perceived IO survey was .70, thus considered adequate for research purposes. Spearman's rho correlation test was used to assess the size and direction of the linear relationship between the various items. Table 3 shows the Spearman's rho results. IO1 and IO5 were strongly correlated, which was expected as according to Chen et al. (2009) these items belong to the same construct, namely 'adequacy of information'. IO2 and IO4 were strongly correlated and IO3 and IO4 strongly negatively correlated. This was also expected as, according to Chen et al. (2009), IO2, IO3 and IO4 belong to the same construct, namely 'too much information'. It was also expected that IO2 and IO3 would strongly correlate, yet, their correlation was only moderate and not statistically significant. Lastly, IO3 and IO7 were strongly positively correlated and IO4 and IO6 strongly negatively. Even though these did not belong to the same construct, it was considered reasonable that these items had a strong correlation, after examining the item content. Fig. 7 shows the percent distributions of responses to the IO items, with positive items inversed.

A one-sample Wilcoxon signed rank test with $\alpha = .05$ was used to compare the median of the IO survey responses against the scale midpoint. Positive items were inversed before the analysis. Results indicated that the median response (2) was significantly lower than the scale midpoint (3), indicating that perceived IO was low.

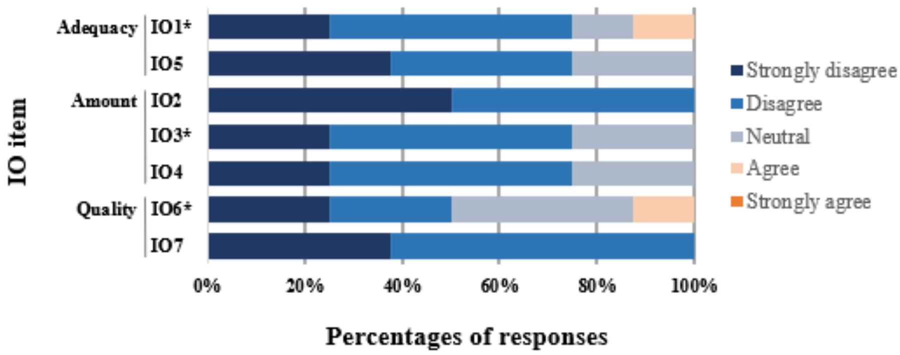

*Figure 7: Percent distribution of responses to perceived IO items. *: inversed items.*

		Type of focus		Consider all evidence		Consider multiple PoIs		Critical reflection	
		TV1	TV5	TV2	TV6	TV3	TV7	TV4	TV8
Type of focus	TV1	1.00							
	TV5	**.84****	1.00						
Consider all evidence	TV2	.01	.31	1.00					
	TV6	**.88****	**.77***	.14	1.00				
Consider multiple PoIs	TV3	**.82***	**.73***	-.07	**.72***	1.00			
	TV7	-.19	.07	.28	-.24	-.17	1.00		
Critical reflection	TV4	.62	.21	-.20	.62	.37	**-.70**	1.00	
	TV8	-.60	-.58	.06	-.57	-.60	.45	-.60	1.00

*Table 4: Spearman's Rho correlation of TV items. * significant at .05, ** at .01*

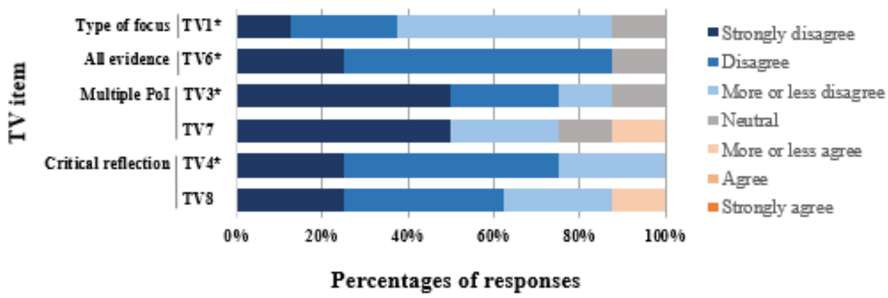

*Figure 8: Percent distribution of responses on TV items. * is item inversed*

5.4 Perceived tunnel vision

Cronbach's alpha for the eight-item perceived TV survey was .41, thus considered inadequate for research purposes. Spearman's rho correlations were used to assess the size and direction of the linear relationship between the various items, and to inspect if there were correlations that could account for the low internal consistency. The results of Spearman's rho are presented in Table 4. There was a strong positive correlation between TV1 and TV5. As these theoretically should be related, a strong correlation was expected. Yet, TV5 was intended to be the inverse of TV1, with evidence focused being the opposite of person focused. Hence, a negative correlation was expected. This indicates that the interpretation of TV5 was different than intended. Therefore, TV5 was removed from further analysis. Moreover, verbal comments indicated that TV2 was difficult to understand. Results showed that participants' answers to TV2 were often inconsistent with the rest of their answers. Additionally, the correlation matrix indicated that the correlation of TV2 was weak with all other items. Therefore, TV2 was removed from further analysis. After removing TV2 and TV5, Cronbach's alpha for the six-item perceived TV survey was .77, so considered adequate for research purposes. Figure 8 shows the percent distributions of participants' answers on the TV items, with positive items inversed.

A one-sample Wilcoxon signed-rank test with an α of .05 was used to compare the median of the responses to the TV survey against the midpoint of the scale. Positive items were inversed before this analysis. Results indicated that the median response (2) was significantly lower than the midpoint of the scale (4), indicating that perceived tunnel vision was low. However, caution is required, as items were removed after seeing the results and it is possible that the response to TV5 indicates that the tool can prevent tunnel vision (by allowing people to focus on the evidence) and encourage it (by allowing people to focus on a person).

5.5 Findings from observations and think aloud

Several findings arose from observations during the user-test. Those findings were either actions performed or comments expressed. Generally, participants had a positive attitude towards the tool and considered it to be intuitive, easy to use, and provide a pleasant visual presentation. They also said that it could be of added value during homicide investigations, providing reasons such as *"provides a good overview of all evidence available and all persons of interest in the investigation"* [EVAL-P4], *"allows you to see where your information is scarce"* [EVAL-P8], *"ensures that there is an objective view"* [EVAL-P1], *"would be a good supplementary tool* [..]*, especially when you have a case with many Persons of Interest."*[EVAL-P6]

Additionally, some general comments and suggestions were given. Some participants felt that the tool would require unnecessary duplication of work and that it needed to be connected to their current information system. Two participants said that tool acceptance could be improved by providing more detail on how to interpret a score and how scores were determined. One emphasized that when the tool gets implemented, it should be seen as an addition and not a substitute. Three expressed appreciation for being included in the design and evaluation process.

Perceived Workload. Some observations were made related to perceived workload. First, participants did not immediately expect that a cell within the matrix overview was clickable (T3) and/or did not notice the top-menu bar (T4), experiencing these tasks to require more effort. Yet, observations showed that when they learned that they could click on a cell or where to look for the evidence tab, they were able to apply this knowledge easily in later tasks (e.g., in T5 and T7). One said: *"Now I knew that there was an evidence page, so I did not have to think about it anymore."* [EVAL-P8]. Moreover, comments also indicated that some felt Task 8 required more thinking, for example: *"This [Task 8] was more of a thinking task, as you have to evaluate the weights."* [EVAL-P5]

Perceived Usability. Participants' actions and comments provided much feedback with regard to the usability of the prototype. For instance, when aiming to obtain detailed information about a specific PoI (T2), half the participants assumed that clicking on the checkbox would allow them to perform additional actions. Additional improvements of the checkbox were pointed out when participants aimed to obtain more information about the PoI with the highest score (T6). Some expected that they could select the checkbox of a PoI and then could click on the evidence page. They expected that this would result in all evidence for that PoI. When this did not work, they expressed feeling lost and unable to finish the task. Moreover, as previously mentioned, some did not immediately expect that a matrix cell was clickable and suggested making this more apparent (T3). Moreover, when adding new evidence, they expected to be able to click on an element in the matrix and give evidence from there (T8).

Perceived Information Overload. With regards to participants' perceived IO, additional observations could be made from actions and verbal expressions. When participants tried to understand where a specific element abbreviation referred to (T3), it became apparent that there was some ambiguity between the element abbreviations. Some participants' focus shifted to 'GMO1' and assumed that this was the required element. For these participants, it took some time to realize that they had to search for 'GO1'. Moreover, during T4 it could be observed that a participant initially experienced there to be much information. When searching for all available tangible evidence, this participant did not notice the filter option and started selecting all tangible evidence available in the evidence list. Yet, when it became apparent that this would require selecting many rows, they noticed the filter option. They indicated that they did not know where to look on the screen at first. Moreover, ambiguity amongst participants was experienced when adding new evidence (T8). Verbal comments indicated that almost all experienced difficulty interpreting the descriptions of the element types. Some expressed feeling lost and did not know which one to select. This ambiguity was also found with the interpretation of the underlying factors, as participants did not understand what was meant or interpreted factors differently, e.g. "*Objectivity is unclear to me, as we are talking about a witness statement*." [EVAL-P6] and "*I do not understand what is meant by 'unique'. I can't say if it is unique, as I do not know all the evidence that is currently registered*." [EVAL-P2]

Perceived Tunnel Vision. With regards to perceived TV, several observations can be made. When receiving the warning notification to evaluate all non-evaluated evidence (T7), several participants did not read the (entire) warning before proceeding. Moreover, when adding new evidence, two participants expressed that it would be helpful if they could immediately evaluate new evidence against existing PoIs. Moreover, one expressed to have preferred to see the initial score as well as the calculated score, yet also pointed out the potential problem: "*I would want to know what the initial score was to test if the calculation was done properly. Yet, I know that there are also disadvantages, as we as humans are not capable of reasoning completely objectively. Thus, some people might get influenced when they see the initial and adjusted scores*." [EVAL-P5].

In regards to TV, the open question led to additional useful insights. Participants' comments indicated that when inspecting the matrix, they immediately focused on the red and orange columns. Participants expressed that they felt as if they should do something with those columns first. Moreover, when answering the questions, these participants took the persons related to the red and orange column into account when formulating their answers, e.g., *"I would prefer to verify Franssen and Hendriks first. Yet, that does not mean that Schipper and de Jong are not interesting, as their scores are close to each other. So I would say the top three and Hendriksen."* [EVAL-P4]. Furthermore, two participants mentioned that they felt as if the matrix alone did not provide them with enough information to answer the question and that they would like to see more details first. Additionally, participants mentioned that they would like to have more insight into the meaning of a score. One participant elaborated on this: *"I would like to give meaning to the scores. For me, it would be helpful if I could see what the maximum score of an element is. For instance, if the maximum score is 50, and the evidence scores 32, I would like to see 32/50."* [EVAL-P3].

6 Discussion

Overall, findings showed participants were greatly interested in and excited about the tool. Moreover, comments indicated that a design process considering end-users was appreciated. Generally, participants perceived high levels of usability and low levels of task load, information overload, and tunnel vision using the tool.

The SUS score (above average score of 78) indicated that the tool had acceptable usability and individual responses showed that participants did not feel as if they had to learn much before using the tool. Yet, some participants perceived the tool to be cumbersome at some points. This was supported by observations where it became apparent that some participants expected to perform certain actions differently. A possible solution to this problem would be training the users before using the tool by adding an onboarding tutorial for new users. This could also prevent problems of overlooking the top-menu bar and the filtering options. Given the tool's intended use, it would be acceptable to have a short initial training.

In terms of task load, the results suggested that most tasks' load was considerably low, indicating that completing the tasks did not require a high cognitive load. Tasks 7 and 8 required more effort; yet, this was expected as those tasks required more thinking. Moreover, Task's 5 considerably low load can be explained by a learning effect: observations showed that some participants had difficulty finding the evidence page in Task 4 at first, but had learned to use it by Task 5.

The information overload survey indicated that participants did not experience 'too much information to deal with' when working with the tool and they felt as if they could 'find the information needed'. Moreover, most responded that they had carefully considered all information available in the tool. Yet, some pointed out that they would have looked more in-depth when they would not have been executing specific tasks. Furthermore, verbal comments indicated that participants did experience some ambiguity when working with the tool, indicating that more clarification might lower perceived IO even further. A possible improvement would be changing the labels of the abbreviations, for instance by replacing them with a text label (e.g. 'age', 'gender', 'camera'). Besides preventing confusion, this could also support the accuracy of the interpretation as well as the aspect of 'recognition rather than recall' as shown in research by [Islam, 2015]. In doing so, one could argue that the labels would lose their strength since they are no longer presenting information regarding element categories, i.e., motive, means, or opportunity. Therefore, one could consider creating a header on top of the elements, similar to the 'individual' and 'group' headers that were already in the prototype. Nonetheless, further research is necessary to determine an optimal solution. Second, the descriptions of the elements should be reconsidered and rephrased, as some were experienced to be unclear and misleading. This improvement suggestion also applies to the phrasing of the underlying factors. To improve the understandability of the elements and underlying factors, it may be useful to provide users with a 'help' button that shows an explanation where necessary. This ensures that the underlying factors are less ambiguous and that users interpret them similarly.

In terms of perceived TV, both the survey and observations indicated that most participants felt as if the tool would support them in 'considering multiple persons of interest', as defined by [Rassin, 2018] and [Posthumus, 2005], and that it would support them in 'reviewing the investigation critically', as defined by [Posthumus, 2005]. Moreover, participants perceived that the tool would support them in considering 'all evidence for all persons of interest', as defined by [Rassin, 2018]. It became clear that participants got triggered by the red and orange cells. However, as stated previously, some caution is needed with the tunnel vision scale interpretation, as the results indicated that participants felt as if the tool both allowed them to 'focus on evidence' as well as 'focus on a person', as defined by [Epskamp-Dudink and Winter, 2020]. This might indicate that the tool both prevents and enables tunnel vision to some extent. Yet, the results also showed that most participants thought that this tool would not make them focus on one PoI solely. Hence, it could indicate that participants felt the tool would make them focus on PoIs, but not necessarily one PoI and so might not induce tunnel vision. This interpretation is supported by responses to the other questions and actions. Yet, further research is needed to substantiate this.

Furthermore, improvements can be made that might support the prevention of TV even further. For instance, users should be able to evaluate existing PoIs against new evidence immediately. Several participants mentioned this to be useful and it could improve the aspect of tunnel vision 'consider all evidence for all persons of interest' as defined by [Posthumus, 2005]. Also, several said that they would prefer to have more insight into the scores, both in the summary when adding evidence and in the matrix overview. Therefore, we consider showing the initial and adjusted score, as well as how they have been calculated. This might not only increase usability, but may also improve the aspect of tunnel vision by creating 'greater transparency of the system' as defined by [Findley and Scott, 2006] and [Posthumus, 2005].

We would have liked to compare our results to those achieved with other methodologies that could be used for investigations and have a prototype and/or tool. For instance, by comparing our perceived TV and IO to that achieved with the collaborative prototype software application ACH Walkthrough tool (ACH-W) [Wilson, Brown and Biddle, 2014], which is based on the Analysis of Competing Hypotheses (ACH) methodology. They aimed to create a prototype that supports collaborative sensemaking. Yet, they do not specifically test their prototypes' effect on TV and IO. We would also have liked to compare it with TRACE developed by [Strömer-Galley et al., 2020]. They created a crowd source-based, analytical technique to support intelligence analysts through the analytical process. It aims to promote critical analysis by reducing users' over-reliance on memory and the likelihood that cognitive biases drive the analysis, in a flexible way. Moreover, it aims to support users in managing cognitive load. They did not explicitly link their design decisions to IO and TV. However, they did study the impact on reasoning and found promising results. Currently, both do not specifically measure perceived TV and IO. Therefore, it is hard at this point to compare the results of our prototype directly to these studies. Yet, as both alternative systems aim to support the analytical process and could help in the prevention of IO and TV, it would be interesting to compare C-SEL to ACH-W and TRACE in a future study.

7 Conclusion

In this paper, we discussed the impact of a system on information overload and tunnel vision of criminal investigators. We presented a methodology for the criminal investigation process (C-SEL) as well as an interface for supporting software that intends to reduce information overload and tunnel vision. A prototype was filled with a fictional homicide investigation and a user study indicated that criminal investigators experienced a significantly low level of information overload and tunnel vision using the prototype. Moreover, the results showed acceptable usability and verbal statements indicated a largely positive attitude towards the prototype. This research clearly shows the opportunity to use interface design artefacts to support the prevention of information overload and tunnel vision.

C-SEL (and its interface for supporting software presented in this paper) can be used to support and steer forthcoming homicide investigations in various ways (e.g., at the level of tactical cues, incorporating interesting people, assessing cues relative to a particular person). It can thus be used as guidance during many different phases of the investigation. In principle, it will not change the police work itself, but can be used by an analyst to steer the investigation in a way that avoids information overload and tunnel vision. C-SEL has been tested on three solved cases [Sutmuller et al., 2020]. To provide further support for its (and its tool's) potential to prevent tunnel vision, it could also be tested on cases in which somebody was wrongly convicted and the real perpetrator was convicted later (e.g. the Schiedam park murder [Posthumus, 2005], to investigate how C-SEL would have prioritized these two people.

Several limitations should be considered in relation to the findings. Firstly, since this was an exploratory study, only a small sample was used. To perform statistical analyses to discover the effect on information overload and tunnel vision more participants are needed. Yet, significant results were found despite the small sample.

Secondly, due to the pandemic, physical contact was impossible, so, all evaluations had to be done online, resulting in different experimental setups amongst participants. For instance, some participants performed the evaluation in a noisy environment, whilst others did not. Moreover, three evaluations were conducted using screen sharing by the researcher. This may have led to participants feeling less immersed in the tool influencing their perceptions. Furthermore, some had a slower internet connection, resulting in the tool responding slower as well, which may have impacted perceived task load and usability. However, whilst this may have influenced opinions, the statistical analysis is not impacted as a within-subjects design was used.

Thirdly, it is hard to determine what design decisions specifically had an effect on the low perceived information overload and tunnel vision. Further research is needed to see individual effects of design artefacts on information overload and tunnel vision.

Fourthly, the high-fidelity prototype was not fully interactive and did not contain all pages, so, some actions were impossible during the study. Though all necessary actions were available, this may have influenced the results to some extent. For instance, participants were unable to see multiple pages in the matrix, limiting the number of PoIs available. Whilst it is likely that in a real homicide investigation there may be more than 30. Future research will evaluate the tool with full functionality.

Fifthly, only subjective measurements were used in this study. Future work should also consider more objective measurements, especially to test tunnel vision. This could be done similar to research by [Fahsing and Ask, 2016]. Likewise, research could compare C-SEL with methods currently mentioned by participants such as mind mapping. Moreover, a comparison could be made between C-SEL and other existing tools that aimto prevent tunnel vision and/or information overload in investigations such as ACH-W and TRACE [Wilson et al., 2014; Stromer-Galley et al., 2021]. This however requires a working implementation of the tool, allowing the users to use it in a complete manner. Thereby, it would be necessary to take the design improvements, as mentioned above, into consideration.

Sixthly, only one version of the tool and its components was tested. Yet, there are different ways to create the same functionality. Therefore, further research should consider using A/B testing to test those differences. This could for instance be useful to test different ways of displaying the scores or showing a PoI. Finally, C-SEL was designed for homicide investigations. Yet, information overload and tunnel vision also arise in other criminal investigations. Future research will determine whether the results can be generalized.

Acknowledgements

The authors would like to thank the police officers who participated in this study.

References

[ACPO, 2006] ACPO: "Murder Investigation Manual."; National Centre for Police Excellence, Wyboston (2006).

[Ask and Granhag, 2005] Ask, K., & Granhag, P. A.: "Motivational sources of confirmation bias in criminal investigations: the need for cognitive closure."; J. of Investigative Psychology and Offender Profiling, 2, 1 (2005), 43-63.

[Baber, 2006] Baber, C.: "Crime scene investigation as distributed cognition."; Pragmatics and Cognition, 14, 2 (2006), 357-385.

[Baddeley, 1992] Baddeley, A.: "Working Memory."; Psychology of Learning and Motivation, 255, 5044 (1992), 556-559.

[Bangor, Kortum and Miller, 2009] Bangor, A., Kortum, P., & Miller, J.: " Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale."; J. of Usability studies, 4, 3 (2009), 114-123.

[Bawden and Robinson, 2009] Bawden, D., & Robinson, L.: "The dark side of information: overload, anxiety and other paradoxes and pathologies."; J. of Inf. Sci, 35, 2 (2009), 180-191.

[Bawden and Robinson, 2020] Bawden, D., & Robinson, L.: "Information Overload: An Overview."; Oxford Encyclopaedia of Political Decision Making, Oxford University Press: Oxford, UK (2020), 1-60.

[Benselin and Ragsdell, 2016] Benselin, J. C., & Ragsdell, G.: "Information overload: The differences that age makes."; J. of Librarianship and Inf. Science, 48, 3 (2016), 284-197.

[Benyon, 2014] Benyon, D.: "Designing interactive systems: A comprehensive guide to HCI, UX and interaction design"; Harlow: Pearson Education Limited (2014).

[Bronkhorst, 2014] Bronkhorst, L.: "Politiële tunnelvisie, kan het worden voorkomen?"; Scriptie Master Strafrecht en Criminology, Rijkuniversiteit Groningen, Groningen (2014).

[Brooke, 1996] Brooke, J.: "SUS - A quick and dirty usability scale."; In Usability evaluation in industry, Taylor and Francis, London (1996), 189-194.

[Brookman, Maguire and Maguire, 2019] Brookman, F., Maguire, E. R., & Maguire, M.: "What Factors Influence Whether Homicide Cases Are Solved? Insights From Qualitative Research With Detectives in Great Britain and the United States."; Homicide Studies, 23, 2 (2019), 145-174.

[Byers, Bittner, and Hill, 1989] Byers, J. C., Bittner, A. C., & Hill, S. G.: "Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary?"; Advances in Industrial Ergonomics and Safety (1989), 481-485.

[Cao, 2008] Cao, Y.: "An information assistant system for the prevention of tunnel vision in crisis management."; Centre for Telematics and Information Technology, University of Twente, Enschede, the Netherlands (2008), 1381-3625.

[Chen, Shang, and Kao, 2009] Chen, Y. C., Shang, R. A., & Kao, C. Y.: "The effects of information overload on consumers' subjective state towards buying decision in the internet shopping environment. Electronic Commerce Research and Applications, 8, 1 (2009), 48–58.

[Chewing and Harrell, 1990] Chewning, E. G., & Harrell, A. M.: "The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task."; Accounting, Organizations and Society, 15, 6 (1990), 527-542.

[Eppler and Mengis, 2004] Eppler, M. J., & Mengis, J.: "The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines."; The Information Society, 20, 5 (2004), 325-344.

[Epskamp-Dudink, 2016] Epskamp-Dudink, C.: "Niet te filmen! Over retrospectief scenario denken in de opsporingspraktijk"; Politieacademie, Apeldoorn (2016).

[Epskamp-Dudink and Winter, 2020] Epskamp-Dudink, C., & Winter, J.: "Benefits of scenario reconstruction in Cold Case investigations." J. of Criminal Psychology, 10, 2 (2020), 65-78.

[Evans, 2017] Evans, D. C.; "Bottlenecks: Aligning UX Design with User Psychology."; Apress, Berkeley, California (2017).

[Farhoomand and Drury, 2002] Farhoomand, A. F., & Drury, D. H.: "Managerial information overload."; Communications of the ACM, 45, 10 (2002), 127-131.

[Findley and Scott, 2006] Findley, K. A., & Scott, M. S.: "The Multiple Dimensions of Tunnel Vision in Criminal Cases."; Wisconsin Law Review, 1023, 2 (2006), 291–398.

[Goette, Han and Leung, 2019] Goette, L., Han, H.-J., & Leung, B. T. K.: "Information Overload and Confirmation bias."; Cambridge Working Papers in Economics (2019), 1-48.

[Gruijter and Poot, 2019] Gruijter, M. de, & Poot, C. J. de.: "The use of rapid identification information at the crime scene; similarities and differences between English and Dutch CSIs."; Policing and Society, 29, 7 (2019), 848-868.

[Fahsing and Ask, 2016] Fahsing, I., & Ask, K.: "The making of an expert detective: the role of experience in English and Norwegian police officers' investigative decision-making. Psychology, Crime & Law, 22, 3 (2016), 203-223.

[Fahsing and Ask, 2017] Fahsing, I., & Ask, K.: "In search of indicators of detective aptitude: police recruits' logical reasoning and ability to generate investigative hypotheses"; J. of Police and Criminal Psychology, 33, 1 (2018), 21-34.

[Figma Inc., 2016] Figma. (2016) Retrieved from www.figma.com.

[Findley and Scott, 2006] Findley, K. A., & Scott, M. S.: "The Multiple Dimensions of Tunnel Vision in Criminal Cases."; Wisconsin Law Review, 291 (2006), 312-13.

[Galbraith, 1974] Galbraith, J. R.: "Organization Design: An Information Processing View."; Interface, 4, 3 (1974), 28-36.

[Groenendaal and Helsloot, 2014] Groenendaal, J., & Helsloot, I.: "Tunnel vision on tunnel vision? A preliminary examination of the tension between precaution and efficacy in major criminal investigations in the Netherlands"; Police Practice & Research, 16, 3 (2014), 224-238.

[Hallihan, Cheong and Shu, 2012) Hallihan, G. M., Cheong, H., & Shu, L. H.: Confirmation and Cognitive Bias in Design Cognition. Proc. of the ASME 2012 Inter. Design Engineering Technical Conf. & Computers and Information in Engineering Conf. New York: ASME (2012).

[Hair, Anderson, Tatham and Black, 1998] Hair, J., Anderson, R., Tatham, R., & Black, W.: "Multivariate Data Analysis." Prentice Hall (1998).

[Hanani, Shapira and Shoval, 2001] Hanani, U., Shapira, B., & Shoval, P.: "Information Filtering: Overview of Issues, Research and Systems."; User Modeling and User-Adapted Interaction, 11, 3 (2001), 203-259.

[Hanshew, 2010] Hanshew, K.: "Daring More Democracy? Internal Security and the Social Democratic Fight against West German Terrorism."; Central Eur. Hist., 43, 1 (2010), 117-147.

[Hart, 2006] Hart, S.G.: "Nasa-Task Load Index (NASA-TLX); 20 Years Later."; Proc. of the Human Factors and Ergonomics Society Annual Meeting, 50, 9 (2006), 904-908.

[Hawk and Dabney, 2014] Hawk, S. R., & Dabney, D. A.: Are All Cases Treated Equal?: Using Goffman's Frame Analysis to Understand How Homicide Detectives Orient to Their Work. British J. of Criminology, 54, 6 (2014), 1129-1147.

[Helsloot and Groenendaal, 2012] Helsloot, I., & Groenendaal J.: "Op zoek naar de balans tussen daadkracht en voorzorg in de opsporing."; Stichting Crisislab (2012).

[Heuer, 1999] Heuer, R. J.: "Analysis of Competing Hypotheses." Psychology of Intelligence Analysis. Centre for the Study of Intelligence, (1999), 95-110.

[Holst, 2020] Holst, A.: "Total data volume worldwide 2010-2014"; (2020), https://www.statista.com/statistics/871513/worldwide-data-created/

[Hoppmann, 2013] Hoppmann, G.: "Die Entwicklung der Rasterfahndung und DNA-Reihenuntersuchung"; Krim (2013), 219-226.

[Innes, 2003] Innes, M.: "Investigating Murder: Detective Work and the Police Response to Criminal Homicide."; Oxford University Press, New York (2003).

[Islam, 2015] Islam, M. N.: "Exploring the intuitiveness of iconic, textual and icon with texts signs for designing user-intuitive web interfaces." In 2015 18th Int. Conf. on Computer and Information Technology (2015), 450-455.

[Jones and Kelly, 2018] Jones, S. L., & Kelly, R.: "Dealing With Information Overload in Multifaceted Personal Informatics Systems."; Human-Computer Interaction, 33, 1 (2018), 1-48.

[Kahneman and Tversky, 1973] Kahneman, D., & Tversky, A.: "On the psychology of prediction."; Psychological Review, 80, 4 (1973), 237-251.

[Koffka, 2013] Koffka, K.: "Principles of Gestalt Psychology"; Harcourt Brace, NY (2013).

[Knauff and Wolf, 2010] Knauff, M., & Wolf, A.G.: "Complex cognition: the science of human reasoning, problem-solving, and decision-making."; Cognitive Processing, 11 (2010), 99-102.

[Knight, 2019] Knight, W.: "The Importance of Visual Design."; In UX for Developers, CA Press: Berkeley, 103-126.

[Lau and Redlawsk, 2001] Lau, R. R., & Redlawsk, D. P.: "Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making."; American J. of Political Science, 45, 4 (2001), 951-971.

[Liedenbaum, Poot, Vergouw and Kouwenberg, 2015] Liedenbaum, C., Poot, C. de, Vergouw, S., & Kouwenberg, R.: "Focus in de opsporing"; Wetenschappelijk Onderzoek- en Documentatiecentrum, Ministerie van Justitie & Veiligheid, Boom Lemma, Den Haag (2015).

[Maze, 2021] Maze. (2021) Retrieved from www.maze.com.

[Meyer, 1998] Meyer, J.: "Information overload in marketing management."; Marketing Intelligence and Planning, 16, 3 (1998), 200-209.

[Nielsen, 2005] Nielsen, J. Ten Usability Heuristics (2005).

[Oviatt, 2006] Oviatt, S.: "Human-Centered Design Meets Cognitive Load Theory: Designing Interfaces that Help People Think."; Proc. of ACM Int. Conf. on Multimedia (2006), 871-880.

[Posthumus, 2005] Posthumus, F.: "Evaluatieonderzoek in de Schiedammer parkmoord"; Openbaar Ministerie (2005).

[Prabaswari, Basumerda and Utomo, 2019] Prabaswari, A. D., Basumerda, C., & Utomo, B. W.L : "The Mental Workload Analysis of Staff in Study Program of Private Educational Organization" IOP Conf. Series: Materials Science and Engineering, 528, 1 (2019).

[Preece, Sharp and Rogers, 2015] Preece, J., Sharp, H. & Rogers, Y.: "Interaction Design: Beyond Human-Computer Interaction"; John Wiley & Sibs Ltd, West Sussex (2015).

[Rassin, 2018] Rassin, E.: "Reducing tunnel vision with a pen-and-paper tool for the weighting of criminal evidence."; J. of Inves. Psychology & Offender Profiling, 15, 2 (2018), 227-233.

[Rossmo, 2016] Rossmo, D. K.: "Case rethinking: a protocol for reviewing criminal investigations."; Police Practice and Research, 17, 3 (2016), 212-228.

[Salet, 2015] Salet, R.: "Opsporing, tegenspraak en veranderende frames: een onderzoek naar tegenspraak in grootschalige rechercheonderzoeken."; Boom Lemma, Den Haag (2015).

[Salet and Terpstra, 2014] Salet, R., & Terpstra, J.: "Critical review in criminal investigation: evaluation of a measure to prevent tunnel vision."; Policing: a journal of policy and practice, 8, 1 (2014), 43-50.

[Savolainen, 2007] Savolainen, R.: "Filtering and withdrawing: strategies for coping with information overload in everyday contexts"; J. of Info. Science, 33, 5 (2007), 611-621.

[Shick, Gordon, and Haka, 1990] Schick, A. G., Gordon, L. A., & Haka, S.: "Information overload: A temporal approach."; Accounting, Organizations & Society, 15, 3 (1990), 199-220.

[Snook and Cullen, 2008] Snook, B., & Cullen, R.: "Bounded Rationality and Criminal Investigations."; In Criminal Investigative Failures (2008), 71-98.

[Stelfox, 2011] Stelfox, P.: "Criminal Investigations: Filling the Skills Gap in Leadership, Management, and Supervision. Policing, 5, 1 (2011), 15-22.

[Stromer-Galley et al., 2021]. Stromer-Galley, J., Rossini, P., Kenski, K., McKernan, B., Clegg, B., Folkestad, J., Osterlund, C., Schooler, L. Boichak, O., Canzonetta, J., Martey, R., Pavlich., C., Tsetsi, E & McCracken, N.: "Flexible versus structured support for reasoning: enhancing analytical reasoning through a flexible analytic technique. Intelligence and National Security", 36(2), 279-298.

[Sutmuller, 2021] Sutmuller, A.D.: "Homicide investigation in the digital era: The development and evaluation of a case-specific elements library (C-SEL)"; Doctor of Philosophy, Delft University of Technology (2021).

[Sutmuller, den Hengst, Baros and van Gelder, 2020] Sutmuller, A. D., den Hengst, M., Baros, A. I., & van Gelder, P.: "Getting the Perpetrator Incorporated and Prioritized in Homicide Investigations: The Development and Evaluation of a Case-Specific Element Library (C-SEL)."; Inter. J. of environmental research and public health, 17, 17 (2020), 6430.

[Tversky and Kahneman, 1974] Tversky, A., & Kahneman, D.: "Judgment under Uncertainty: Heuristics and Biases."; Science, 185, 4157 (1974), 1124-1131.

[van Koppen, Hessing, Merckelbach and Crombach, 2002] van Koppen, P., Hessing, D., Merckelbach, H. L. G. J., & Crombag, H. F. M.: "Het recht van binnen: Psychologie van het recht."; Deventer: Kluwer (2002).

[van Nimwegen, van Oostendorp, Burgos and Koper, 2006] van Nimwegen, C., van Oostendorp, H. Burgos, D., & Koper, R.: "Does an Interface with Less Assistance Provoke More Thoughtful Behavior?"; Inter. Conf. of the Learning Sciences, Bloomington, Indiana (2006), 785-791.

[van Wijk, van Leiden and Hardeman, 2017] van Wijk, A., Leiden, I., and Hardeman, M.: "Modus operandi van de recherche: de recherchepraktijk in moord- en verkrachtingszaken."; Politie & Wetenschap, Politiekunde 87 (2017).

[Wilson, 2012] Wilson, L.: "Criminal Major Case Management Persons of Interest Priority Assessment Tool (POIPAT)."; CRC Press: Boca Raton, FL, USA (2012).

[Wilson, Brown & Biddle, 2014]. Wilson, J., Brown, J.M., Biddle, R.: "ACH walkthrough: A distributed multi-device tool for collaborative security analysis". In Proceedings of the 2014 ACM Workshop on Security Information Workers, 9-16.

[Zamanirad, Benatallah, Barukh, Rodriguez and Nouri, 2020] Zamanirad, S., Benatallah, B., Barukh, C., Rodriguez, C., & Nouri, R.: "Dynamic Event Type Recognition and Tagging for Data-driven Insights in Law-Enforcement."; Computing, 102, 7 (2020), 1627-1651.

Appendix. Perceived Tunnel Vision Questionnaire

[7-point Likert scale. 1 = Completely disagree, 5 = Completely agree]

1. This tool would ensure that I have an evidence-based focus.
2. This tool would prevent me from considering all evidence for all Persons of Interest
3. This tool would ensure that I consider different Persons of Interest.
4. This tool would ensure that I look critically at the investigation.
5. This tool would ensure that I have a person-based focus.
6. This tool would ensure that I consider all evidence for all Persons of Interest.
7. This tool would ensure that I only consider one Person of Interest.
8. This tool would prevent me from looking critically at the investigation.