LONDON’S GLOBAL UNIVERSITY

Possible Sources and Impact of Stress on Decision-Making in Forensic Science

Mohammed A. Almazrouei

In partial fulfilment of the requirements for the degree of Doctor of Philosophy

University College London | September 2022

Supervisors | Prof. Ruth M. Morgan and Dr. Itiel E. Dror

Declaration

I, Mohammed Ali Almazrouei, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: *Mohammed Almazrouei*

Date: *20 September 2022*

nothing is impossible | لا شيء مستحيل

H.H Sheikh Mohammed bin Rashid Al Maktoum | صاحب السمو الشيخ محمد بن راشد آل مكتوم

Acknowledgments

It is time that I reflect on my PhD journey. With its ups and downs, there were people that were by my side that I cannot thank enough.

Huge thanks to my supervisors, Prof. Ruth Morgan and Dr. Itiel Dror, for their guidance and support throughout the PhD journey. Also, I am very grateful for a third supervisor, Assoc. Prof. Ifat Levy, for welcoming and working with me during an exchange program at Yale University.

My mother, my father, brothers and sister, my Aunt Ghaya, cousins Saif and Khalid –I can't imagine how my journey would have been without you all.

In addition, during the PhD time, I met many in London that I consider life-long friends. In no particular order, I am very thankful to my friends: Salim, Shaikha, Nicola, Amir, Beshayer, Somaya, Ahmad (Bu Ameera), Laila, Fabio, Pete, Lee, Julian, Rouda, Ronald, Hamad, Maitha, Alex and Nick.

Not to forget, this would not have been possible, without the scholarship from Abu Dhabi Police (ADP). Thanks very much to ADP leaders for giving me this amazing opportunity, and for everyone at the Police Attaché in the UK and the ADP Education Department for their support.

Contents

List of Figures

List of Tables

Abstract

Workplace stress has been shown to impact the quality of decision-making made by professionals in multiple domains, including medicine and policing. However, there remains a lack of research addressing the influence that workplace stress may have on the quality of forensic examiners' decision-making. Forensic experts work in pressurised environments, and they make important decisions that may affect legal outcomes. Hence, it is critical to understand stress and its potential negative or positive impact on forensic experts' decision-making.

This research sought to explore the extent to which the forensic experts experience stress in the workplace, and its potential impact on the decision-making that they make in the forensic science process. The first part of the thesis presents data on the possible sources of stress and pressures, from surveys of practicing forensic examiners (Study 1 and Study 2). The findings suggested that some experts experienced high stress from workplace factors, including implicit pressures from stakeholders that interact with the experts to report certain forensic conclusions over others. Experts varied in their opinion whether stress affected their decision-making.

The second part of the thesis explored the possible impact of stress on forensic science decision-making. An online method for testing the impact of stress on participants was developed, and it was found effective in inducing stress in human participants (Study 3). Then, the possible impact of this stressor on a fingerprinting comparison task with both novices and fingerprint experts was explored (Study 4). Findings revealed a complex relationship between stress and expert decision-making. Whilst stress improved the performance of both novices and experts on fingerprint assessments, stress also caused reduced risk-taking, especially when the fingerprints were difficult.

Three underlying themes emerged from this research: the importance of addressing common stress factors, the positive impact of stress on decision-making, and the complexity of evaluating stress and its impact. These themes form the basis of a new model outlining the 'Stressor-Stress Response in Expert Decision-Making' which may contribute to the development of stress-optimisation strategies to enhance forensic expert performance.

Impact Statement

The work undertaken during this research has potential impact in practice, academia, and public policy. In terms of practice, the thesis has put forward stress-optimisation strategies that could be implemented in the forensic science workplace to enhance expert decision-making (for example, emotional intelligence training to supervisors/managers of forensic experts). These were based on insights from current research on the possible sources of stress and its impact on forensic science judgments.

In terms of academia, this thesis has impact on research in the forensic science discipline, and beyond. Research undertaken here differs from other stress research in forensic science in two main aspects. First, discussion and data focused on the forensic expert decision-making, as opposed to solely well-being as in most of the other research. Second, the experimental approach to study the relationship between stress and decision-making performance was a first in forensic science. This focused understanding on the decision-making of experts is critical because it is the forensic experts themselves who make decisions that are relied upon by other stakeholders (for example, the police for investigations).

The online stress method that was developed in this thesis is the first that was effectively deployed to induce stress in human subjects online, and without the presence of researchers as stress agents (either in person or virtually). This method may contribute to advancements in stress research, beyond forensic science, amidst the increasing recognition of the value of online research after the adjustments made to research practices during the restrictions from the Covid-19 pandemic. Such advancements may include accessing a large sample of participants quickly and increasing the diversity of participants doing research. As part of the stress-inducing method development, a toolkit of measures was put together in order to enhance the quality of online research (for example, including a commitment statement, use of captcha to detect bot responses, etc). These measures contribute to enhancing the internal validity of any online research, not just stress research.

In terms of public policy, this thesis has potential to contribute to the considerations of the quality of life and well-being of professionals outside the forensic science domain (for example, insights on personal stressors outside the workplace and how they may carry through to influence the workplace environment). In addition, data from this thesis feeds into Sustainable Development Goal 3 of the United Nations on (i.e., "ensure healthy lives and promote well-being for all at all ages").

This research has been published in four international, peer-reviewed journals (a fifth paper is currently under review), as well as presented in international conferences and to forensic service providers, and presented through public engagement activities with lay persons in order to maximise dissemination to a range of disciplinary academic audiences.

UCL Research Paper Declaration Form

The UCL Research Paper Declaration Form has been completed five times to declare parts of the thesis that have been published or submitted for publication (i.e., Chapters 2-6).

Chapter 2

1. For a research manuscript that has already been published (if not yet published, please skip to section 2):		
a) Where was the work published? (e.g. journal name)	International Journal of Law, Crime and Justice	
b) Who published the work? (e.g. Elsevier/Oxford University Press):	Elsevier	
c) When was the work published?	01/12/2019	
d) Was the work subject to academic peer review?	Yes	
e) Have you retained the copyright for the work?	Yes	
[If no, please seek permission from the relevant publisher and check the box next to the below statement]: ☐	*I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.*	
2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):		
a) Has the manuscript been uploaded to a preprint server? (e.g. medRxiv):		If yes, which server? Click or tap here to enter text.
b) Where is the work intended to be published? (e.g. names of journals that you are planning to submit to)	Click or tap here to enter text.	

c) List the manuscript's authors in the intended authorship order:	Click or tap here to enter text.
d) Stage of publication	Please select.

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):
Mohammed A. Almazrouei (conceptualisation, write up); Itiel E. Dror (conceptualisation, write up); Ruth M. Morgan (conceptualisation, write up)
4. In which chapter(s) of your thesis can this material be found?
Chapter 2 (Section 2.1.3 on "Communication in Forensic Science")
5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work):

Candidate:	Mohammed Almazrouei	Date:	30/08/2022
Supervisor/ Senior Author (where appropriate):	Ruth M Morgan	Date:	08/09/2022

Chapter 3

1. For a research manuscript that has already been published (if not yet published, please skip to section 2):	
a) Where was the work published? (e.g. journal name)	Journal of Forensic Sciences
b) Who published the work? (e.g. Elsevier/Oxford University Press):	Wiley
c) When was the work published?	01/11/2020
d) Was the work subject to academic peer review?	Yes
e) Have you retained the copyright for the work?	Yes
[If no, please seek permission from the relevant publisher and check the box next to the below statement]:	
☐	*I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.*

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):		
a) Has the manuscript been uploaded to a preprint server? (e.g. medRxiv):		If yes, which server? Click or tap here to enter text.
b) Where is the work intended to be published? (e.g. names of journals that you are planning to submit to)	Click or tap here to enter text.	
c) List the manuscript's authors in the intended authorship order:	Click or tap here to enter text.	

d) Stage of publication	Please select.

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):
Mohammed A. Almazrouei (conceptualisation, data analysis, write up); Itiel E. Dror (conceptualisation, data collection, write up); Ruth M. Morgan (conceptualisation, write up)
4. In which chapter(s) of your thesis can this material be found?
Chapter 3
5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work):

Candidate:	Mohammed Almazrouei	Date:	30/08/2022
Supervisor/ Senior Author (where appropriate):	Ruth M Morgan	Date:	08/09/2022

Chapter 4

1. For a research manuscript that has already been published (if not yet published, please skip to section 2):	
a) Where was the work published? (e.g. journal name)	Forensic Science International: Mind and Law
b) Who published the work? (e.g. Elsevier/Oxford University Press):	Elsevier
c) When was the work published?	01/11/2021
d) Was the work subject to academic peer review?	Yes
e) Have you retained the copyright for the work?	Yes
[If no, please seek permission from the relevant publisher and check the box next to the below statement]: ☐ *I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.*	

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):		
a) Has the manuscript been uploaded to a preprint server? (e.g. medRxiv):		If yes, which server? Click or tap here to enter text.
b) Where is the work intended to be published? (e.g. names of journals that you are planning to submit to)	Click or tap here to enter text.	
c) List the manuscript's authors in the intended authorship order:	Click or tap here to enter text.	

d) Stage of publication	Please select.

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):
Mohammed A. Almazrouei (conceptualisation, data analysis, write up); Ruth M. Morgan (conceptualisation, write up); Itiel E. Dror (conceptualisation, data collection, write up)
4. In which chapter(s) of your thesis can this material be found?
Chapter 4
5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work):

Candidate:	Mohammed Almazrouei	Date:	30/08/2022
Supervisor/ Senior Author (where appropriate):	Ruth M Morgan	Date:	08/09/2022

Chapter 5

<table>
<tr><td colspan="3">1. For a research manuscript that has already been published (if not yet published, please skip to section 2):</td></tr>
<tr><td>a) Where was the work published? (e.g. journal name)</td><td colspan="2">Behavior Research Methods</td></tr>
<tr><td>b) Who published the work? (e.g. Elsevier/Oxford University Press):</td><td colspan="2">Springer</td></tr>
<tr><td>c) When was the work published?</td><td colspan="2">25/07/2022</td></tr>
<tr><td>d) Was the work subject to academic peer review?</td><td colspan="2">Yes</td></tr>
<tr><td>e) Have you retained the copyright for the work?</td><td colspan="2">Yes</td></tr>
<tr><td colspan="3">[If no, please seek permission from the relevant publisher and check the box next to the below statement]:

☐ I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.</td></tr>
<tr><td colspan="3">2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):</td></tr>
<tr><td>a) Has the manuscript been uploaded to a preprint server? (e.g. medRxiv):</td><td></td><td>If yes, which server? Click or tap here to enter text.</td></tr>
<tr><td>b) Where is the work intended to be published? (e.g. names of journals that you are planning to submit to)</td><td colspan="2">Click or tap here to enter text.</td></tr>
<tr><td>c) List the manuscript's authors in the intended authorship order:</td><td colspan="2">Click or tap here to enter text.</td></tr>
</table>

d) Stage of publication		Please select.	
3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):			
Mohammed A. Almazrouei (conceptualisation, data collection, data analysis, write up); Ruth M. Morgan (conceptualisation, write up); Itiel E. Dror (conceptualisation, write up)			
4. In which chapter(s) of your thesis can this material be found?			
Chapter 5			
5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work):			
Candidate:	Mohammed Almazrouei	Date:	30/08/2022
Supervisor/ Senior Author (where appropriate):	Ruth M Morgan	Date:	08/09/2022

Chapter 6

1. For a research manuscript that has already been published (if not yet published, please skip to section 2):		
a) Where was the work published? (e.g. journal name)	Click or tap here to enter text.	
b) Who published the work? (e.g. Elsevier/Oxford University Press):	Click or tap here to enter text.	
c) When was the work published?	Click or tap to enter a date.	
d) Was the work subject to academic peer review?		
e) Have you retained the copyright for the work?		
[If no, please seek permission from the relevant publisher and check the box next to the below statement]: ☐ *I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.*		
2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):		
a) Has the manuscript been uploaded to a preprint server? (e.g. medRxiv):	No	If yes, which server? Click or tap here to enter text.
b) Where is the work intended to be published? (e.g. names of journals that you are planning to submit to)	Forensic Science International	
c) List the manuscript's authors in the intended authorship order:	Mohammed A. Almazrouei; Itiel E. Dror; Ruth M. Morgan	

d) Stage of publication		Submitted	
3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):			
Mohammed A. Almazrouei (conceptualisation, data collection, data analysis, write up); Itiel E. Dror (conceptualisation, write up); Ruth M. Morgan (conceptualisation, write up);			
4. In which chapter(s) of your thesis can this material be found?			
Chapter 6			
5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work):			
Candidate:	Mohammed Almazrouei	Date:	30/08/2022
Supervisor/ Senior Author (where appropriate):	Ruth M Morgan	Date:	08/09/2022

Chapter 1 Introduction

1.1 Overview

Workplace stress has been shown to impact the quality of decision-making made by professionals in a variety of domains, including medicine (e.g., Arora et al., 2010), policing (e.g., Akinola & Mendes, 2012), the military (e.g., Kavanagh, 2005), management (e.g., Gok & Atsan, 2016), and psychology (e.g., Dror et al., 1999; Kerstholt, 1994; Yu et al., 2015). However, research regarding the impact of workplace stress upon forensic expert decision-making is scarce (Jeanguenat & Dror, 2018). Workplace environments of forensic experts have been characterised as uncertain (information could be ambiguous or contradictory), time pressured (Helsloot & Groenendaal, 2011), and significantly consequential for investigation outcomes (Smit et al., 2018). Forensic examiners face multiple sources of workplace stress (Jeanguenat & Dror, 2018; Kelty & Gordon, 2015). Some stress factors are common across other occupations, such as high workload, and some are specific to forensic science contexts, such as exposure to bloody crime scenes (Jeanguenat & Dror, 2018; National Institute of Justice, 2019)

Undertaking research to understand and evaluate the possible sources of workplace stress and their impact on the decision-making of forensic examiners is important. This is because of the critical decisions that examiners need to make in casework to reach conclusions (Dror & Stoel, 2014), and communicate insights and evaluative interpretations to stakeholders (Almazrouei et al., 2019; Dror & Pierce, 2020). These decisions are undertaken in complex and often ambiguous contexts, and are impacted by different factors (Dror, 2020a; Dror & Cole, 2010) to produce insights that assist intelligence-led policing and/or legal proceedings (Morgan, 2017b, 2017a; Morgan et al., 2018).

The acknowledgement of subjectivity in decision-making in forensic science (Forensic Science Regulator, 2015; House of Lords Science and Technology Select Committee, 2019; National Academy of Sciences, 2009; National Commision on Forensic Science, 2015) and instances where miscarriages of justice were identified (Office of the Inspector General, 2006; The Fingerprint Inquiry, 2011) caused a new focus in forensic science research. Specifically, research efforts investigating factors influencing decision-making in forensic science (see recent reviews in Cooper & Meterko (2019), and Kukucka & Dror (2022)), including workplace stress factors (Jeanguenat & Dror, 2018; National Institute of Justice, 2019) noticeably increased.

1.2 Aims and Research Questions

Therefore, the main aim of this PhD thesis was to explore stress, as a human factor in forensic science contexts, in order to identify the characteristics of workplace environments that enable optimal decision-making. To address this aim, there were two main objectives:

a) to obtain empirical data to better identify the sources of workplace stress experienced by forensic experts, and
b) to undertake experiments to investigate the impact of stress on judgments that are made at critical stages in the forensic science process.

To address the aim and objectives, this thesis focussed on two main research questions:

1. What degree of workplace stress (and feedback) do forensic experts experience?
2. Does stress impact conclusions that are reached by forensic experts? Specifically, does stress influence conclusions reached by fingerprint experts, and does this influence differ between experts and novices?

1.2.1 Research Question 1

Addressing research question 1 contributes data to better understand the possible sources of workplace stress as experienced by forensic examiners who perform casework (Holt et al., 2017; Jeanguenat & Dror, 2018). This may have implications for developing relevant evidence-based approaches to improve the wellbeing of experts as well as their decision-making performance. In addition, considering the context within which decisions are being made (such as stress from explicit or implicit feedback at work) may help ensuring there is transparency in this process to mitigate conditions that exert pressure on examiners to make 'expected' decisions.

To address research question 1, forensic experts were asked to complete two questionnaires about their stress and its sources (Chapters 3 and 4). In addition, the first questionnaire included questions on the implicit and explicit feedback experts receive from their interactions with key stakeholders, such as their managers/supervisors or the police. The follow-up questionnaire included questions that investigated additional sources of workplace stress (specifically, stress from the nature of the case and working in high profile cases, and stress from the circumstances at the workplace), whether examiners receive support from management, and whether examiners believed that the stress they experienced affected their judgments.

1.2.2 Research Question 2

Investigating research question 2 aims to fill a different gap in the forensic science literature. There is virtually no experimental, peer-reviewed research that has studied the impact of stress on decisions being made in forensic science contexts (Jeanguenat & Dror, 2018). An experimental approach can help understanding the relationship between stress and expert decision-making. Hence, experimental data could offer a complementary picture to the self-reported data in research question 1. Such a holistic understanding may help devise appropriate strategies for optimal working environments in forensic science.

To address research question 2, a series of online experiments were conducted (Chapters 5 and 6). A stress-inducing method was developed to stress participants online, without the presence of researchers. This stress method was adapted from the Trier Mental Challenge Test stress protocol (Kirschbaum et al., 1991). After collecting data testing this method, two online experiments were conducted to investigate the impact of stress on fingerprint assessments made by novices (Experiment One) and experts (Experiment Two). Understanding whether stress influences decisions during fingerprint analysis is critical, as fingerprints are highly used and carries significant weight in court decision-making (Mustonen et al., 2015), and insights from these studies may potentially be applicable to other comparative forensic science fields, like handwriting, toolmarks, and other expert domains.

1.3 Structure of the Thesis

This thesis presents a literature review (Chapter 2), two chapters that present the findings from two surveys (Chapters 3 and 4), two chapters that present experimental studies (Chapters 5 and 6), a discussion (Chapter 7) and a conclusion (Chapter 8).

Chapter 2 provides an overview of the pertinent published literature. This thesis is an interdisciplinary research project and therefore this chapter brings together three key research areas: forensic science decision-making, workplace stress, and fingerprinting. The literature review brings together theories, empirical research and casework examples of stress and decision-making. Because of the lack of published research addressing workplace stress in forensic science, some inferences and examples are drawn from other specialised domains, such as medicine and policing. This chapter is structured around three main topics, starting with an overview of decision-making in forensic science. Here, the importance of forensic expert decision-making, and communication with the forensic examiners, in the forensic science process is presented. Concerns about subjectivity of decisions in forensic science, and cognitive factors that may influence forensic experts who make the actual casework, are addressed and explored. The second part focuses on stress, its characteristics and how it relates to expert performance. In addition, previous research on the sources of stress in forensic science are presented. Third, fingerprint analysis is introduced as a specific pattern recognition domain that formed the basis of the experimental work (research question 2). This part discusses subjectivity in fingerprint assessments and how conclusions are reached. This chapter concludes with an articulation of the current gaps that exist in the knowledge regarding the possible sources and impacts of stress on forensic expert decision-making.

Chapter 3 reveals the extent to which forensic examiners experience workplace stress and feedback, to address objective (a) of the thesis. It presents a novel study that investigates the possible sources of workplace stress (and feedback) experienced by forensic examiners from various fields. Since the data from this questionnaire was collected from one large forensic laboratory, it minimises inter-laboratory differences in the working cultures which could have confounded the findings of previous published studies. The study also involves an in-depth analysis of workplace stress and feedback, as they vary by forensic science fields and years of experience, to better understand the moderating factors of workplace stress and feedback. The findings suggested that forensic experts felt their high levels of stress originated more from the workplace (e.g., stress from backlogs) than from personal reasons (e.g., financial issues). The data also revealed a concerning finding in that a few experts, from the same lab, strongly felt high implicit pressures in reporting forensic science conclusions.

Chapter 4 presents the findings of a follow-up questionnaire on the sources of workplace stress, and it included crime scene examiners in addition to laboratory-based examiners (i.e., this chapter further addresses objective (a)). As in the first questionnaire, this study involved a detailed analysis of workplace stress as it varies by the forensic science fields and years of experience, with sex considered as an additional factor. In addition, this study sought to identify whether examiners felt that they receive support from management and whether they felt that stress affected their decision-making. The findings demonstrated that stress from managers and/or supervisors and from case backlogs were the only stress factors, among the other stressors explored in this study, that predicted workplace stress. Examiners felt that they receive moderate support from their management, and this level of support was not associated with workplace stress. Demographics of experts played a role in their reported stress (e.g., crime scene examiners reported higher stress levels than analytical examiners due to the nature of cases they are involved in). The participants were divided on whether they felt that stress affected their judgments, which supported the need to explore this question in an experimental approach (i.e., objective (b)).

Chapter 5 presents a new method that was needed to address objective (b) of the thesis. Specifically, this method was developed to cause stress to human subjects online, such as forensic experts, without the presence of researchers (either in person or virtually). In this method, participants were asked to answer general knowledge and mathematical questions selected specifically for this study under either a control condition (without feedback and no time limits) or a stress condition (with feedback and with time pressure). This way, stress conditions of social evaluative threats (when one is judged negatively by others, such as displaying negative feedback) and uncontrollability (when nothing can be done to avoid negative consequences or change a situation, such as imposing time pressure) are induced (Dickerson & Kemeny, 2004). The online stress method was found to be effective, and therefore has potential to enable advancements in stress research, amidst the growing recognition of online research (Kirschbaum, 2021). Indeed, the method developed in this thesis has been published in the high impact specialised journal, *Behavior Research Methods* (Almazrouei et al., 2022)

Chapter 6 presents insights from a consideration of the impact of the online stressor (developed in Chapter 5) on forensic science decision-making. Fingerprint assessment was used for this study because it is frequently utilised in casework (Mustonen et al., 2015). Insights on the influence of stress may apply to other pattern recognition tasks (e.g., toolmarks comparison). This chapter presents the influence of stress on conclusions, confidence levels and response times of novices (Experiment One) and fingerprint experts (Experiment Two). Overall, the findings suggested that stress improved the performance of both novices and experts on fingerprint assessments, but mainly for same-source evidence. Moreover, the stressor had an impact on the confidence levels and response times of novices, but not experts. These findings offer a starting point to better understand the possible impact of stress on decision-making in forensic science contexts, thus further contributing to objective (b) of the thesis.

Chapter 7 presents a synthesised overview of key themes that have emerged from Chapters 2-6. Three themes were identified: Theme 1 addressing common stress factors is a priority; Theme 2 the positive impact of stress on decision-making; and Theme 3 evaluating stress and its impact is not simple or straightforward. These themes contribute to the development of a model of stressors, stress responses and expert decision-making, which is presented in this chapter. Finally, limitations of the empirical studies in the thesis, and a consideration of potentially fruitful areas of future research on stress and stress-optimisation strategies in practice are presented.

Chapter 8 provides the conclusions that have been derived from the work presented in this thesis that address the possible sources and impact of stress on decision-making in forensic science. It also offers the implications of these findings both in forensic science practice and research.

1.4 Additional Considerations

Parts of the literature review (see Section 2.1.3) have been published in an article that sets out a theoretical foundation for the empirical chapters. In addition, the empirical chapters have either been published or submitted for publication (4 articles; see Table 1-1).

Table 1-1: List of publications.

PhD thesis chapters	**Publication status**	**Details of publication**
Chapter 2	Published	Almazrouei, M. A., Dror, I. E., & Morgan, R. M. (2019). The forensic disclosure model: What should be disclosed to, and by, forensic experts? *International Journal of Law, Crime and Justice*, 59, 100330. https://doi.org/10.1016/j.ijlcj.2019.05.003
Chapter 3	Published	Almazrouei, M. A., Dror, I. E., & Morgan, R. M. (2020). Organizational and human factors affecting forensic decision-making: Workplace stress and feedback. *Journal of Forensic Sciences*, 1968–1977. https://doi.org/10.1111/1556-4029.14542
Chapter 4	Published	Almazrouei, M. A., Morgan, R. M., & Dror, I. E. (2021). Stress and support in the workplace: The perspective of forensic examiners. *Forensic Science International: Mind and Law*, 2, 100059. https://doi.org/10.1016/j.fsiml.2021.100059
Chapter 5	Published	Almazrouei, M. A., Morgan, R. M., & Dror, I. E. (2022). A method to induce stress in online research environments. *Behavior Research Methods*. https://doi.org/10.3758/s13428-022-01915-3
Chapter 6	Submitted	Almazrouei, M. A., Dror, I. E., & Morgan, R. M. The impact of stress on forensic decision-making.

Chapter 2 Literature Review

2.1 Decision-Making in Forensic Science

2.1.1 The Role of Expert Decision-Making in Forensic Science

The decision-making of forensic experts is an integral part throughout the forensic process (see Figure 2-1). Forensic science decisions can include forming a strategy for crime scene examination, prioritising exhibits for laboratory analysis, interpreting evidence, and communicating expert witness opinions to fact-finders (Morgan, 2017a; Morgan et al., 2018; Morgan & Bull, 2007; Roux et al., 2012). Critically, decisions made early on at the crime scene can affect subsequent decisions, which may influence the police investigation and judicial outcome of a case (Dror, 2015; Earwaker et al., 2020; Morgan et al., 2018; Nakhaeizadeh et al., 2017).

Therefore, for accurate, reproducible, and transparent crime reconstructions to take place, an integral consideration of human decision-making needs to be made at every stage of the forensic science process (i.e., component 4 of the Forensic Reconstruction of Trace Evidence conceptual model (Morgan, 2017a; Morgan et al., 2020). Even with technological advances in forensic analysis, it is still the human forensic experts who act as the "main instrument" in the forensic analysis and comparison. Specifically, the experts are involved in many of the forensic science tasks, including pattern recognition tasks (Dror & Stoel, 2014) of visually comparing patterns to decide whether they come from the same source (e.g., whether two bullets have "sufficient similarity" to have been fired from the same gun (Dror & Stoel, 2014) or assessing the peaks of a mixture DNA profile (Jeanguenat et al., 2017)).

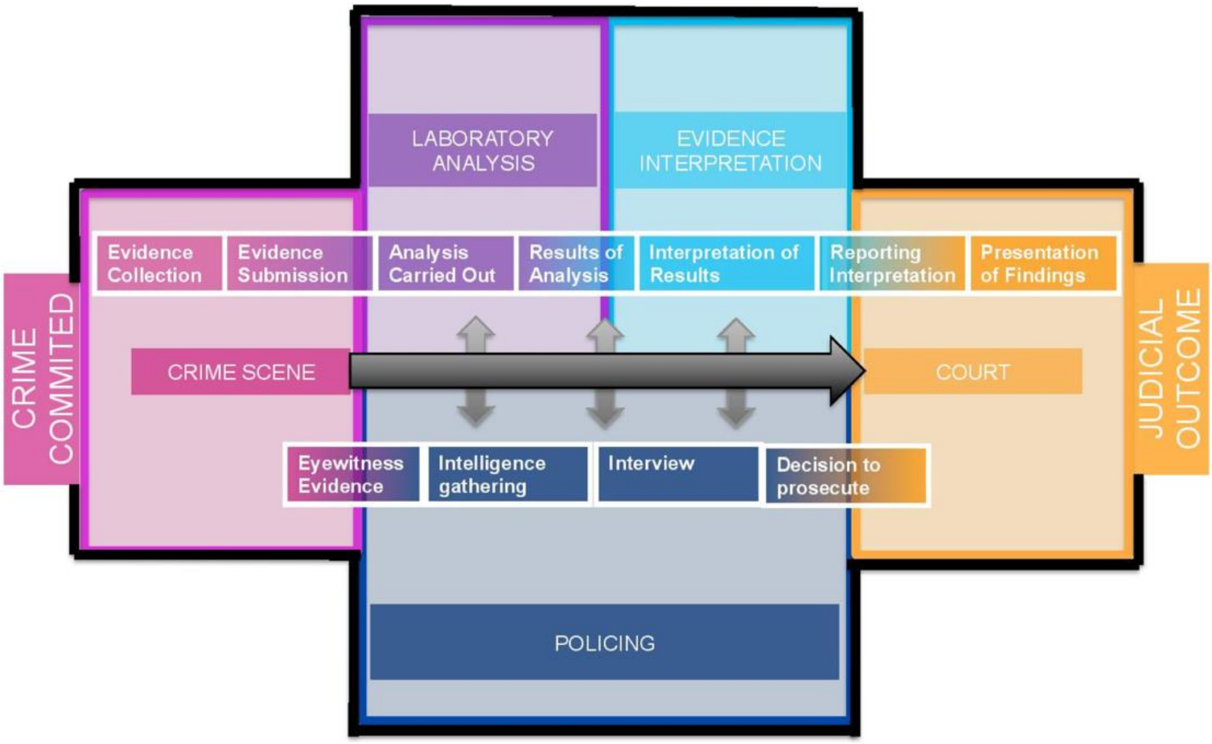

Figure 2-1**:** Decision-making in forensic science from crime scene to court (Morgan et al., 2018).

Given the critical role of human decision-making in forensic science, it is important to offer clear, transparent and reproducible judgments (National Academy of Sciences, 2009; Saks & Koehler, 2005). However, it is well established in a wide range of disciplines that decision-making can be considered t0 be subjective given the incorporation of both explicit and tacit forms of knowledge (Morgan, 2017b, 2017b). All human decision-making is influenced by different factors (Dror & Cole, 2010; Saks et al., 2003). Forensic experts are no exception; they too are impacted by these factors (Cooper & Meterko, 2019). In addition, their working environment contains elements of time pressure, operating in a context of 'high stakes' (Helsloot & Groenendaal, 2011) and uncertainty (see Principle 5 of the Sydney Declaration (Roux et al., 2022)). Therefore, there have been increasing concerns raised about the subjectivity of decision-making in forensic science (Kassin et al., 2013), and how to increase the transparency of how decisions are made and conclusions reached at each stage of the forensic science process (Almazrouei, 2020; Heavey et al., 2022).

2.1.2 Concerns About Decision-Making in Forensic Science

Examples of miscarriages of justice and reports from multiple national bodies worldwide have revealed repeatedly that evaluative conclusions can be subjective and susceptible to different human factors. For example, hundreds of miscarriages of justice have been identified in the United States through the Innocence Project and its work into cases where DNA has led to exonerations (Innocence Project, 2022). It was reported that 52% of wrongful convictions involved the 'misapplication of forensic science' which includes "convictions based on forensic evidence that is unreliable or invalid and expert testimony that is misleading" along with other factors (Innocence Project, 2022). In England and Wales, Smit et al. (2018) identified 235 cases (from January 2010 to December 2016) that were upheld at the Court of Appeal that contained misleading criminal evidence, including forensic science evidence. It was reported that the weight of forensic science evidence (e.g., its relevance and prohibitive value) was miscommunicated to, or misunderstood by, the triers of fact, which led to wrongful convictions (Smit et al., 2018).

An infamous case of evidence that was misinterpreted was that of Brandon Mayfield (Kassin et al., 2013; Office of the Inspector General, 2006; Stacey, 2004). A number of FBI fingerprint experts, as well as a fingerprint expert hired by the defence, all erroneously matched Mayfield's fingerprint to the fingermark recovered from the scene of the 2004 Madrid train bombings. Following an independent investigation by the Office of the Inspector General, cognitive bias was deemed a contributing factor to the erroneous identification (Office of the Inspector General, 2006). Several cognitive and psychological issues were identified in the case. For example, the fingerprint examiners were exposed to contextual information that led them to target Mayfield as the suspect (Kassin et al., 2013). The case also involved time pressure, being a high-profile case, and increasing the need for closure (i.e., the desire to provide clear-cut judgments (Ask & Granhag, 2005)) in this case by positively identifying Mayfield as opposed to reporting an exclusion or an inconclusive decision (Kassin et al., 2013).

Discussions of cognitive bias in forensic science have intensified since the publication of the National Academy of Sciences report (2009; see also Found, 2015). The report identified problems with many forensic science fields, such as toolmarks and firearms, bloodstain pattern analysis, handwriting, and even fingerprints (which had been regarded infallible until recently (Mnookin, 2008)). The NAS (2009) report stated that there were issues with reliability, accuracy, and the potential for cognitive bias in forensic science. The report suggested that the forensic science fields:

> "need to develop rigorous protocols to guide these subjective interpretations and pursue equally rigorous research and evaluation programs. The development of such research programs can benefit significantly from other areas, notably from the large body of research on the evaluation of observer performance in diagnostic medicine and from the findings of cognitive psychology on the potential for bias and error in human observers." (NAS report, 2009, p. 8).

Other national reports followed, which further emphasised the issues of subjectivity and bias in decision-making in forensic science. These include reports from the United States (e.g., Executive Office of the President's Council of Advisors on Science and Technology, 2016; National Commision on Forensic Science, 2015) and from the United Kingdom (e.g., Forensic Science Regulator, 2015; Government Chief Scientific Adviser, 2015; House of Lords Science and Technology Select Committee, 2019; The Fingerprint Inquiry, 2011).

2.1.3 Communication in Forensic Science

During casework, forensic examiners communicate and receive feedback from a variety of stakeholders that can be classified into forensic services, investigative services, legal and external stakeholders categories (see Figure 2-2). For example, within the forensic services domain, forensic examiners communicate with management and/or supervisors for various reasons, such as to reach resolutions in the conclusions reached (Mustonen et al., 2015). It is worth noting that the external stakeholders who communicate with the forensic experts can be further divided into regulatory, such as government authorities, and public domains, such as the media (Dror & Pierce, 2020).

Communication and information-sharing between the forensic services domain and investigative domain is important for providing forensic intelligence (i.e., providing accurate and timely forensic data that can assist police decision-making for investigation and intelligence purposes (Raymond & Julian, 2015; Roux et al., 2012)) (see Figure 2-1). Therefore, effective forensic intelligence requires feedback and collaboration between the forensic experts and other stakeholders (such as the police (Raymond & Julian, 2015)), in a more holistic approach to minimise uncoordinated efforts (Roux et al., 2021). Nonetheless, the communication to the forensic experts should be fit-for-purpose and limited to task-relevant information (Dror et al., 2017; Gardner et al., 2019). Otherwise, cross-contamination can occur when irrelevant and biasing information cascades from one stage of the forensic science process to another (i.e., a "bias cascade effect" (Dror, 2018)). In addition, a "bias snowball effect" can occur when a variety of task-irrelevant information is integrated to form an expert opinion (Dror, 2018), such as viewing horrific photographs of the victims before bitemark assessments (Osborne et al., 2014).

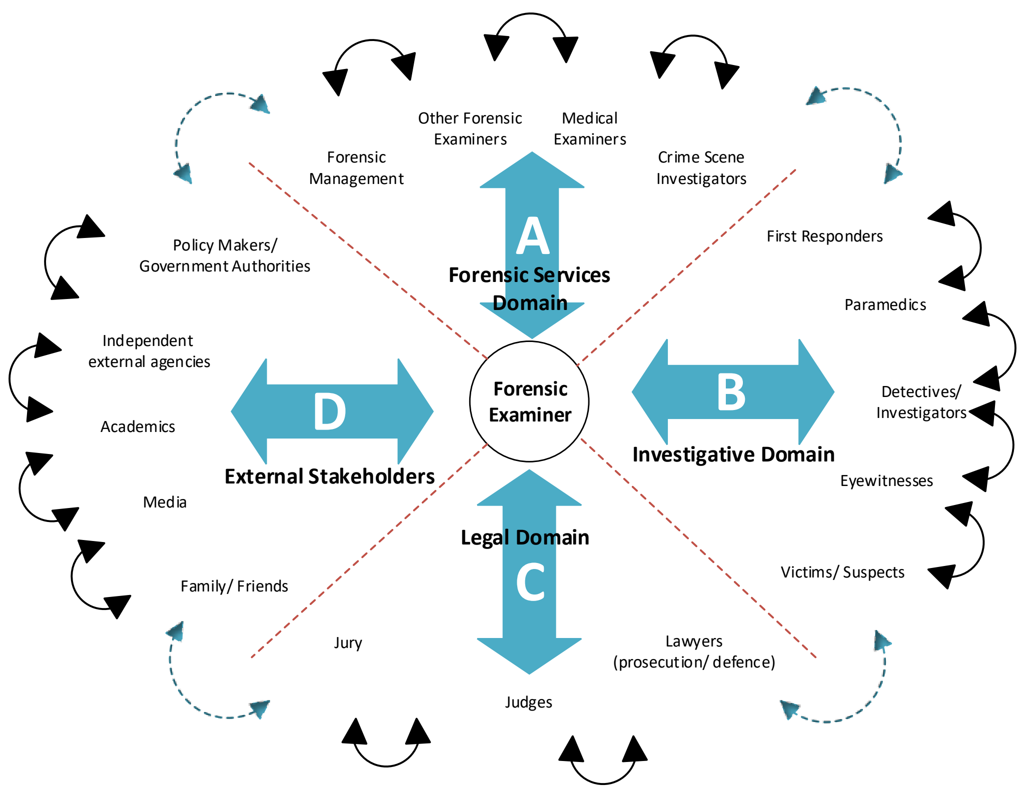

Figure 2-2: Interactions and communications of forensic examiners with stakeholders (Almazrouei et al., 2019).

Forensic examiners communicate with legal stakeholders to provide scientific evidence in the form of written reports and/or verbal testimonies (Arscott et al., 2017; Dror et al., 2015; Howes, 2015). Testimonies of forensic experts possess significant weight in the criminal justice system because experts typically present their opinions as if they are impartial and scientific, and the courts view them as such (Dror et al., 2015). However, the experts may be overconfident and overstate their opinions on the examined evidence, due to the inability of humans to evaluate their actual knowledge and abilities (Kukucka et al., 2017; Page et al., 2012).

Another issue in adversarial legal systems stemming from the communication of forensic examiners with legal stakeholders is the possibility of "adversarial allegiance." As the experts are normally recruited by either the defence or the prosecution to give an expert opinion, the decision-making of the experts can be biased towards the party that retained them (Murrie et al., 2009, 2013).

In the United States, Federal Rule of Evidence 702 governs the admissibility of forensic expert testimony and evidence (Federal Rules of Evidence, 2020; Risinger et al., 2002)—whilst recognising that different states have different standards of admissibility for expert testimony (Lesciotto, 2015). Federal Rule 702 accounts for key relevant cases such as *Daubert v. Merrell Dow Pharmaceuticals, Inc.* and *Kumho Tire Co. v. Carmichael* (Federal Rules of Evidence, 2020; Risinger et al., 2002). It requires that the testimony of the expert be "the product of reliable principles and methods" and that "the expert has reliably applied the principles and methods to the facts of the case" (Federal Rules of Evidence, 2020, para. 1). Similarly, in the UK, Criminal Procedure Rule number 33.2(a) explicitly states that the expert has a duty to the court to provide "objective and unbiased" expert opinions (Criminal Procedure Rules, 2014, p. 1).

Researchers have argued that these requirements would address the issue of cognitive bias in the forensic science evidence presented to court (Dror et al., 2015; Risinger et al., 2002). Yet, in the United States for example, it would be rare to exclude forensic science evidence under Federal Rule 702 (Risinger et al., 2002). Therefore, a potential practical approach would be to control and be transparent about what is disclosed to, and by, the forensic examiners (Almazrouei et al., 2019), particularly given that the risk of cognitive bias is considered unavoidable and inherent in the methodology (Dror & Cole, 2010).

2.1.4 Cognition of Decision-Making in Forensic Science

People in daily life as well as professional work process information. Humans interact with incoming information, referred to as "bottom-up" information (Dror & Stoel, 2014). The bottom up information is processed by the brain and cognitive system process to reach a decision (Dror & Bucht, 2012; Kassin et al., 2013). In the forensic science context, bottom-up information is essentially the data generated from traces, such as bloodstain patterns (Taylor et al., 2016a). The processing of these bottom-up data is dynamic and sequential in nature (Dror & Bucht, 2012; Dror & Langenburg, 2019). When forensic experts make decisions, not all the information or the data from the trace may be examined before a conclusion is made. Instead, each piece of information may be assessed sequentially until a threshold for making a conclusion is reached (even before examining all the available data from the evidence). For instance, a fingerprint comparison involves sequentially examining a latent fingermark and a reference fingerprint and accumulating data until a sufficient "similarity" is reached to report an identification, or "dissimilarity" is reached for an exclusion (Dror & Langenburg, 2019). This pattern of decision-making takes place across many domains (see Decision Field Theory (Busemeyer & Townsend, 1992, 1993)).

The interpretation of bottom-up information is mediated by top-down mechanisms, beyond the actual data, such as knowledge, experience, motivations (Kassin et al., 2013) and stress (Jeanguenat & Dror, 2018). Top-down processing is critical because the human brain has limited capacity to process all the bottom-up information. For instance, selective attention enables the brain to select and process only some bottom-up information while other information is ignored (Kahneman, 2003; Saks et al., 2003). With more experience and knowledge, humans develop these top-down processes and become experts. So, expertise, including forensic expertise, entails well-developed top-down cognitive mechanisms that result in improved performance compared with that of laypersons (Busey & Dror, 2011). With more experience and training, the tasks that initially required effort and had a high cognitive load on the forensic experts are done faster (Kellman et al., 2014). This way, experts develop heuristics (mental shortcuts) to judge part of the information and ignore other parts. Hence, information processing becomes more automatic and efficient (Gigerenzer & Gaissmaier, 2011).

However, these top-down mechanisms can also result in vulnerabilities (Busey & Dror, 2011). A potential vulnerability is "cognitive bias", which is "the class of effects through which an individual's preexisting beliefs, expectations, motives, and situational context influence the collection, perception, or interpretation of evidence, or their resulting judgments, decisions, or confidence" (Spellman et al., 2022, p.5; this definition is a modified version of the one outlined in Kassin et al. (2013)). For example, once the forensic experts develop an expectation, "tunnel vision" results (Spellman et al., 2022). In this situation, the experts are searching for certain information, ignoring other information, and ultimately the forensic science evidence can be misinterpreted (Kerstholt et al., 2010). Thus, heuristic thinking is useful to make efficient decisions, but can cause systematic errors and biased judgments (Kassin et al., 2013). It is critical to emphasise that these vulnerabilities are unintentional because the experts are not aware of their occurrence (as opposed to the intentional and unethical decisions to bias forensic conclusions (Dror & Cole, 2010)).

2.2 Factors Impacting Decision-Making in Forensic Science

2.2.1 An Eight-Level Taxonomy

Dror (2020) devised a taxonomy that outlines biasing factors that can affect decision-making in forensic science throughout the forensic science process (from the crime scene to court; see Figure 2-3). These factors could affect the observations (e.g., observing peaks in a DNA mixture), the conclusions (e.g., matching suspect and reference mixture DNA profiles), or both, without forensic examiners necessarily being aware of their impact (Dror & Hampikian, 2011; Jeanguenat et al., 2017; see also Dror, 2016 for the Hierarchy of Expert Performance Framework).

The structure of the taxonomy starts with the factors specific to the case being investigated (e.g., the contextual information effect of knowing about the DNA evidence when making fingerprint decisions (Stevenage & Bennett, 2017; see Category A in Figure 2-3). Then, moving down the taxonomy, the factors are specific to the examiners doing the work (Category B). Such factors include their experiences, their working environment, and the culture they work in. At the very bottom, the factors relate to human nature that we are all subject to (Category C). This taxonomy is helpful as it unpacks the different sources that may bias expert judgments (Dror, 2020a), and it has been demonstrated to be useful in recognising specific biasing factors in multiple forensic science fields, such as forensic evaluations (Zapf & Dror, 2017) and digital forensic examinations (Sunde & Dror, 2019).

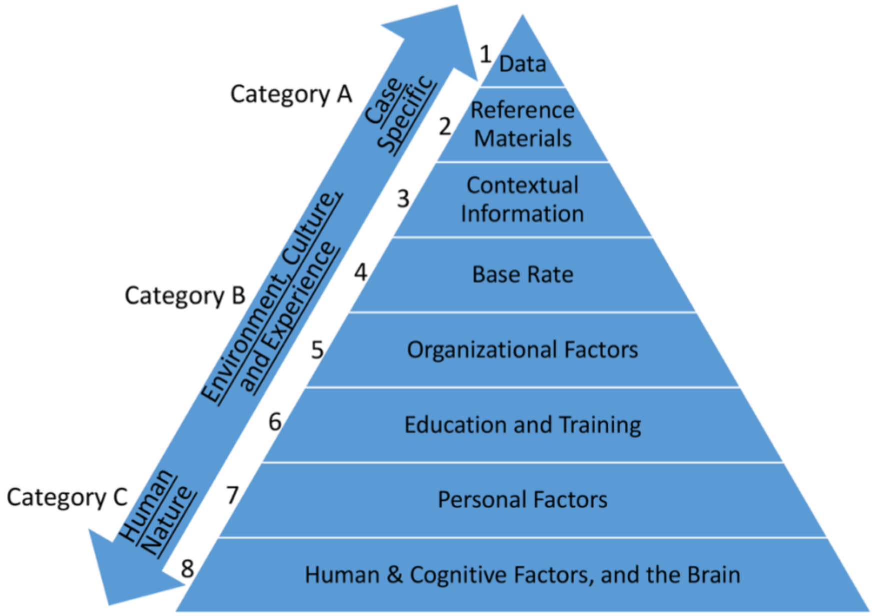

Figure 2-3: Eight factors of bias that may affect expert decision-making (Dror, 2020).

2.2.2 Contextual Bias

Context can create situations which affect judgements in forensic science (Saks et al., 2003). For example, “reference materials” may affect decision-making if a single suspect reference is presented instead of multiple reference samples as fillers (Kassin et al., 2013). This form of presentation can result in the formation of pre-existing expectations, whereby the examiner seeks information that supports their beliefs (Ask & Granhag, 2005; Kassin et al., 2013; Kukucka & Kassin, 2014), hence, creating a confirmation bias (Kassin et al., 2013; see Section 2.1.4).

The effect of context on decision-making in forensic science has now been extensively studied (Cooper & Meterko, 2019). The first empirical study on contextual effects dates back to 1984 (Found, 2015). In this study, it was found that contextual evidence, which created the belief that the suspect is guilty, impacted judgments of forged signatures (Miller, 1984). Since then, various forensic fields have exposed such potential biases — from bloodstain pattern analysis (Taylor et al., 2016a, 2016b) and bitemark analysis (Osborne et al., 2014) to more established fields, such as forensic toxicology (Hamnett & Jack, 2019) and DNA analysis (Dror & Hampikian, 2011).

Additionally, various types of biasing contextual information have been studied, such as confessions of guilt (Kukucka & Kassin, 2014) and the nature of the case in question (e.g., van den Eeden et al., 2018). In van den Eeden et al. (2018), contextual information (indicating a suicide, violent death, or no context) affected the first impression of the participants on the scene and the number of traces they collected. Not only has the type of context been extensively studied, but context has also been shown to influence and cascade between the different stages of the forensic science process. For instance, Nakhaeizadeh et al. (2017) found that exposure to contextual information (specifically, the subtle context of female in comparison to neutral clothing on the skeletal remains) influenced subsequent decisions and the interpretation of the sex assessment of the skeletal remains in the laboratory.

Despite the aforementioned studies, research on the effect of context on the decision-making in forensic science remains a critical issue for further research so as to minimise cognitive bias (Kassin et al., 2013). This is particularly important, considering the resistance of some forensic examiners to the fact that experts are susceptible to cognitive bias (Kukucka et al., 2017; Oliver, 2018; Page et al., 2012), and the presence of concerns from some researchers on the paradigm of cognitive bias research in forensic science (Champod, 2014; Curley et al., 2020).

Furthermore, more empirical research on the impact of contextual factors on forensic science decisions is needed because of the presence of conflicting empirical findings (e.g., null findings (Kerstholt et al., 2007, 2010)) or merely the methodological flaws in some of the previous studies (Cooper & Meterko, 2019). Yet, as will be illustrated in Sections 2.2.3 and 2.3, it is arguably equally important to expand the research efforts onto other human factors, such as emotions (Dror et al., 2005; Hall & Player, 2008; Osborne et al., 2014) and workplace stress (Jeanguenat & Dror, 2018).

2.2.3 Motivational and Emotional Bias

Several studies have investigated the motivational and emotional factors that could influence decision-making in forensic science (Charlton et al., 2010; Dror et al., 2005; Hall & Player, 2008; Osborne et al., 2014; Osborne & Zajac, 2015). For instance, one study that employed semi-structured interviews with 13 fingerprint examiners illustrated that the examiners expressed personal interests in solving crimes and catching criminals, particularly for serious crimes (Charlton et al., 2010). These motivations could be enhanced by the need for closure (Ask & Granhag, 2005; Charlton et al., 2010). The consequences of the need for closure on cognition are critical. For instance, people tend to “freeze” their thinking when they are motivated to achieve closure, and they become reluctant to consider alternative solutions (Ask & Granhag, 2005). It has been asserted that in criminal investigations, the usual working hypothesis is that the suspect is guilty, and the police investigators are motivated to look for evidence that confirms this hypothesis (Ask & Granhag, 2005).

In addition to motivational factors, forensic experts can operate under emotionally charged contexts (Osborne et al., 2014). Emotional factors can influence forensic science decisions (Dror et al., 2005; Osborne et al., 2014; Osborne & Zajac, 2015). For instance, Dror et al. (2005) found that emotional case details (e.g., horrific crime scene photographs) increased the ‘match’ decisions of fingerprints. This study, however, did not allow for inconclusive decisions to be made, which does not replicate real-world fingerprint casework (Dror et al., 2005). Ten years later, using a similar design, Osborne and Zajac (2015) included an inconclusive decision (specifically, “unsure”) as an option. In this study, the emotional context did not affect the match decisions. However, participants who received the emotional information reported fewer non-match fingerprint decisions compared with the participants who received neutral context (Osborne & Zajac, 2015).

Other researchers found that emotional context does *not* influence fingerprint decision-making (Hall & Player, 2008). In this study, experienced fingerprint experts were asked to judge pairs of fingerprints either in the context of forgery (low emotional context) or murder (high emotional context). Although experts self-reported that their analysis was affected by the context, the conclusions did not differ between the high and low emotional contexts (Hall & Player, 2008). It has been debated in the published literature that this null finding could have resulted from various methodological factors (such as weak emotional manipulation or the experts knowing they were taking part in this study, as opposed to being a true null finding (Dror, 2009; Kassin et al., 2013; Saks, 2009)).

There is a consensus among researchers that emotions and stress are closely related (Du et al., 2018). For instance, the cognitive appraisal of stress can generate negative emotions, such as anger and anxiety (Du et al., 2018). However, empirical studies have been able to identify the impact of factors focused on emotions (e.g., Osborne et al., 2014) distinctly from stress (e.g., Arora et al., 2010) on human decision-making.

2.3 Stress and Decision-Making in Forensic Science

2.3.1 Characteristics of Stress

There is no universally accepted definition of stress or workplace stress (Adderley et al., 2012; Epel et al., 2018). A commonly used definition of stress comes from the theory of stress and coping (Lazarus & Folkman, 1984). This theory defines stress as when a person perceives the demands of the environment to be larger than their ability to cope by meeting, lessening, or altering these demands (Epel et al., 2018; Lazarus & Folkman, 1984). From this perspective, workplace stress broadly refers to the responses employees develop due to being exposed to demands at work that exceed their available resources and coping capabilities (Anshel, 2000; Kelty & Gordon, 2015). However, these definitions appear to be limited to the impact of high levels of stress on individual. The low levels of stress are also important for assessing expert performance (e.g. low stress can lead to boredom (Driskell et al., 2014)).

Despite there being no generally accepted definition of stress, there are general characteristics that have been recognised as being associated with stress (Epel et al., 2018). Stress is contextual, and dependant on the interactions of individuals with the environment (Epel et al., 2018; Kelty & Gordon, 2015; Lazarus & Folkman, 1984). Additionally, stress is dynamic, which means it can change over time (Kelty & Gordon, 2015; Lazarus & Folkman, 1984). For instance, daily discrimination against a worker by their supervisor (i.e., short-term, intense stress) can lead to long-term, chronic stress and disease (Epel et al., 2018).

Stress involves both physiological and psychological responses in an individual (Benson & Casey, 2013; Quick & Henderson, 2016). Under stress, people often engage in a physiological reaction commonly referred to as “fight or flight” (Benson & Casey, 2013). In this stress response, the brain triggers the autonomic nervous system to prepare the person to fight or flee from the problem. This reaction is accompanied by physiological changes, such as increases in stress hormone levels from the adrenal glands, an increase in heart rate, tightening of the muscles, and suppression of the immune system (Benson & Casey, 2013). A psychological response of “cognitive appraisal”, which involves perceptions of people to a stressful event, is a key component of the stress response. In other words, it is more about the individual’s perception of stress that makes a situation stressful rather than a stressor per se (Epel et al., 2018; Lazarus & Folkman, 1984).

Individual-level responses to stress vary, making some individuals better able to cope and perform under stress than others (Cooper & Marshall, 1976).

2.3.2 Stress and Performance

Stress does not always have a negative effect. Stress can have a positive impact on human decision-making performance (Kowalski-Trakofler et al., 2003; Paton & Flin, 1999; Yerkes & Dodson, 1908). The Yerkes-Dodson law empirically shows an inverted-U-shape relationship between stress and performance (Yerkes & Dodson, 1908). Performance is low at low levels of stress then increases with increased stress. However, this increased performance continues only until the level of stress is moderate — a "eustress stage" (Quick & Henderson, 2016). This can push individuals to meet deadlines (Benson & Casey, 2013; Jeanguenat & Dror, 2018).

Once stress increases beyond the eustress stage and moves towards high levels of stress, performance starts to drop (Benson & Casey, 2013; Yerkes & Dodson, 1908). LeBlanc et al. (2005) asked 30 paramedics to calculate drug dosage after working in a highly stressful scenario and found that acute stress could increase medical errors. Additionally, repeated exposure to stress has been shown to impair the cognitive ability of individuals (Deligkaris et al., 2014) and the well-being of forensic examiners at the workplace (e.g., Holt & Blevins, 2011). This is because repeated stress can result in negative workplace experiences. These occupational experiences can include causing physical (e.g., stomach distress and heart disease), psychological (e.g., anger and job dissatisfaction) and behavioural reactions (e.g., substance use, smoking and absenteeism) (Benson & Casey, 2013; Spector, 2012). High levels of stress or prolonged stress within the workplace can lead to poor work performance and decreased productively (Driskell et al., 2014). Equally important, stress can lead to diminished cognitive performance, mainly by affecting working memory (Deligkaris et al., 2014).

2.3.3 Stress and Working Memory

Working memory is thought to be a main cognitive function that can explain the association between stress and cognitive performance (see Deligkaris et al., 2014, for a review). This is based on the role working memory plays in many major cognitive functions needed for accurate performance, such as attention (Baddeley & Logie, 1999; Deligkaris et al., 2014; Gutshall et al., 2017).

Stress has been found to have an impact on working memory (Gutshall et al., 2017). The influence of stress on attention, as mediated by working memory, is still not clear (Deligkaris et al., 2014). Several studies have reported that stress narrows attention span (i.e., causes tunnel vision) under acute stress conditions. This occurs because the decision maker adapts to a simpler way of information processing to minimise nonessential information and focus more on the task at hand (Dror, 2007; Kowalski-Trakofler et al., 2003). In contrast, some researchers found that individuals under stressful conditions are more susceptible to distraction because of broadening of attention span (Keogh & French, 2001).

2.3.4 Risk-Taking Under Stress

The same decision problem may produce a different judgment if processed by a different decision-making system (Dror, 2007). There is a growing body of published literature that suggests that humans operate under two decision-making systems: system 1 is fast and intuitive, and system 2 is deliberative and logical but slower (Evans & Stanovich, 2013; Kahneman, 2003). This dual-system processing has been subject to criticism by some researchers who have argued that there is a continuum of the processing styles rather than there being discrete types (Osman, 2004). Yet, it can still offer a valuable model in understanding human decision-making behaviour (Evans & Stanovich, 2013), including making decisions under stressful situations that involve risk-taking (Reyna, 2004). For instance, it is evident that under time pressures, system 1 can offer a fast, heuristic decision based on intuition. This ready-made decision is based on the familiarity and experience of the decision maker (Kahneman, 2003), i.e., the top-down cognitive processing discussed earlier in Section 2.1.4. However, often good decisions are reached when both systems are involved because each system can support and constraint different aspects of the decision (Dror, 2007).

Cognitive strategies in human decision-making also involve risk-taking when under stress (Kerstholt, 1994; Maule et al., 2000). In the policing domain, for example, a police officer might decide to shoot a person who could be innocent, or might *not* shoot a person who could undertake a terrorist act (Dror, 2007) Risk-taking involves a complex equation of assessing different alternative choices and their consequences before making a decision —the "payoff matrix" (Kornbrot, 1988). The decision to take a risk depends on decision parameters (e.g., the complexity of the decision, number of alternative choices), internal factors (e.g., the state of mind of the decision maker), and external factors (e.g., stress and context (Dror, 2007)). For instance, Dror et al. (1999) found that time pressure decreased thresholds to make decisions. This means that less evidence is required before reaching a conclusion. This dynamic risk-taking under stressful conditions is also argued to be applicable to the forensic science discipline, as inferred from the cognitive model in forensic science (Dror & Langenburg, 2019). For example, fingerprint experts may opt for the less risky "inconclusive" conclusions because they are typically not challenged in courts, and are most often not verified —as opposed to identification decisions (Dror & Langenburg, 2019).

2.4 Sources of Workplace Stress in Forensic Science

2.4.1 Common Organisational-Level Sources

There are three main sources of stress at work (Cooper & Marshall, 1976): organisational-level factors (e.g., promotions and career development), extra-organisational factors (e.g., financial and family problems), and individual-level factors (e.g., tolerance for ambiguity; see also Section 2.3.1 for a discussion on individual-level responses to stress) (Cooper & Marshall, 1976; Epel et al., 2019; Kelty et al., 2021). These factors of workplace stress form part of the model of stress at work by Cooper and Marshall (1976), which is still being utilised as a framework to research workplace stress (e.g., Johnson et al., 2005).

The organisational-level factors are intrinsic to the job and the workplace environment (Cooper & Marshall, 1976). Forensic experts face organisational-level stress factors, which are either common across other domains or specific to the forensic science context (Jeanguenat & Dror, 2018; Kelty & Gordon, 2015). Common stress factors can originate from workload, from interactions with stakeholders (e.g., colleagues and managers), or from other organisation-specific aspects (e.g., work shifts and salaries; Kelty & Gordon, 2015).

Case workload has increased in the forensic science discipline, which adds pressure for the forensic examiners (National Institute of Justice, 2019). For example, forensic examiners who had more working hours per week reported higher levels of stress (Holt et al., 2017). There are various causes for the increased forensic caseload. Staff shortage is a common cause in the crime scene field (Kelty & Gordon, 2015). This shortage can affect staff morale particularly when there is an unexpected increase in major crimes that require the findings to be reported in less time (Kelty & Gordon, 2015). Technological advancement is another cause of increased forensic caseload because forensic science is being used more often in nonviolent crimes (Becker et al., 2005; Jeanguenat & Dror, 2018; National Institute of Justice, 2019). This kind of culture promotes continuous pressure to take more and more cases.

Relationships in the workplace are another common organisation-level stress factor in forensic science (Jeanguenat & Dror, 2018). They can be one of the primary causes of stress among criminal justice employees in general (Cullen et al., 1985; Holt et al., 2017; Johnson et al., 2005). As forensic examiners interact with multiple stakeholders — in the legal, investigative, forensic, regulatory, and public domains (Almazrouei et al., 2019; Dror & Pierce, 2020) — possibilities for task ambiguity or competing demands are possible for the examiners (Holt et al., 2017; Holt & Blevins, 2011). Such conflicting, unclear communication to the forensic examiners were found to be a significant factor for increased levels of stress in the forensic science discipline (Holt et al., 2017).

2.4.2 Context-Dependant Organisational-Level Sources

Several stress factors specific to the context of forensic science have been identified (Jeanguenat & Dror, 2018). One of these stress factors is the intensified scrutiny of forensic techniques and approaches and criticisms of their validity (e.g., National Academy of Sciences, 2009). This can create adverse working environments (Jeanguenat & Dror, 2018). Additionally, there are unreasonable expectations placed on the forensic experts by management and supervisors to not make mistakes due to the high stakes of the forensic results to stakeholders (Charlton et al., 2010; Murrie et al., 2019; Mustonen et al., 2015). Another unique stress factor is that forensic science across the board is underfunded despite increased demand (House of Lords Science and Technology Select Committee, 2019; Morgan & Levin, 2019; National Institute of Justice, 2019). Forensic experts may have a sense of job insecurity and uncertainty because their salaries and laboratory equipment are dependent on securing government funding (Jeanguenat & Dror, 2018; National Institute of Justice, 2019).

Forensic examiners can be directly exposed to emotionally distressing elements of crime scenes. These elements can come from violent crimes against children (Burruss et al., 2018; Kelty & Gordon, 2015), bloody and violent scenes (Salinas & Webb, 2018), and examining decomposed bodies (Iorga et al., 2016). For example, in Kelty and Gordon (2015), 9 of 19 crime scene examiners reported the presence of friends or family of victims at the crime scene to be one of the highest sources of workplace stress. They try to manage this stress by attempting to cognitively detach themselves and focusing on the scientific evidence (Kelty & Gordon, 2015). Alternatively, examiners can be indirectly exposed to case details and photographs when they get involved with managers to prepare the case strategy to triage the samples for examination (Jeanguenat & Dror, 2018).

2.4.3 Extra-Organisational Sources

There are several extra-organisational sources of workplace stress that can influence the physical and mental well-being of individuals in the workplace (Cooper & Marshall, 1976). These can include family problems (e.g., relationship with their spouse) and financial difficulties (Bell et al., 2012; Burke, 1994).

Stress from the personal life can "spill over" into the workplace and vice versa (Sok et al., 2014). For instance, in investigating 139 academics, Bell et al. (2012) found that individuals with higher perceived levels of stress at the workplace had a poorer balance between the workplace and personal life. This imbalance caused increased levels of work-life conflict (Bell et al., 2012). High work-life conflict has been shown to reduce job satisfaction and lower performance in the workplace (Burke, 1994; Frank et al., 2017; Hall et al., 2010). However, it is critical to emphasise that the personal life of individuals can serve as a coping mechanism for stress in addition to being a source of stress (for example, communicating feelings with the spouse (Kelty & Gordon, 2015)).

2.4.4 Moderating Factors of Stress

It has been found that considering the different attributes of individuals (such as their core field, sex, and years of experience) can help with understanding the moderating variables of stress (Holt et al., 2017; Kavanagh, 2005). Workplace stress within the same agency can differ according to the field of expertise. For example, the digital forensic field faces different stress factors to other forensic science fields (e.g., the job responsibilities of digital examiners can involve examining emotionally disturbing child pornography images and videos (Burruss et al., 2018; Holt & Blevins, 2011)). Different fields can have different caseloads, which can affect the stress from working on too many cases. For instance, in examining 4,205 criminal cases from five U.S. jurisdictions, it was found that biological evidence was more represented in rape cases, whereas latent prints dominated cases of burglaries and robberies (Peterson et al., 2013). Field-specific workloads can change. For instance, the recent dramatic increase in demand of digital evidence caused an approximate threefold increase in turnaround time (National Institute of Justice, 2019).

Considering sex differences when looking at the occupational experiences in regard to forensic workplace stress is vital. Women make up most of the forensic science discipline workforce. For example, in a study of 15 U.S. forensic laboratories, 58% of the scientific employees were female (Houck, 2009). Several studies reported that women can experience higher stress levels than men at the workplace for reasons such as having additional family responsibilities outside the workplace (Sharma et al., 2016) and differences in coping styles (Matud, 2004). However, there is still no clear relationship between sex and forensic science workplace stress. For instance, Holt and Blevins (2011) found that female digital forensic examiners experienced less stress than male counterparts. The opposite was reported in a subsequent study using a bigger sample size and looking into various forensic fields (Holt et al., 2017). Holt and Blevins (2011) did not provide an explanation for their findings, which opposed their hypothesis, because the duties of male and female digital examiners are the same.

Similar to stress varying according to the field of expertise and sex, there is still no clear pattern in how years of experience in forensic work may affect stress (Holt et al., 2017). In addition, contradictory findings–yet plausible causes of workplace stress, as it relates to experience– have been reported. For instance, some argue that more experienced individuals report higher stress for reasons such as having more responsibilities or being in a supervisory role (Holt & Blevins, 2011). Others reported that more experienced individuals would be less stressed because they have fewer stressful day-to-day tasks (Patterson, 2003), or because they develop adaptive coping strategies for stress (Gutshall et al., 2017).

2.5 Stress and Fingerprint as Evidence

2.5.1 Stress and Fingerprint Decision-Making

As indicated in Sections 2.2.22.2.3and 2.2.3, a number of published studies have investigated the impact of human factors, such as context (e.g., Dror et al., 2006; Smalarz et al., 2016; Stevenage & Bennett, 2017) and emotions (e.g., Hall & Player, 2008; Osborne & Zajac, 2015) on fingerprint decision-making . However, there is still a lack of literature studying the impact of stress on the decision-making in forensic science (Jeanguenat & Dror, 2018), including its impact on fingerprint decision-making.

It is contended by Ulery et al. (2017) that systematic bias can occur when the forensic organisation encourages the decision maker to make one decision over another. Such stress factors can vary among laboratories and cases, and responses of experts can vary according to these stressors (Ulery et al., 2017). For instance, some fingerprint experts reported that they were discouraged from making inconclusive decisions when the latent and known prints were of value and included a large area for comparison (Ulery et al., (2011); see Section 2.3.4 on a discussion on risk-taking). Another study found that some fingerprint experts seek conclusive outcomes due to a "need for closure", especially when working with serious crimes (Charlton et al., 2010). Therefore, the stress environment in which the examiner operates can be a contributing factor in the conclusions reached by fingerprint experts (Ulery et al., 2017).

Importantly, biased fingerprint conclusions were discovered in high profile case studies (e.g., Office of the Inspector General, 2006; The Fingerprint Inquiry, 2011), which led to doubts about the claims that fingerprint identification is accurate and infallible (for example, see Leadbetter, 2007). Reasons for these claims include the beliefs that fingerprints are permanent and unique to the individual (Cole, 2006) and the overconfidence of fingerprint examiners in their abilities or in the strength of fingerprint evidence (Mnookin, 2008). However, the validity of these beliefs has been criticised, with the individualisation conclusion to be with no scientific basis (Saks & Koehler, 2008). Similarly, the method of fingerprint examination and formulating conclusions have been questioned (Dror & Langenburg, 2019; Kellman et al., 2014; Mnookin, 2008), as will be illustrated in the next section.

2.5.2 Fingerprint Examination and Conclusions

Examining fingerprints is a complex task. Even with technologies like the Automated Fingerprint Identification System (AFIS), the assessment of fingerprints relating to crime still fundamentally relies on the visual-based decisions of fingerprint examiners (Dror et al., 2012; Thompson & Tangen, 2014; VanderKolk, 2011). This is partially because fingermarks from crime scenes typically have less data than fingerprints collected under more controlled situations (for example, reference fingerprints for comparison (Dror & Cole, 2010; Towler et al., 2018)). In addition, detailed fingerprint examination involves assessing the perceptual elements within the fingerprint area (Busey & Dror, 2011). This assessment, as with other types of pattern recognition domains, such as facial recognition and aircraft identification, requires human expertise (Busey & Dror, 2011).

Fingerprint experts often reach their conclusions using the ACE-V approach (analysis, comparison, evaluation, and verification). First, examiners assess the suitability of the latent fingermark from the crime scene (analysis). Then, examiners compare the suitable latent mark with the reference fingerprint from the suspect, searching for similarities and differences (comparison). In the evaluation stage, the similarities and differences between the mark/ print are assessed to reach a conclusion. In the verification stage, if conducted, a second examiner verifies the conclusion made by the first examiner (Langenburg et al., 2009; Stevenage & Bennett, 2017; VanderKolk, 2011). It is contended that the ACE-V approach is not a formalised method (Kellman et al., 2014; Mnookin, 2008), and it varies across jurisdictions/ labs (Langenburg, 2011; Stevenage & Pitfield, 2016). This is because there are no clear metrics or specifications on how to analyse the latent mark, or how to compare or evaluate the prints; hence, there is no specificity on how to assess and what counts as sufficient to make a conclusion.

At the end of the ACE-V approach, fingerprint examiners typically report a categorical conclusion: identification, exclusion, or inconclusive. The examiner makes an identification conclusion (i.e., the latent fingermark was made by the known fingerprint) if there are sufficient corresponding features (or similarities) between the two prints. An exclusion conclusion is reported (i.e., the latent fingermark was not made by the known fingerprint) wherein the examiner decides that there are sufficient differences (or dissimilarities) between the marks and prints. If the fingerprint examiner cannot decide whether the latent fingermark can be identified or excluded from the known fingerprint, an inconclusive conclusion is made (Langenburg et al., 2009; Stevenage & Bennett, 2017; VanderKolk, 2011). As indicated earlier in this section, there is a variability across jurisdictions on the threshold of reaching these conclusions (Langenburg, 2011; Stevenage & Pitfield, 2016). Not only do different jurisdictions use different thresholds, but examiners working within the same lab may have different thresholds. Furthermore, the same examiner, examining the same pair of prints, may use different thresholds at different times (Dror, 2016; Ulery et al., 2012).

2.6 Identified Gaps

2.6.1 Sources of Workplace Stress in Forensic Science

A few published studies have addressed the understanding of stress felt by forensic examiners in the workplace, such as looking at crime scene examiners exposed to horrific crimes (e.g., Kelty & Gordon, 2015; Yoo et al., 2013) and forensic odontologists exposed to mass casualties (e.g., Webb et al., 2002). Webb et al. (2002) argued that research on the psychological consequences of stress on forensic science professionals is lacking because of the general belief that professionals involved in emergency situations, such as forensic odontologists in mass disasters, are expected to deal with stress and demands as part of their job. It appears that most of the relevant research carried out so far has mainly focused on the stress experienced by crime scene examiners (e.g., Adderley et al., 2012; Craven et al., 2022; Kelty & Gordon, 2015; Salinas & Webb, 2018) or forensic examiners exposed to crimes against children, such as forensic interviewers (e.g., Bonach & Heckert, 2012; Brady et al., 2019) and digital forensic examiners (e.g., Burruss et al., 2018; Holt & Blevins, 2011; Seigfried-Spellar, 2018).

This gap in the evidence-base has driven some national bodies to call for more research into understanding workplace stress in forensic science. For instance, in 2020, the U.S. National Institute of Justice awarded more than $4 million in funding to this line of research (National Institute of Justice, 2020). Additionally, The American Society of Crime Laboratory Directors recently formed a Trauma and Stress Working Group (American Society of Crime Laboratory Directors, 2019). Notably, stress and mental health research on professional disciplines, including forensic science discipline, has gained momentum as a result of the COVID-19 pandemic (De Kock et al., 2021; Fournier et al., 2022; Puzzo et al., 2022). However, research into the workplace stress of core fields, such as forensic biology and fingerprinting, is still lacking (Holt et al., 2017).

The most comprehensive study investigating workplace stress surveyed 670 forensic examiners from various forensic science fields (Holt et al., 2017). This study included examiners from different state and federal forensic laboratories in the United States (Holt et al., 2017). However, inter-laboratory differences in the working cultures could have confounded the findings of the study. There is therefore a need to account for this possibility by targeting forensic examiners from within the same laboratory. Furthermore, this study, along with the other aforementioned studies, focused on the perceived influence of stress on the well-being of examiners (such as the trauma from stress (Burruss et al., 2018; Yoo et al., 2013)). However, they neglected data collection, or at minimum a focused discussion on its potential impact on the *decision-making* in forensic science (Jeanguenat & Dror, 2018).

Research addressing the stress experienced by forensic experts that considers demographical factors such as forensic science field, sex, and years of experience is also lacking (Holt et al., 2017). Such variations can help understand the factors that moderate stress, and how different factors play a role in creating, reducing, and managing stress (Holt et al., 2017; Kavanagh, 2005). Establishing the factors that forensic experts perceive to be stressful has implications for developing relevant, evidence-based approaches to improving the well-being of experts (Holt et al., 2017) and their decision-making performance (Jeanguenat & Dror, 2018).

2.6.2 The Impact of Stress on the Decision-Making in Forensic Science

Workplace stress has been shown to have an impact on the quality of decisions made by professionals in a variety of specialised domains, such as medicine (e.g., Arora et al., 2010) and policing (e.g., Akinola & Mendes, 2012). In the medical domain, for instance, a review of 22 empirical studies indicated that high levels of stress factors (such as bleeding, time pressure and procedural complexity) can affect the performance of surgeons (Arora et al., 2010). The relationship between stress and cognitive decision-making has "clear implications for professions that are characterised by high levels of work pressure and intense cognitive demands" (Deligkaris et al., 2014, p. 118), such as those in forensic science (Helsloot & Groenendaal, 2011; Jeanguenat & Dror, 2018). There is a lack of the aforementioned experimental paradigms in the forensic science fields (Jeanguenat & Dror, 2018), including fingerprinting.

The fingerprint field was chosen for the experimental studies in this thesis because fingerprint evidence is used frequently and can carry significant weight in court decision-making (Mustonen et al., 2015), so understating whether stress influences those decisions is critical. Furthermore, there is evidence from research (e.g., Charlton et al., 2010; Ulery et al., 2017) and casework (e.g., Kassin et al., 2013) to suggest that stress can affect fingerprint decision-making.

To date, no peer-reviewed studies have explicitly investigated the impact of stress on fingerprint decision-making (and, more broadly, on any forensic field). The small number of studies that included a stress factor in their design were limited in the following respects. First, the main aims, method, and findings did not explicitly address the direct impact of stress on the judgments in forensic science. For example, in the crime scene field (Helsloot & Groenendaal, 2011), forensic team leaders were asked to make decisions in settings that resemble real casework (i.e., ambiguous crime scene that involves making important decisions under time constraints after receiving contextual information at different times — hence, the implicit stress). Similarly, in the fingerprint field, the main aim of Stevenage and Bennett (2017) was to study the impact of contextual bias from prior knowledge of DNA results on fingerprint conclusions, not the impact of stress arising from the time pressure induced.

Furthermore, the design of the limited studies that indirectly assessed the impact of a stressor (predominantly time pressure) on fingerprint decision-making might not have been ecologically valid (Kellman et al., 2014; Stevenage & Bennett, 2017; Thompson & Tangen, 2014; Zou et al., 2021). The time allowed to make the fingerprint assessments was extremely short (e.g., two seconds to provide a fingerprint decision (Stevenage & Bennett, 2017)). In forensic settings, fingerprint examiners typically have ample time to make fingerprint evaluations (Kellman et al., 2014), and even when there is pressure to report conclusions faster, the deadline would not be so short. Additionally, these studies in the fingerprint field forced the participants to make either an identification or an exclusion (Stevenage & Bennett, 2017; Thompson & Tangen, 2014). In real casework, inconclusive decisions are allowed (Dror & Langenburg, 2019; Kellman et al., 2014).

To address these gaps, it is critical to examine the experiences (self-reported) and behaviour (experimental) of forensic experts. Forensic experts are the ones who perform the actual forensic casework (Almazrouei et al., 2019). Further, their decisions can affect subsequent police or court decision-making (Earwaker et al., 2020; Kelty & Gordon, 2015; Morgan et al., 2018). Collecting both self-reporting and experimental data can offer a more holistic understanding of the impact of workplace stress on the well-being of forensic experts, as well as their decision-making performance. This can ultimately provide an evidence-base for more efficient, accurate, reproducible and robust crime reconstructions.

2.7 Aims and Research Questions

Therefore, the aim of this PhD thesis was to explore the extent to which the forensic experts experience stress in the workplace, and its potential impact on their decision-making. Specifically, this research sought to provide data on the possible sources of stress as felt by the examiners, including the explicit and implicit pressures from the feedback they receive during casework (objective (a)). In addition, to address objective (b), experimental work was undertaken to explore the possible influence of stress on forensic science decision-making by considering whether stress has a positive or a negative impact on the quality of experts' decisions, and willingness to take risks by reporting inconclusive decisions.

This thesis addresses the following main research questions:

- Research question 1: What degree of workplace stress (and feedback) do forensic experts experience? To address research question 1, forensic examiners were asked to complete two surveys about the stress they feel and its sources (Chapters 3 and 4). Answering this question contributes to objective (a) of the thesis.
- Research question 2: Does stress impact conclusions reached in crime reconstructions by forensic experts? Specifically, does stress influence conclusions reached by fingerprint experts, and does this influence differ between experts and novices? To address research question 2, online experiments were conducted (Chapters 6), after developing a method to stress participants online, without the presence of researchers (Chapter 5). Addressing this question contributes to objective (b) of the thesis.

Findings from this thesis offer a holistic understanding on the possible sources (self-reported) and impact (experimental) of stress on decision-making in forensic science contexts. Hence, this research may help devise appropriate strategies for optimal working environments for forensic experts who perform the actual casework.

Chapter 3 Organisational and Human Factors Affecting Forensic Decision-Making: Workplace Stress and Feedback (Study 1)

3.1 Introduction

Forensic examiners operate in a stressful environment (Holt et al., 2017; Kelty & Gordon, 2015; National Institute of Justice, 2019). For example, Holt and Blevins (2011) surveyed 56 digital forensic examiners and found that around 68% were working under a lot of pressure at work. Participates in this study reported a number of coping mechanisms, such as drinking alcohol and smoking. As stress becomes high, performance and quality of decisions start to drop (Yerkes & Dodson, 1908). In forensic science, quality of judgments includes accuracy, but also other issues, such as confidence levels, documentation of the decision-making process, reporting of the conclusions, ability to justify the decisions and their presentation in court (Dror & Pierce, 2020; see also (Dror, 2016) for Hierarchy of Expert Performance).

Feedback is a critical factor in its own right, that can impact well-being and performance (Choi et al., 2018), as it can have implications for the motivation, expectations and the decision-making of forensic examiners (e.g., questions 8 and 9 in (Kukucka et al., 2017)). Therefore, understanding the ways feedback given to forensic examiners and how it may affect their decision-making, is important for understanding the context in which decisions are made (Almazrouei et al., 2019; Dror & Pierce, 2020). This has the potential to impact the entire crime reconstruction process (Morgan et al., 2020).

Human factors are not independent, and often affect one another. For example, stress and emotions are closely related, as stress can generate negative emotions (Du et al., 2018). Similarly, stress and feedback are related (e.g., pressures from feedback can cause stress). Importantly, such pressures can impact conclusions (Ulery et al., 2017):

> "Errors and disagreements among examiners may be due to in part . . . [to] systemic pressures encouraging some decisions more than others. These pressures will vary by agency or among cases, and examiners' responses to these pressures will vary." (p. 66)

The study reported here deals with these human factors of stress and feedback that can affect decision-making. A questionnaire was designed to contain questions about stress and feedback (see Appendix A). For clarity in presenting the findings, this chapter is divided into two parts. The first part focuses on stress experienced at the workplace, examining the existence of and sources of stress in forensic science laboratories. The second part addresses the feedback provided, examining how it is perceived by practicing forensic examiners.

3.2 Part One: Workplace Stress

Research on workplace stress factors in forensic science have generally been neglected in the published literature (Jeanguenat & Dror, 2018; National Institute of Justice, 2019). There is a lack of research addressing workplace stress of examiners working in forensic science in general, and specifically across core forensic science fields (such latent prints and forensic chemistry) (Holt et al., 2017), and across different stages of their career. It is argued that research on the psychological consequences of stress experienced by forensic science professionals is lacking because of the general belief that professionals involved in emergency situations are expected to deal with stress and demands as part of their job (Webb et al., 2002).

Research on stress experienced by forensic experts can help in understanding the factors that moderate stress, and how different factors play a role in creating, reducing, and managing stress (Holt et al., 2017; Kavanagh, 2005). This may have implications for developing relevant evidence-based approaches to improve the wellbeing of experts as well as their decision-making performance. Therefore, this study explores the factors that may cause forensic science examiners to feel stress. It was of interest to examine the contribution of stresses attributed to the workplace as opposed to personal factors; whether there were differences in the stresses felt by examiners working in different forensic science fields; and whether the years of experience moderated the level of stress experienced.

3.2.1 Method

3.2.1.1 Questionnaire

Following established approaches in decision-making studies within the forensic science discipline (Gardner et al., 2019; Hamnett & Jack, 2019; Kukucka et al., 2017), and studies addressing perceptions of workplace stress factors (e.g., Burruss et al., 2018; Holt et al., 2017), a questionnaire was designed to examine workplace stress (Part One) and feedback (Part Two).

Part One contained questions to ascertain whether forensic examiners had felt stressed at work, and how much of the stress they attributed to personal reasons (e.g., family, medical, and/or financial matters) as opposed to relating the stress to the workplace (see Figure 3-1). The participants were required to rank their responses on a seven-point Likert-type scale. The participants were also asked to provide demographic information on their primary forensic field and years of experience.

3.2.1.2 Participants

A total of 150 forensic examiners from a major forensic laboratory in the United States took part in the study (71% response rate; $N = 212$). All the participants were practicing forensic examiners, and they were from the same forensic laboratory, so that it was possible to examine and compare variables (e.g., fields of expertise and years of experience) without introducing inter-laboratory variations.

Forensic examiners identified their primary fields as: biology/ DNA ($n = 42$), latent prints ($n = 40$), controlled substances ($n = 24$), forensic alcohol ($n = 7$), toxicology ($n = 4$), firearms ($n = 9$), and trace evidence ($n = 5$). Nineteen (13%) did not report their primary field, and three latent print examiners stated that they also work as crime scene examiners as a secondary field. The fields were grouped together on the basis of the type of expertise deployed, giving three field categories: forensic biology ($n = 42$; DNA and biology), latent prints ($n = 40$), and forensic chemistry ($n = 35$; controlled substances, toxicology, and forensic alcohol). The remaining fields (trace evidence, firearms, and crime scene investigation as a secondary field) were excluded from the analysis by field of expertise, because they contained low participant numbers and did not fit within any of the three main field categories.

The mean years of experience was 12 ($SD = 9.7$ years, with a range from 1 to 47 years; did not respond: $n = 12$). Four examiners provided a qualitative written response to the question about their years of experience (e.g., “many” or “lots”) or the number written was illegible and thus not included in the years of experience analysis (i.e., 16 participants (11%) were excluded from the analysis by experience, leaving 134 participants). Following the accepted approach in the published literature to categorise data, such as the years of experience (e.g., Holt et al., 2017; Yoo et al., 2013), the years of experience were grouped into categories of comparable sample sizes: early-career (0 to 5, n = 36); mid-career (6 to 10, n = 28) and (11 to 20, n = 40); and late-career (>20, with n = 30).

3.2.1.3 Statistical Analysis

Both descriptive and inferential statistics were applied, using SPSS (version 25), to measure the reported stress levels in general, and to examine stress by field and years of experience. Following previous research (Holt et al., 2017), the seven-point Likert-type scale responses were converted to an ordinal, categorical scale of low, moderate, and high scores: scores 1–2 as low (i.e., low feelings of stress), scores 3–5 as medium, and scores 6–7 as high (i.e., strong feelings of stress). Equal categories of low and high scores were made as per previous published research (Holt et al., 2017). However, it should be emphasised that some of the neighbouring scores (e.g., scores 2 and 3) are grouped in different categories (i.e., low and medium) and this is reflected in the interpretation. Likert scales can be categorised (e.g., Kukucka et al., 2017) and can be statistically treated at an ordinal level (Jamieson, 2004). This categorisation helps to examine the variability of stress experienced by the examiners.

A chi-square test (goodness of fit) was used to determine whether the categorical responses for each question differed significantly (i.e., low *vs.* high stress scores; see Figure 3-1). An alpha significance level of 0.05 was used for all the statistical tests. In addition to the significance testing, the means and standard deviations are reported.

One-way ANOVA and post hoc (Bonferroni) were used to compare the mean workplace stress levels across the categories of forensic fields and years-of-experience. In case that the homogeneity of variance assumption was not met, as assessed by Levene's test, then a one-way Welch ANOVA and post hoc (Games-Howell) were used instead. In addition to comparing the means, a chi square test was used to test whether the responses of the high scores for the three categories of forensic fields differed significantly from one another. The stress scores were particularly important at the high levels where the influence of stress on the well-being and performance of forensic examiners can be most critical (Benson & Casey, 2013; Deligkaris et al., 2014; Yerkes & Dodson, 1908).

3.2.2 Results

3.2.2.1 Workplace Stress

One in three forensic examiners (36%, n = 53) reported that they often experience stress while at the workplace (low *vs*. high stress scores, χ^2 (1, N = 79) = 9.23, p = .002; M = 4.61, SD = 1.90; see Figure 3-1). For the high stress levels felt by the examiners, stress was attributed more from the workplace (i.e., 25%, n = 37, from management and/ or supervisors (χ^2 (1, N = 96) = 5.04, p = .025; M = 3.62, SD = 2.16), and 20%, n = 29, from backlog pressure (χ^2 (1, N = 95) = 14.41, p < .001; M = 3.30, SD = 2.05)) than from the personal life (11%, n = 16; χ^2 (1, N = 84) = 32.19, p < .001; M = 3.14, SD = 1.85).

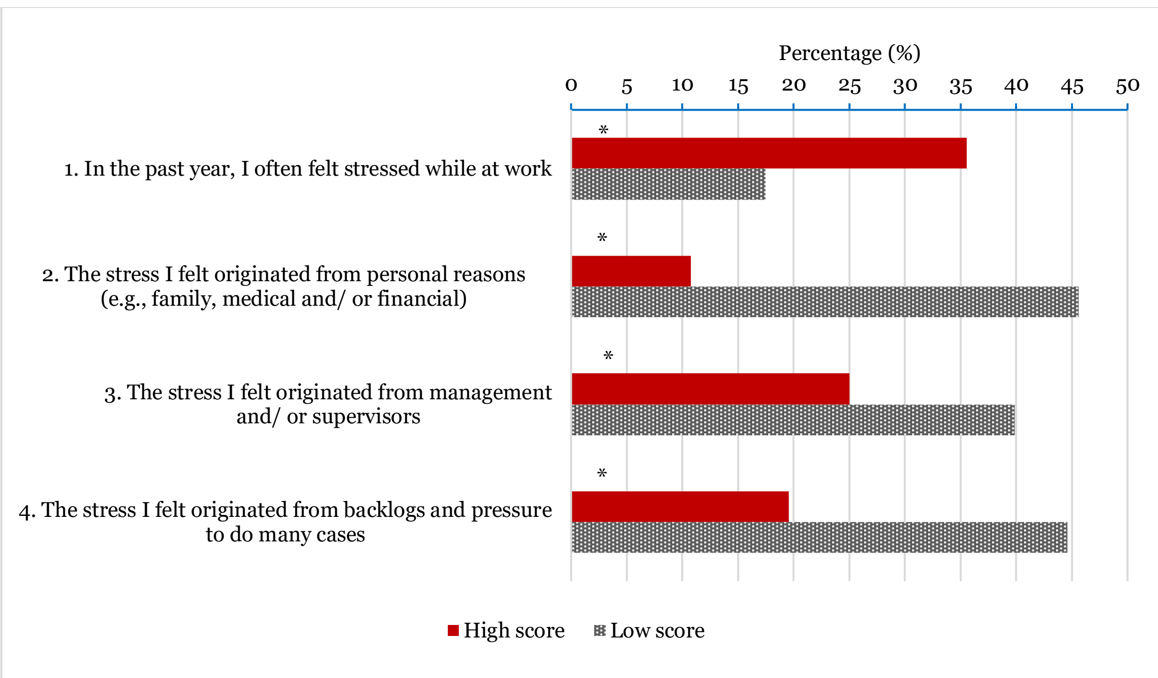

Figure 3-1: Scores of stress levels. * p < 0.05 for χ2 of low *vs*. high stress level scores.

3.2.2.2 *Stress by Field and Experience*

On average, moderate workplace stress (question 1) were felt by all forensic field categories: biologists (M = 5.02, SD = 1.94), latent print examiners (M = 4.75, SD = 1.77) and forensic chemists (M = 4.09, SD = 1.92). Whilst the mean stress levels did not vary across the three field categories (questions 1-4, $p > .05$), the level of high stress differed from backlog pressure only, χ^2 (2, N = 24) = 7.75, p = .021. The percentage of forensic biologists (34%, n = 14) who strongly felt that their stress originated from backlog pressure was higher than the other fields, i.e. latent print examiners (18%, n = 17) and forensic chemists (9%, n = 3).

The mean stress levels varied across experience groups, but only due to stress from management and/ or supervisors (question 3, *Welch's* $F(3, 67.7) = 6.01$, p = .001) and backlog stress (question 4, *Welch's* $F(3, 67.7) = 8.15$, $p < .001$; see Table 3-1). There were no interactions between the forensic field and years of experience on the reported stress levels (univariate ANOVA for questions 1-4, $p > .05$).

Table 3-1: Mean responses for questions (3), (4) and (7) where significant findings were found among the means of four experience groups in the current study.

Question		0-5	6-10	11-20	>20
3.	Management stress	2.53 (1.63) $^{a, b, c}$	4.21 (2.83) ^{a}	3.70 (2.12) ^{b}	4.20 (2.28) ^{c}
4.	Backlog stress	2.06 (1.51)$^{d, e, f}$	3.37 (1.98)^{d}	3.98 (2.19)^{e}	3.50 (1.94)^{f}
7.	Feedback on expected conclusions	2.18 (1.62)^{g}	2.89 (1.85)	2.76 (1.48)	3.33 (1.94)^{g}

$^{a, b, c, d, e, f}$ p<0.05, Post hoc (Games-Howell); ^{g} p<0.05, post hoc (Bonferroni)

3.2.3 *Discussion*

3.2.3.1 *Workplace Stress*

On average, forensic examiners in this study reported a moderate frequency in feeling stressed at the workplace (question 1, M = 4.61, SD = 1.90). However, there was variability in the data as reflected by the standard deviations, and by the low and high stress scores (see Figure 3-1). Variability is expected given individual differences in responding to stress factors (Epel et al., 2018; Lazarus & Folkman, 1984). Also worth noting is that although question 1 asked examiners on the *frequency* of their stress at work (i.e., "often"), the responses to this question can also reflect their *level* of stress. It is generally reasonable to assume that people who feel stressed more frequently also feel higher levels of stress (e.g., see transdisciplinary model of stress that describes 'stress' as a set of integrated processes, including the history of stressors in the life of an individual (Epel et al., 2018)).

In this study, 36% of the forensic examiners strongly felt that they are often stressed at work. Published research from other domains has shown that repeated exposure to stress or when stress levels are high, the well-being (Benson & Casey, 2013) and decision-making performance drops (Deligkaris et al., 2014; Yerkes & Dodson, 1908). For example, LeBlanc et al. (2005) asked 30 paramedics to calculate drug dosage after working in a highly stressful scenario and found that intense stress increased medical errors.

The data from the study reported here concerns the feelings experienced by forensic examiners. It does not include objective measures of the performance and quality of decisions of the participants. Hence, the data reported does not show the nature of the causational relationship, if any, between high stress and performance. Higher levels of stress can impact performance in a number of ways. These data cannot ascertain the impact, but clearly shows that stress is felt by forensic examiners, and hence warrants further research.

Future research needs to experimentally examine the impact of stress on the decision-making performance in the forensic science context, as has been studied in other specialised domains (see for example, Arora et al. (2010) for a review of studies that investigated the impact of stress in the medical domain). Such experimental research is important given the critical nature of forensic science decisions within the criminal justice system (Morgan, 2017a; Morgan et al., 2020).

In the current study, 17% of forensic examiners reported feelings stressed at work relatively infrequently (if they felt stressed at all). It has been observed in some contexts that low levels of stress can lead to underload, boredom and lower performance (Driskell et al., 2014). Conversely, moderate stress can improve performance (Yerkes & Dodson, 1908), as it can, among other things, push individuals to meet deadlines (Jeanguenat & Dror, 2018). Hence, the published literature addressing stress suggests that there could be benefits in maintaining moderate stress levels at the workplace of forensic examiners (by, for example, providing new, interesting tasks to motivate underloaded, low stressed individuals (Driskell et al., 2014)).

The findings of this study suggest that the forensic laboratory management and/ or supervision contribute to the stress levels felt by the forensic examiners (the way the question was framed in the survey does not allow us to determine if it was the laboratory management or the supervisor that created the stress, or both –it is only possible to identify that there was stress felt and it was attributed to either or both of these factors). Published research addressing stress suggests that relationships in the workplace are a common organisational-level stress factor, and that they can be one of the primary causes of stress among criminal justice employees in general (Cullen et al., 1985; Holt et al., 2017). Hence, it would appear that forensic management and/ or supervisors may play a key role in optimising the stress levels and well-being of forensic examiners.

Similarly, the findings of the current study reveal that backlogs and pressure to complete many cases can contribute to the stress felt by the forensic examiners. It has been suggested in the published literature that pressure from case backlog is intensified by the increase of requests from prosecutors and law enforcement agencies for rapid forensic analysis and reports (e.g., Houck & Speaker, 2020), in addition to increasing forensic service requests for non-violent crimes in an under-resourced and overtaxed forensic science environment (Jeanguenat & Dror, 2018; National Academy of Sciences, 2009). However, it is acknowledged that backlog pressure is a complex measure and can vary from one forensic organisation to another (National Institute of Justice, 2019).

The findings show that more examiners strongly felt that their stress originated from the workplace than arising due to personal reasons. It is, however, important to note that the questions posed in the this study did not directly relate personal and workplace causes of stress in one question so as to offer the opportunity for examiners to rate one type of stress factor directly against the other. Further research on personal life stress is needed, as it has been suggested in the published literature that stress from the personal life can affect the work-life balance, increase work-life conflict, reduce job satisfaction and lower performance in the workplace (Burke, 1994; Hall et al., 2010).

3.2.3.2 Stress by Field and Experience

On average, forensic biologists, forensic chemists, and latent print examiners reported moderate frequencies or levels of stress at the workplace (again, it is important to note that there were individual differences even within the same forensic science field). Previous research targeting specific forensic fields yielded inconsistent findings. For instance, forensic odontologists reported low stress levels when attending mass casualty incidents, for reasons such as having sense of achievement and obtaining invaluable professional experience (Webb et al., 2002), whereas digital forensic examiners reported moderate levels of stress in undertaking their roles (e.g., examining child pornography (Holt & Blevins, 2011)). These previous studies were conducted across laboratories, hence, it is not possible to attribute the different findings to the forensic fields, because these differences may arise from other confounding factors, such as the general workplace culture and environment in the laboratory.

The results from this study, within a single laboratory, allows for a better comparison across forensic fields. These data indicate that high levels of stress from backlog pressure varies among the three fields; specifically, more forensic biologists strongly felt stress from backlog pressure in comparison to forensic chemists and latent print examiners. However, as previously mentioned, backlog is a complex measure and has been shown to vary across forensic organisations – even within the same field of expertise - and can change with time (National Institute of Justice, 2019). The dynamic and complex nature of backlog pressure suggests that each forensic organisation may be well advised to evaluate the way they communicate their own backlogs among the different forensic fields, and how it can influence the well-being and performance of their forensic examiners.

The findings also reveal that mid and late career examiners— i.e., over 5 years of experience— felt more stress originating from management and/ or supervision and from backlogs in comparison to early career examiners—i.e., under 5 years of experience (there were no interactions between field of expertise and years-of-experience categories in all the stress questions). A previous study suggested that examiners with more experience have more workload responsibilities, such as having a supervisory role (Holt & Blevins, 2011), which may go some way towards offering insight to this trend that was observed in this study.

There are differences in the levels of workplace stress across occupations (Johnson et al., 2005). There is insufficient understanding and data about stress in forensic science to enable a meaningful comparison to other occupations. This study is one of the first to address workplace stress from various forensic science fields (with statistical comparisons of examiners working in primary fields, such as forensic biology and chemistry). In addition, since data were collected from one laboratory, the data does not necessarily generalise to other forensic laboratories. However, there are good reasons to believe that forensic science is a high stress occupation in comparison to typical working environments (Jeanguenat & Dror, 2018; National Institute of Justice, 2019). Working environment and organisational culture are human factors that impact forensic decision-making (Dror, 2020a).

3.3 Part Two: Workplace Feedback

Feedback is a key component of the conceptual model of communication in forensic science presented by Howes (2015). Additionally, feedback received by forensic examiners who perform casework analysis and interpretation, is an important component of monitoring and improving performance, and motivating and rewarding examiners for hard work (e.g., Choi et al., 2018). Feedback can be explicit (messages that can be directly codified and articulated) (Ellis et al., 2006; Morgan, 2017b), such as an immediate supervisor saying "well-done" to the examiner. Feedback can also be implicit, meaning that messages are not direct and less codified (Ellis et al., 2006; Morgan, 2017b). An example of implicit feedback would be the supervisor "smiling" to the examiner, which can cause subjective interpretation and experiences of emotions (Söderkvist et al., 2018).

Stress and pressure resulting from explicit and/ or implicit feedback can influence forensic science judgments. In an earlier study, some fingerprint examiners reported that they were not allowed or were discouraged from making inconclusive decisions when the latent mark and known prints were of value and included a large area for comparison (Ulery et al., 2011). Moreover, Kassin et al. (2013) discussed that a contributing factor of the misidentification in the 2004 Madrid train bombings was the increased 'need for closure' (i.e., the desire to provide clear-cut judgments (Ask & Granhag, 2005)), which resulted in a subsequently established erroneous identification of Mayfield. It is salient that an independent investigation report on this case stated that the criteria for reaching an inconclusive result could lead to implicit pressures on an examiner to reach an identification when making a difficult comparison of marks, particularly when the case was very serious (Office of the Inspector General, 2006).

Previous published research has started to look into the possible relationships between perceived feedback and forensic expert decision-making (e.g., questions 8 and 9 in Kukucka et al. (2017)). Yet its impact and scope are still largely unexplored. This current study assessed the explicit and implicit feedback, as felt by the forensic examiners with the following key actors (see Figure 2-2): forensic management and/ or supervisors (the forensic services domain), police investigators (the investigative domain) and legal advocates (the legal domain). These have been identified as actors that can impact decisions made during crime scene work, laboratory analysis, and/ or judicial procedures (Julian & Kelty, 2015; Kelty et al., 2018; Murrie et al., 2013; National Academy of Sciences, 2009).

Therefore, the second part of this current study sought to identify the level of explicit and implicit feedback as felt by the forensic examiners, and whether the feedback varied by forensic science field of expertise or years of experience.

3.3.1 Method

The same methodology was followed as outlined in Part One, with the only difference being the inclusion of three questions on feedback. Specifically, the feedback questions addressed whether forensic examiners received feedback about their work from stakeholders, such as from management, supervisors, police investigators and/ or legal advocates (i.e., explicit feedback; see question 5 in Figure 3-2). In addition, questions 6 and 7 asked whether the forensic examiners felt that the stakeholders appreciated them more when they help to solve a case (such as when finding a 'match' rather than 'inconclusive') and whether the examiners sometimes felt they know what the stakeholders expect or want their conclusions to be (i.e., implicit feedback; Figure 3-2).

3.3.2 Results

3.3.2.1 Workplace Feedback

About half (49%, n = 71; M = 3.06, SD = 1.93) of forensic experts reported low scores for feeling that management, supervisors, police investigators and/ or legal advocates appreciated it more when they were helping to solve cases, and that sometimes they felt they knew what these stakeholders wanted or expected their conclusions to be (53%, n = 77; M = 2.75, SD = 1.77). Nevertheless, some examiners, albeit a small minority, reported high scores for feeling such feedback and expectations, 14%, n = 20, $\chi^2(1, N = 91) = 28.58$, $p < .001$ and 8%, n = 11, $\chi^2(1, N = 88) = 49.50$, $p < .001$, respectively. Examiners were equally divided (27%, n = 40, high scores *vs.* 28%, n = 42, low scores; $p > .05$) on whether they receive explicit feedback (M = 3.95, SD = 2.00; see Figure 3-2).

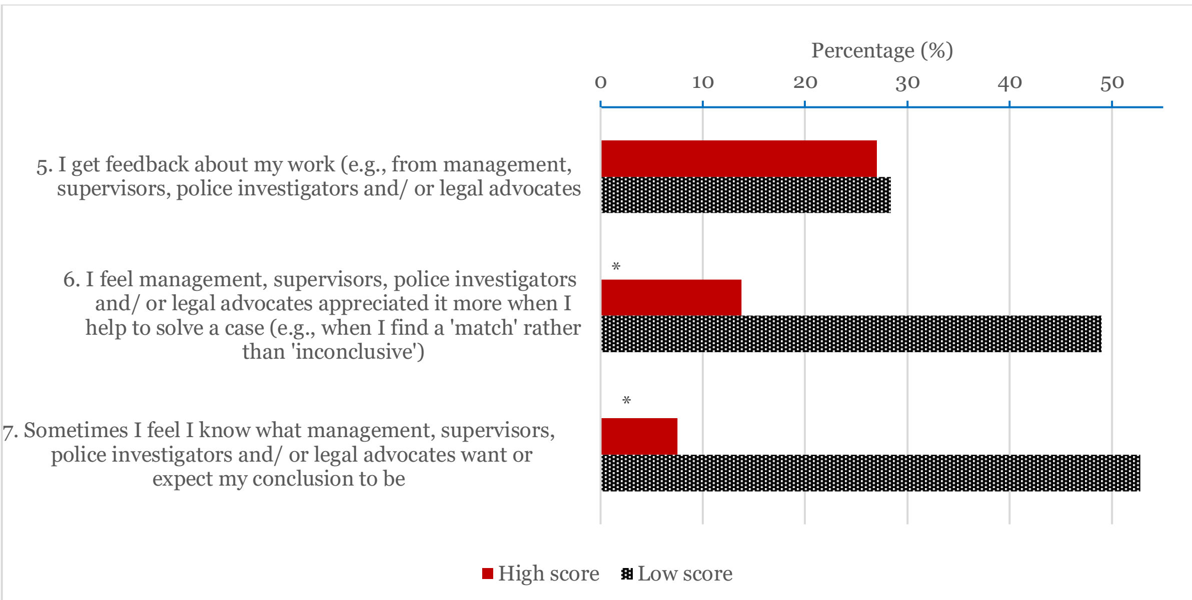

Figure 3-2: Scores of explicit and implicit feedback. * p <0.05 for χ2 of low *vs.* high.

3.3.2.2 Feedback by Field and Experience

On average, most forensic biologists (M = 4.49, SD = 2.06), forensic chemists (M = 3.77, SD = 2.00) and latent print examiners (M = 3.62, SD = 1.96) felt they received moderate explicit feedback from their management, supervisors, police investigators and/ or legal advocates. Both the explicit and implicit mean feedback levels did not significantly differ by field of expertise (questions 5-7, $p > .05$). However, for the high scores of the explicit feedback question, more forensic biologists (41%, n = 17) reported receiving feedback than latent print examiners (21%, n = 8) and forensic chemists (20%, n = 7; approaching statistical significance, $\chi^2(2, N = 32) = 5.69$, $p = .058$).

Question 7 on expected conclusions was the only feedback question that varied by experience (approaching significance, $F(3, 126) = 2.54$, $p = .060$; see Table 3-1). There were no interactions between the forensic science field and years of experience on the reported feedback levels (univariate ANOVA for questions 5-7, $p > .05$).

3.3.3 Discussion

3.3.3.1 Explicit Feedback

Forensic examiners were divided on whether they receive low or high amounts of explicit feedback about their work from the stakeholders they interact with. Additionally, on average, forensic examiners reported receiving similar levels of explicit feedback across the investigated forensic science fields and experience groups. However, more forensic biologists reported receiving high levels of explicit feedback than the latent print examiners and forensic chemists did, whilst at the same time, more forensic biologists reported experiencing high levels of stress from backlog pressure than the other two fields of expertise (see Part One). The data, however, do not include measures to inform an understanding of how such feedback impacts the well-being and the performance of the forensic examiners. Therefore, in order to consider the explicit feedback within the crime reconstruction process further, it will be important for future research to identify what type and level of feedback is warranted (Almazrouei et al., 2019; Dror & Pierce, 2020; Morgan et al., 2018).

3.3.3.2 Implicit Feedback

A few forensic examiners strongly felt that sometimes they knew what stakeholders wanted their conclusions to be (question 7, 8%; see Figure 3-2). Despite being a low proportion, this finding on implicit feedback is concerning because each forensic examiner is involved in casework analysis and interpretation (Dror, 2018). The findings also show that a higher level of implicit feedback was felt by late career (>20 group) in comparison to early career examiners (0-5 group), in terms of what stakeholders wanted or expected their conclusions to be (see Table 3-1). This finding is consistent with previous research, which found that 63.6% of forensic examiners agreed (i.e., slightly agreed, agreed and strongly agreed) that on occasions they know what conclusions they are expected to find (Kukucka et al., 2017), and that forensic examiners can be pressured to extend opinions beyond their scientific findings (Becker et al., 2005).

To be clear, the aforementioned findings do not demonstrate that the examiners are in fact being pressured by the stakeholders to reach expected conclusions. Rather, the data illustrate what the examiners perceive and feel as implicit pressure. It is the perception and feeling of stress that makes a situation stressful rather than there being an actual stress factor (Epel et al., 2018; Lazarus & Folkman, 1984). It is important to consider the context within which decisions are being made to ensure there is transparency in this process to mitigate conditions that exert pressure on examiners to make 'expected' decisions.

The findings from this study demonstrate that some (question 6, 14%) forensic experts felt strongly that stakeholders in the forensic services, investigative and legal domains appreciated it more when they reported conclusions of high certainty (e.g., a clear-cut, match conclusion as opposed to inconclusive). While this is a low percentage of the sample, this high implicit feedback score is also concerning. It shows that some active casework scientists may feel an implicit pressure to reach certain conclusions. As stated earlier, it is the 'cognitive appraisal' of the individual to the situation that makes it pressurising (Epel et al., 2018; Lazarus & Folkman, 1984), even in the absence of such pressures. It is of course important to note that these data cannot indicate whether conclusions are being influenced by such implicit pressures.

3.4 General Discussion

Taking the stress and feedback findings together, many of the forensic examiners in this study perceived that they operate under pressure, and that the level of pressure varies by field and experience, during casework and reporting conclusions. The findings emphasise that one must consider the operating environment that forensic examiners work in, and the importance of managing the levels of workplace stress and feedback.

The insights from the data provide a valuable but limited insight into the possible relationships between feedback, stress and forensic decision-making. This study clearly cannot identify and characterise the relationships but indicates that this could be a fruitful avenue for future studies. Additionally, as detailed earlier, human factors (such as stress and feedback) are interrelated and affect one another (Du et al., 2018). Hence, it is possible that the questions addressing the feelings of examiners regarding implicit feedback (i.e., questions 6 and 7) can be related to stress and/or other human factors.

The current study further contributes to the forensic science literature by synthesising relevant stress and feedback literature from other domains. It offers a focused theoretical discussion, along with empirical data, on how workplace stress and feedback can affect forensic science judgments (whereas most of the previous research mainly focused on the relationship between stress and well-being of forensic examiners (e.g., Burruss et al., 2018; Holt et al., 2017)). In addition, the current study unpacks the notion of feedback, an under-researched but important human factor in forensic science. It is hoped that this study will drive further research directed towards workplace feedback and its potential effects on expert decision-making.

The published literature suggests that there can be individual differences in perceiving and coping with stress (Epel et al., 2018; Lazarus & Folkman, 1984). This means that forensic examiners can perceive and cope with stress and feedback differently, even among those examiners who work in the same laboratory and forensic field, and have the same years of experience. The current data accounts for inter-laboratory variations (Roux et al., 2021), as it has been collected from a single laboratory. However, differences in individual stress perceptions and coping styles were not investigated, and so should be considered in future research and also in practice.

It is important to note that self-reporting from a participant of how they feel about stress or feedback can offer valuable and informative insights. However, individuals cannot accurately describe the rationale of their decision-making, as this often involves unpacking complex cognitive processes (Gardner et al., 2019; Nisbett & Wilson, 1977). It is possible, for example, that the workplace stress felt by the forensic examiners is originating from personal reasons (Hall et al., 2010), and it could have been difficult for participants to separate the workplace from personal causes of stress. In addition, the responses of forensic examiners may have been affected by social desirability bias (Chung & Monroe, 2003), in particular for the implicit feedback questions. Although the current study included a large sample size of 150 practicing forensic examiners from the same laboratory, it may not be representative to forensic laboratories worldwide. The reported levels of stress and feedback may vary in other jurisdictions that have different working environments and cultures.

3.5 Conclusion

This study surveyed active forensic examiners with different fields of expertise and years of experience working within one laboratory. The examiners reported feeling varying levels of workplace stress, and levels of explicit and implicit feedback. More high levels of stress were reported to originate from the workplace (specifically, stress from backlogs and pressure to do many cases, and management and/or supervisors) than from stress derived from personal reasons outside the workplace. More forensic biologists perceived high levels of backlog pressure than latent print examiners and forensic chemists. Mid and late career examiners (i.e., over 5 years of experience) reported higher stress levels originating from management and/ or supervision, as well as backlog pressure in comparison to early career examiners (i.e., less than 5 years of experience).

It was concerning that a few forensic examiners sometimes felt strongly that they knew what the stakeholders in the forensic services, investigative and/ or legal domains expected or wanted their conclusions to be, and that some forensic examiners also strongly felt that the same stakeholders appreciated it more when they helped to solve a case (e.g., by finding a match as opposed to inconclusive).

In a broader context, the creation of working environments that can address the negative impacts of the types of stress examiners are exposed to will be valuable. It is also important to be aware of the impact of both explicit and implicit feedback, and to develop practices that ensure the positive assistance and timely explicit feedback. This may include preventive risk management measures (Dror & Pierce, 2020), such as the evaluation of the how backlogs are measured and communicated to forensic examiners across different fields of expertise. It is also important to consider the context within which decisions are being made to ensure there is transparency in this process to mitigate conditions that exert pressure on examiners to make ‘expected’ decisions.

To further address objective (a) of this thesis, a follow up study was conducted (Chapter 4) to explore other stress factors (such as stress originating from crime scene work), the level of support that forensic examiners experience, and whether experts think that they are influenced by stress or not.

Chapter 4 Workplace Stress and Support: The Perspective of Forensic Experts (Study 2)

4.1 Introduction

Stress is not necessarily negative (Benson & Casey, 2013; Yerkes & Dodson, 1908) as stress, at moderate levels, is recognised to be a motivating factor (Driskell et al., 2014). However, research that assesses levels of support and the sources of workplace stress and their potential effects on forensic examiners' well-being and decision-making is still lacking (Jeanguenat & Dror, 2018; National Institute of Justice, 2019). Such research efforts are needed to keep pace with other professional domains, such as medicine (e.g., Arora et al., 2010; Zavala et al., 2018), terrorism (Corner & Gill, 2019) and policing (e.g., Akinola & Mendes, 2012; Cullen et al., 1985). To date in the forensic science published literature there have been very few studies that have considered organisational factors and their implications for decision-making in casework across different forensic science fields and career stages (Almazrouei et al., 2020; Holt et al., 2017).

Constructive relationships and adequate support are primary factors associated with stress (or lack thereof) among criminal justice employees in general (Cullen et al., 1985; Holt et al., 2017; Johnson et al., 2005). Forensic examiners interact and develop relationships with multiple stakeholders, some external to their workplace, such as investigators and lawyers, and some within their workplace (e.g., managers and supervisors (Almazrouei et al., 2019; Dror & Pierce, 2020)). Communications between examiners and top-level management and immediate supervisors occur for various reasons, such as to manage caseload, review cases, verify conclusions or reach resolutions in disputed conclusions (Mustonen et al., 2015). These interactions can be a source of stress but can also be supportive and reduce stress. For example, it has been identified that the higher the level of perceived management and supervisory support, the lower the level of workplace stress (Holt et al., 2017).

This chapter develops the findings from Chapter 3 by identifying the perceived sources of workplace stress, along with considerations of whether examiners receive support from management, and whether examiners believed the stress they experienced affected their judgements.

4.2 Method

4.2.1 The Questionnaire

A questionnaire was designed to record the feelings experienced by forensic examiners regarding workplace stress and support, in a similar manner to previous studies addressing the perceptions of workplace stress (e.g., Burruss et al., 2018; Holt et al., 2017; Holt & Blevins, 2011). The questionnaire contained 10 questions about the sources of stress (questions 1–3, 6–10) and about support from management (questions 5 and 6). These questions required the examiners to rank their responses on a seven-point Likert-type scale.

An additional question was included that linked stress to the decision-making of forensic examiners: 'In your opinion, are your own judgements influenced by stress?' For this question, examiners could answer 'yes', 'no' or 'don't know'. This is the same question asked by Kukucka et al. (2017), but the term *cognitive bias* was replaced by *stress*. The examiners were also asked to provide demographic information about their field of expertise, sex, years of experience and whether they were active in casework or retired.

4.2.2 Participants

In total, 41 forensic examiners from two forensic laboratories participated in this study. The mean years of experience for the forensic examiners was 14.4 (SD = 8.2; range = 2 to 31). The experience of participants was categorised in groups of comparable sizes (see Table 4-1). Forensic examiners reported that they worked within 11 primary fields of expertise. For the analysis by field, the reported fields were categorised into one of two broad categories: crime scene examination (n = 11, 27%) or analytical (n = 19, 46%, i.e. fields that primarily have analytical casework within the forensic laboratory, which include document examination, firearms examination, DNA, fingerprint examination and chemical criminalistics). A few (n = 3, 7%) forensic examiners did not report their field or reported that their primary field did not fall into any of the two broad field categories (n = 8, 20%), and so these examiners were not included in the analysis by field of expertise (see Table 4-1). Both descriptive and inferential statistics were applied to measure the reported stress and support levels. Unless otherwise clarified, the assumptions for the statistical tests used were assessed and fully met.

Table 4-1: Demographical information of participants.

	n	Valid%
Work Status		
Active	38	93
Retired	0	0
Did not report	3	7
Sex		
Male	18	44
Female	22	54
Did not report	1	2
Years of Experience*		
1–6	7	17
7–10	7	17
11–15	8	20
16-20	7	17
>20	8	20
Did not report	4	10
Field of Expertise*		
Crime scene examination	11	27
Document examination†	3	7
Firearms examination†	3	7
DNA†	4	10
Fingerprint examination†	8	20
Chemical criminalistics†	1	2
Facial recognition‡	3	7
Forensic medicine‡	1	2
Fire investigation‡	2	5
Digital investigation‡	1	2
Imaging‡	1	2
Did not report	3	7

*The percentages do not add to exactly 100% due to rounding.
† Analytical examiners.
‡ primary field do not fall into any of the two broad field categories.

4.3 Results

4.3.1 Workplace Stress and Support

The mean response to each question addressing the feelings of stress encountered or support provided in the workplace is shown in Table 4-2. Figure 4-1 illustrates the reported feelings of stress and the support the forensic examiners received. The widest variations were observed in the feelings respondents had in terms of management support, (questions 4 and 5), where 50% of the data were between scores 2 and 5, with additional responses ranging from the extreme low score of 1 to the extreme high score of 7.

When converting the whole data set (n = 402) into standardised z-scores, 13 data entries (3.2%) had absolute z-scores between 1.96 and 3.29 (no absolute z-scores were above 3.29). The obtained z-score percentages were lower than the suggested cut-offs, as outlined by Field (2018; see Table 4-3 for details). Hence, no further statistical treatment, such as exclusion of outliers, was required (Field, 2018).

Table 4-2: Means and standard deviations for the 10 questions on workplace stress and support.

Question	***M (SD)***
1. How often do you feel generally stressed?	3.61 (1.26)
2. How often do you feel stressed at work?	3.85 (1.39)
3. How often do you feel stressed because of management/supervisors?	3.95 (1.47)
4. Do you feel that your management is concerned with your wellbeing?	3.85 (1.81)
5. Do you receive support from your management?	3.98 (1.86)
6. How often do you feel stressed from backlogs and the need to do many cases?	3.43 (1.55)
7. Was the source of stress related to the nature of cases (e.g. terrorism, murder, rape)	1.87 (1.11)
8. Was the source of stress related to high-profile cases (i.e. media coverage)	1.97 (1.31)
9. Was the source of stress related to the circumstances at your work (e.g. pressure exerted by investigators/prosecution, competition with colleagues)?	2.88 (1.70)
10. Was the source of stress related to personal reasons?	2.70 (1.29)

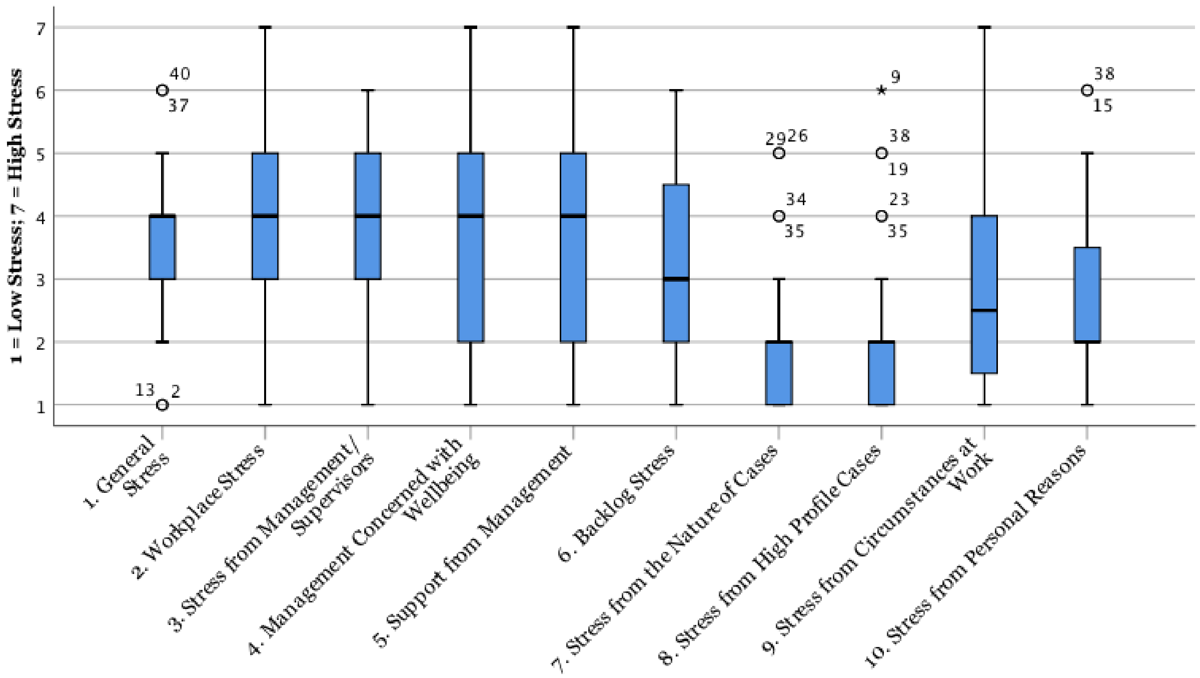

Figure 4-1: A box plot for questions 1–10 on workplace stress and support.

Histograms and Q-Q plots were assessed to confirm that the data were normally distributed for each of the 10 questions, which was the case for all questions except for questions 7, 8 and 9 (where the data were skewed). Hence, non-parametric tests (e.g., Mann-Whitney U) were used in the analysis of these three questions.

In a manner akin to Yoo et al. (2013), a stepwise multiple regression analysis was run to develop a model that predicted the general stress (question 1) of forensic examiners. Specifically, backward stepwise regression was chosen for this analysis because it provided a regression model with only the significant predictors (the insignificant predictors are removed from the model without having a substantial effect on how well the data fit the model) and because it is more preferable than forward regression (Field, 2018). Of all the predictors (questions 2–10), only workplace stress (question 2, $B = 0.714$, $SE_B = 0.076$, $\beta = 0.786$, $p < 0.001$) and personal stress (question 10, $B = 0.303$, $SE_B = 0.083$, $\beta = 0.305$, $p = 0.001$) were statistically significant predictors of general stress in the model, $F(2, 37) = 54.203$, $p < 0.001$, adj. $R^2 = 0.732$ (see Figure 4-2).

Table 4-3: Percentages of standardised z-scores to objectively assess for outliers.

Absolute z-scores	**% cut-offs (Field, 2018)**	**Current study (%)**
Greater than 1.96	≈ 5%	3.2%
Greater than 2.58	≈ 1%	0.7%
Greater than 3.29	0%	0%

Given that in the first model workplace stress was a stronger predictor of general stress than stress due to personal reasons (β of 0.786 *vs.* 0.305, respectively), another series of backward stepwise regressions was run to develop a second model to predict workplace stress (thereby excluding personal reasons (question 10) and general stress (question 1) as predictors in this second model). Stress from case backlogs and the need to do many cases (question 6, $B = 0.431$, $SE_B = 0.107$, $p < 0.001$) and stress from managers or supervisors (question 3, $B = 0.407$, $SE_B = 0.120$, $p = 0.002$) were the only significant predictors in model 2, $F(2, 35) = 21.262$, $p < 0.001$, adj. $R^2 = 0.523$. The two stress factors were of comparable strength in predicting workplace stress (i.e. β of 0.488 *vs.* 0.412, respectively; see Figure 4-2).

Pearson correlations were conducted to test the relationships of management support (questions 4 and 5) with stress from the workplace (question 2) and with stress from managers or supervisors (question 3). No statistically significant relationships were found between management support and either workplace stress or stress from management/supervisors, $p > 0.05$.

4.3.2 Effects of Field, Sex and Experience

Two-tailed t-tests were applied to determine if there were differences in stress and support levels between participants within each field category and between male and female examiners. The Mann-Whitney U test was used for questions 7, 8 and 9. Crime scene examiners (mean rank = 21.05) reported feeling significantly more stressed than analytical examiners (mean rank = 11.31) as a result of the nature of the cases that they were dealing with (question 7; $U = 32.50$, $z = -3.27$,

p = 0.002, r^2 = 0.37). Similarly, the score for personal reasons as a reported source of stress (question 10, approaching significance; $t(30)$ = -1.98, p = 0.057, d = -0.75, 95% CI[-1.84, 0.03]) was higher for crime scene examiners (M = 3.27, SD = 1.27) compared with analytical examiners (M = 2.37, SD = 1.17; see Figure 4-2). The responses to the remaining questions did not significantly vary by field of expertise (i.e., all at p > 0.05).

Female forensic examiners reported feeling more stressed in general (question 1, M = 4.27, SD = 1.08; $t(40)$ = 4.26, p < 0.001, d = 1.36, 95% CI[-0.76, 2.12]) and at the workplace (question 2, M = 4.45, SD = 1.10; $t(40)$ = 3.12, p = 0.003, d = 0.99, 95% CI[0.43, 2.03]) relative to male examiners (M = 2.83, SD = 1.04 and M = 3.22, SD = 1.40, respectively). However, the sources of stress (questions 3 and 6–10) and view of management support (questions 4 and 5) did not significantly differ between female and male examiners (i.e., p > 0.05).

A one-way ANOVA was conducted to determine if the perceived levels of stress and support were different for the different years of experience groups. A Kruskal-Willis H test and post hoc analysis (with the Bonferroni correction for multiple comparisons) were used for questions 7, 8 and 9. When reported stress levels varied significantly across experience groups, it was due to circumstances at work (question 9, $\chi^2(4)$ = 14.16, p = 0.007, η_H^2 = 0.32) or personal reasons (question 10, $F(4, 32)$ = 2.81, p = 0.042, η_p^2 = 0.26). The reported stress levels resulting from workplace circumstances were higher for 11–15 years of experience (mean rank = 29.69) than for 7–10 years of experience (mean rank = 10.14), with an adjusted p = 0.004. No statistically significant variations were found among the experience groups for reported stress from personal reasons (p > 0.05, post hoc [Bonferroni]). Univariate analysis of variance showed no significant interactions between field, sex and experience for any of the 10 questions, p > 0.05.

4.3.3 Stress and Decision-Making

Forensic examiners were divided on whether they thought their judgements were influenced by stress; 39% (n = 16) answered ‘yes’ to this question, while 22% (n = 9) answered ‘no’, and the rest of examiners (39%, n = 16) were unsure. Responses did not vary significantly by field (p = 1.000, Fisher’s exact test), sex (p = 0.722, Fisher’s exact test) or experience (p = 0.517, Fisher’s exact test).

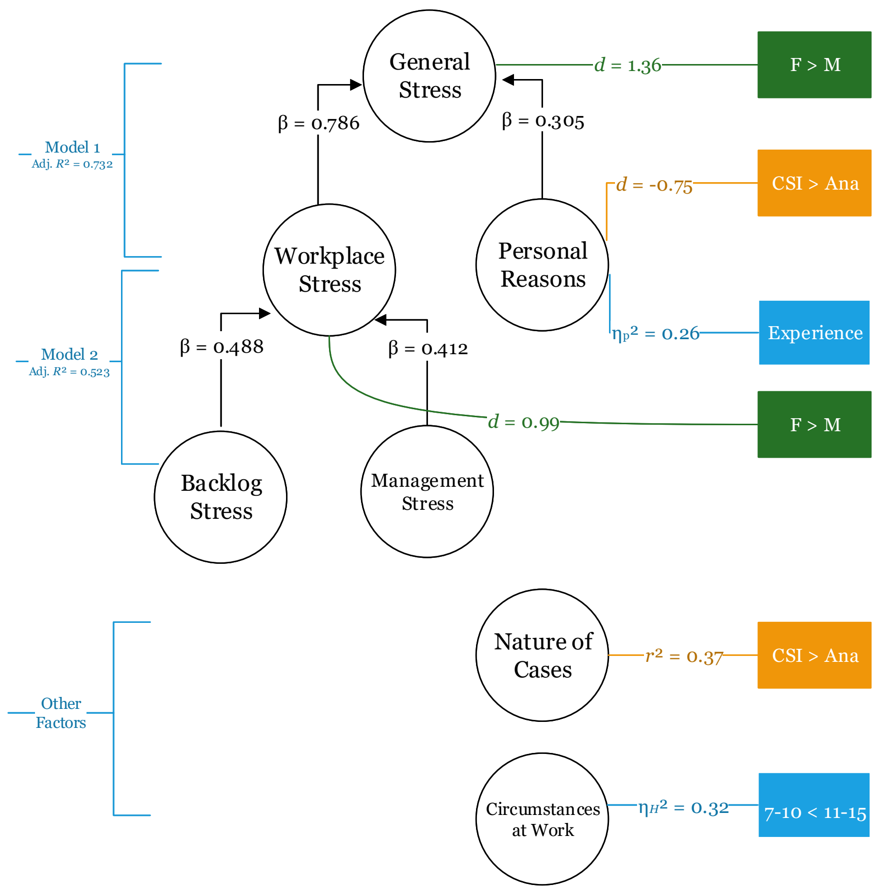

Figure 4-2: A summary of the results showing the significant findings at an alpha level of 0.05. Regression models 1 and 2 (Adjusted R^2; standardised β); stress by field of expertise (orange box; CSI = crime scene investigation field; Ana = analytical field; Cohen's *d*; Mann-Whitney U r^2); stress by sex (green box; F = female; M = male); and stress by experience (light blue box; one-way ANOVA η_p^2; Kruskal-Willis H η_H^2).

4.4 Discussion

4.4.1 Workplace Stress and Support

Forensic examiners reported a range of feelings of stress and views of support levels (low to high scores in all the questions; see Figure 4-1). On average, examiners reported feeling a moderate level of stress in general (question 1) and at the workplace (question 2). Findings in the published literature have suggested that the wellbeing and performance of an individual is optimum at moderate stress levels and deteriorates at either high or low stress levels (Benson & Casey, 2013; Yerkes & Dodson, 1908). It should be noted that questions 1, 2, 3 and 6 included the term 'often' which relates to the *frequency* of stress, but the responses can also reflect the *level* of stress. Hence, it can be considered reasonable to assume that examiners who are stressed more frequently also feel higher stress levels (Almazrouei et al., 2020; Epel et al., 2018).

Stress deriving from workplace and personal factors were significant predictors of the reported general stress of forensic examiners. The first regression model, containing these two factors alone, accounts for 73.2% of the variability in the general stress of examiners (see Figure 4-2). Additionally, reported stress from the workplace was 2.5 times stronger than personal reasons as a predictor of general stress. This finding suggests the workplace environment and culture where forensic examiners operate is an important factor in the general wellbeing of forensic examiners.

Female examiners reported feeling more stressed than male examiners from both general stress and workplace stress. Previous research reported women can experience higher stress levels than men at the workplace for reasons such as having additional family responsibilities outside the workplace (Sharma et al., 2016) and differences in coping styles (Matud, 2004). However, the data of this current study did not identify the specific sources of stress that influence female examiners differently to male examiners (i.e., $p > 0.05$ for questions 3, 6–10). Therefore, future research could usefully investigate the variability of causes of perceived stress.

Given the importance of understanding the contributing factors to workplace stress, a second regression was run. Model 2 identified management and case backlog as factors that were significant predictors of perceived stress, accounting for 52.3% of the variability of perceived workplace stress. These two factors were also found to contribute more to the *high* stress levels felt by forensic examiners than personal reasons (Almazrouei et al., 2020). This is unsurprising given that stress caused by mangers/ supervisors and case backlogs are common organisational-level sources of stress that are documented in other domains outside forensic science (Cooper & Marshall, 1976; Jeanguenat & Dror, 2018).

Stress that arises from outside the work environment, such as from personal reasons, can affect performance at the workplace and vice versa (Bell et al., 2012; Sok et al., 2014). In this study, perceived stress as a result of personal factors (such as financial and family issues) was a significant predictor of feelings of general stress, and crime scene examiners reported higher stress levels from personal reasons (albeit, approaching significance) compared with analytical examiners (see Figure 4-2). Previous research found that shift work was a major source of stress to crime scene examiners, as it impacts their inability to make plans and keep commitments in their personal life (Kelty & Gordon, 2015). In addition, stress from personal reasons varied with years of experience; however, it should be emphasised that the number of years of experience that a forensic examiner has, can also be correlated to age (e.g., Patterson, 2003), which also correlates to other variables. Hence, it is not possible to attribute the findings to experience *per se* as it may be due to a correlation to other factors rather than causation.

Neither the nature of cases nor working in high-profile cases were reported to be major sources of stress (see low mean scores in questions 7 and 8, Table 4-2). Field-specific differences were found in reported levels of stress from the nature of cases, such as working at murder scenes, where crime scene examiners felt more stressed than analytical examiners. Typically, analytical examiners are not exposed to stress elements from a crime scene, such as bloody scenes (Jeanguenat & Dror, 2018), or stress from managing critical decisions at a crime scene under time pressure (Helsloot & Groenendaal, 2011). These differences in working environments and tasks may provide insights into why crime scene examiners reported feeling more stressed than analytical examiners working on the same type of case.

Similarly, stress from circumstances at work, such as feeling pressure from investigators or prosecutors or enduring competition from colleagues, was relatively low (question 9). Post hoc analysis revealed differences between the years of experience groups. Examiners in the 11–15 years of experience group felt more pressure as a result of circumstances at work than examiners in the 7–10 years of experience groups. This may be a result of the differences in roles and responsibilities for examiners who have more experience, or related to other correlated factors, such as age (see above).

Relationships in the workplace, including managerial and supervisory support, can be important factors related with stress (Cullen et al., 1985; Holt et al., 2017; Johnson et al., 2005). In this study, on average, forensic examiners reported feeling that their management was moderately concerned with their wellbeing and that they received moderate support from management (see Table 4-2). Management support (questions 4 and 5) was not a significant predictor of either the general stress (question 1) or workplace stress (question 2) reported by forensic examiners. Also, the correlation between the findings from these four questions were insignificant. In contrast, a previously published study found management and supervisory support were significant predictors of reduced stress and increased job satisfaction (Holt et al., 2017). The different findings may be due to different working environments in different laboratories.

4.4.2 Stress and Decision-Making

Examiners were divided as to whether stress affected their judgments. Some forensic examiners (39%) felt that stress affected their judgements. To enable clear and transparent forensic science judgments, it has been argued that having a decision-making environment that manages the risks of stress (National Institute of Justice, 2019) and uncertainties (Georgiou et al., 2020; Morgan et al., 2018; see also Dror and Pierce (2020) for quality control and risk management) is important. However, the findings from this study are derived from self-reporting responses, and such responses are outputs of highly complex cognitive information and processing (Gardner et al., 2019; Nisbett & Wilson, 1977). A perception that stress may have influenced a judgement does not necessarily mean the decisions and conclusions made have been influenced by a single stressor or combination of stressors (Almazrouei et al., 2020).

It is worthy of note that extensive empirical research from other domains indicates that stress influences expert decision-making (Akinola & Mendes, 2012; Arora et al., 2010; Corner & Gill, 2019; Gok & Atsan, 2016; Yu et al., 2015). However, in this study, some examiners in the forensic services domain (22%) said that stress did not bias their own judgements. Different explanations may exist for this finding. It could mean examiners do not think their judgements are influenced at all — with or without stress. Such a bias blind spot has been identified when an expert does not believe context (including bias from stress) affects their *own* decision-making and conclusions, but that it can affect others (Kukucka et al., 2017; Page et al., 2012). Alternatively, this finding could indicate that stress does not affect the decisions of examiners perhaps due to examiners being more attentive when stressed. This would be a fruitful area for further research that addresses the multivariate complexity of the impact of stress on decisions within crime reconstructions.

The findings of this study provide insights into the sources of stress for forensic examiners, their feelings on the support they receive in the workplace, and their perceptions of whether stress affects their judgements. However, it is important to consider the findings in this study with caution, due to limited statistical power from the relatively small sample size of forensic experts. In addition, it is important to note that this study includes data from more than one laboratory which may have potentially introduced confounding factors. This is because each has its own working culture and work practices (such as case backlogs and managerial support), and also due to the variations in the demographics of expert participants recruited from each lab.

4.5 Conclusion

This study surveyed forensic examiners working in different fields of forensic expertise and with different years of experience on their feelings of stress and support in the workplace. On average, examiners reported feeling moderate stress levels. Workplace and personal stress factors were significant predictors of general stress. Stress from management and/ or supervisors and case backlogs were significant predictors of workplace stress. Management support was not a significant predictor and was not associated with either general stress or workplace stress.

Feelings of stress that arise as a result of the type of case, from working in high-profile cases, and from circumstances at work (such as enduring pressure by investigators or prosecutors) was relatively low (mean scores of these stress factors were below 3 (out of 7)). Crime scene examiners reported feeling higher stress than analytical examiners from personal reasons and from the nature of cases they were involved with. Male examiners reported feeling less stressed than female examiners from both general stressors and workplace stressors. Examiners within the 7-10 years of experience group reported feeling less stress due to circumstances at work than those within the 11-15 years of experience group at work.

Going forward, gaining a greater understanding of the positive and negative impacts of stress, and the feelings examiners experience of stress in the workplace will be highly valuable for the development of a working culture that addresses the negative impacts of stress on forensic science examiners and their judgements.

In this study, examiners were divided by their opinion on whether stress affected their judgements. There are different plausible explanations for this, but it is evident that the impact of stress on forensic science decision-making should be explored further. An experimental approach offers a more objective assessment of the possible impact of stress on forensic science judgments, compared with self-reported data as in Chapter 4. Hence, Chapters 5 and 6 present data that was collected to address this, thereby contributing to objective (b) of the thesis. A new stress-inducing method was developed and presented in Chapter 5. Then, the possible impact of this stressor on a forensic science task was explored in Chapter 6.

Chapter 5 A Method to Induce Stress in Human Subjects in Online Research (Study 3)

5.1 Introduction

Generating stress in human subjects for research can be a challenging task (Ferreira, 2019). This is because, on the one hand, the experimental design needs to effectively generate stress but, on the other hand, avoid long-term effects on the participants (Ferreira, 2019). Adding to this challenge is the variability in how individuals perceive and react to the same stress factor (Epel et al., 2018; Lazarus & Folkman, 1984).

It has been observed that using only participants that can attend and participate in a study in person can have an impact on the diversity of the participant sample (Upadhyay & Lipkovich, 2020). Added to this, the value of being able to carry out online experiments has been highlighted particularly during the coronavirus pandemic (Wigginton et al., 2020) when much of the face-to-face research involving human subjects was paused worldwide. There has therefore been growing recognition of the value of creating opportunities for studies to be delivered online rather than face-to-face, including stress-inducing studies (Kirschbaum, 2021).

A meta-analysis of 208 laboratory-based stress studies found that the combination of social–evaluative threats (when one is judged negatively by others, such as receiving negative feedback) and uncontrollability (when nothing can be done to avoid negative consequences or change a situation, such as having a time limit for completing a task) were the stress factors that produce the greatest stress response in human subjects (Dickerson & Kemeny, 2004). Therefore, methods that combine social–evaluative threats and uncontrollability elements, such as the Trier Social Stress Test (TSST; Kirschbaum et al., 1993), considered the "gold standard" for inducing experimental stress in human subjects (Allen et al., 2017; Le et al., 2020), have potential for effectively inducing stress in an online setting.

Several studies have been conducted to try and validate online versions of TSST, delivered through virtual reality tools (e.g., Zimmer et al., 2019), and more recently delivered by video conferencing online (Eagle et al., 2021; Gunnar et al., 2021; Harvie et al., 2021). However, some of these internet-delivered studies did not include a control group (Eagle et al., 2021; Gunnar et al., 2021), which limits the opportunity to understand and interpret the outcomes of the stress manipulation, for example, by not accounting for potential additional psychological stress as a result of video conferencing (Riedl, 2022). One study included a control group (Harvie et al., 2021), but required the (virtual) presence of at least three experimenters (i.e., the researcher and two panellists) in each video conferencing session, which limits online stress studies to live tasks in which the presence of the researchers is required nevertheless (virtually rather than in-person).

Therefore, in this study, alternative stressors were considered that combine social–evaluative threats and uncontrollability yet were still feasibly operationalised in an internet-delivered environment without the need of the researchers to be present. One such stressor is the Trier Mental Challenge Test Stress Protocol originally developed by Kirschbaum et al. (1991)—referred to here as the 'Mental Challenge Test'. In the Mental Challenge Test, participants are asked through programmed software to answer a number of arithmetic questions without a calculator under a time limit and receive feedback, such as "wrong" for incorrect answers (Kirschbaum et al., 1991). The studies that utilised the Mental Challenge Test were computer-assisted, yet, to date they have been conducted in the presence of the researchers (Allendorfer et al., 2014, 2019; Dedovic et al., 2005; Kirschbaum et al., 1991).

This study presents a method that has been developed for inducing stress in an online setting, without the presence of researchers (either in-person or virtually). This method may enable advancements in stress research, by accessing large number of international participants rapidly and in a cost-effective manner. In this method, participants were asked to answer a number of general knowledge and mathematical questions selected specifically for this study under stress conditions of social evaluative threats (such as displaying negative feedback) and uncontrollability (such as imposing time limits).

5.2 Method

5.2.1 Participants

Data were collected from 120 participants through the Prolific platform in a single session. Two participants in the stress group withdrew their data and were excluded from analysis. The final sample consisted of 118 participants, of whom $N = 66$, 56% were in the control group and $N = 52$, 44% in the stress group (see Table 5-1). Thirteen participants dropped out ($n = 11$ from the stress group and $n = 2$ from the control group). A drop-out is counted when a participant starts answering the mathematical and general knowledge questions then drops out by exiting the study.

Table 5-1: Demographical information of participants.

	Mean (SD)	Range
Age	33.3 (7.0)	25-59
	n	Valid%
Sex		
Male	58	49.2
Female	60	50.8
Highest Degree Completed		
High school diploma/ A-levels or equivalent	18	15.3
Technical/ community college	9	7.6
Undergraduate degree (BA/BSc/Other)	46	39.0
Graduate degree (MA/MSc/MPhil/Other)	37	31.4
Doctorate degree (PhD/Other)	6	5.1
Other*	2	1.7

*The two participants reported PGCE (postgraduate certificate in education) as their highest completed education. Their data were coded within the 'graduate degree' holders, since PGCE is an advanced education after the bachelor's degree.

5.2.2 Stress Procedure

Participants signed the consent form and were then given instructions about the exercise (see Figure 5-1). The consent form and instructions were carefully written to offer fully informed consent, but without revealing the specific aim of the study (i.e., inducing stress to participants). Then, participants were randomly allocated into either the stress or the control group through Qualtrics. The stress group were shown a warning message that performance was being monitored. They were then asked to answer a block of eight random mathematical/general knowledge questions with time limits and with feedback given (i.e., Stress Block A; see Appendices B-1, B-2 and B-3 for further details on the feedback messages and mathematical/general knowledge questions). If a participant answered a question incorrectly, a "**WRONG!**" message in red would appear immediately on the screen. Conversely, a neutral "**OK**" message appeared in grey if a question was answered correctly. If the time allocated to the question ran out, a "**TIME OUT!**" message appeared in red.

At the end of the mathematical/general knowledge question block, either a neutral message or a negative message was given to participants, depending on their performance (compared to a preset criterion score of three correct answers). If the participant scored three correct answers or lower in this block, then a negative message would appear explicitly comparing the individual score with those of other participants. This had the potential to further increase the social evaluative threat component of stress (Dickerson & Kemeny, 2004; Kirschbaum et al., 1991). If the participant scored four or more questions correctly in this block, a neutral message would appear that had no reference to individual or group performance. This approach was repeated in two more blocks (i.e., Stress Blocks B and C). The control group was asked to complete a comparable number and genre of questions but without feedback or a time limit. Questions were randomised through Qualtrics. To prevent and detect cheating or random responses, a range of quality assurance measures were included, such as adding a commitment statement, including a tool to detect potential bot responses and attention check questions (see Appendix B-1).

After three blocks of mathematical/ general knowledge questions, the participants were asked to complete the state anxiety scale (Spielberger et al., 1983) and a visual analogue scale on stress, referred to as 'VAS-stress' scale from here onwards. Next, participants were asked to provide their demographic information of age, sex and their highest level of education. Participants were then asked to complete the trait anxiety scale (Spielberger et al., 1983). At the end of the experiment, participants were debriefed that this study specifically aimed to induce momentary stress. In the debrief, participants were given the opportunity to withdraw their data without giving a reason and without it affecting the rights and benefits (such as payment) to which they were entitled, or it having any negative repercussions for them.

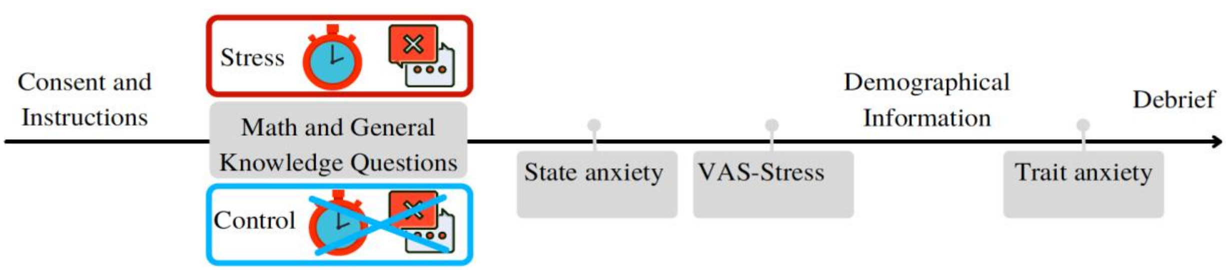

Figure 5-1: Graphic timeline of the experimental procedure.

5.2.3 Stress Manipulation Check

The effectiveness of the stress manipulation was assessed using two self-reported measures. First, to capture the situational anxiety levels of participants (i.e., the anxiety feelings in the present moment; see Appendix B-4), the state scale of the State–Trait Anxiety Inventory (STAI) was used (Spielberger et al., 1983). This state anxiety scale is a validated and commonly used measure for various stress manipulations (Arora et al., 2010; LeBlanc et al., 2005; Spielberger et al., 1983; Tanida et al., 2007). The scale consists of 20 statements (e.g., I feel nervous) for which users indicate their degree of agreement on a 4-point scale, in regard to how they feel 'right now' (score range is from 20 to 80 (Spielberger et al., 1983)). Second, following the approach of Le et al. (2020), participants were asked to report their stress levels on a VAS-stress, retrospectively: "Looking back, how stressed did you feel throughout answering the mathematical and general knowledge questions?" The participants rated their feelings from 0% (not stressed at all) to 100% (extremely stressed).

5.2.4 Trait Anxiety

Participants were also asked to complete the STAI trait anxiety scale (Spielberger et al., 1983; see Appendix B-5) to ensure that the background anxiety levels of participants do not confound the reported state anxiety or VAS-stress levels. The trait scale consists of 20 statements that measure how people 'generally' feel (score range from 20 to 80). The STAI manual recommends placing the trait anxiety scale, after the state anxiety scale if both scales are administered together, because the former measures a more stable anxiety construct that should not be affected with situational stress (Spielberger et al., 1983). Accordingly, the trait anxiety scale was placed at the end of the experiment.

5.3 Results

5.3.1 Overall Stress and Trait Anxiety

The mean stress levels, as measured by the state anxiety scale, was significantly higher for the stress group (M = 48.89, SD = 13.01) than for the control group (M = 34.35, SD = 10.66), M = -14.54, 95%CI [-18.85, -10.22], $t(116) = -6.67$, $p < .001$, Cohen's d = -1.24. In addition, participants in the stress group (M = 73.17, SD = 24.01) reported higher VAS-stress ratings than the control group (M = 30.55, SD = 22.90). This was also a statistically significant difference, M = -42.63, 95%CI [-51.22, -34.04], $t(116) = -9.83$, $p < .001$, d = -1.82. On average, the stress (M = 45.79, SD

= 11.30) and non-stress groups (*M* = 41.58, *SD* = 12.37) were comparable in terms of their background stress (i.e., trait anxiety levels), *M* = -4.21 , 95%CI [-8.59, 0.16], *t*(116) = -1.91, *p* = .059, *d* = -.35.

5.3.2 Trait anxiety as a Stress Moderator

Two linear regression models were run to investigate whether the trait anxiety or the demographical variables (i.e., age, sex and education) moderated the reported state anxiety or VAS-stress scores. In both models, the trait anxiety was the only factor ($p < .001$) that moderated the dependent variables. In addition, trait anxiety was significantly correlated with both state anxiety ($r(118) = .55$, $p < .001$) and VAS-stress scale ($r(118) = .33$, $p < .001$).

Hence, it was necessary to account for trait anxiety, as a background stress, to further understand the effectiveness of the online stressor presented here. To do so, participants were divided into three homogenous groups in terms of reported trait anxiety levels: low, moderate and high anxiety (this approach is similar to Horikawa and Yagi (2012)). The high anxiety group (*N* = 35; *n* = 15 in the control condition and *n* = 20 in the stress condition) were those whose trait scores were 0.5 *SD* above the mean trait score of 43.43 (*SD* = 12.04). Conversely, the low anxiety group (*N* = 40; *n* = 27 in the control condition and *n* = 13 in the stress condition) were those whose trait scores were 0.5 *SD* below the mean trait score. The rest of participants (*N* = 43; *n* = 24 in the control condition and *n* = 19 in the stress condition) were classified to have moderate trait anxiety levels.

The state anxiety levels varied significantly between the stress and control conditions, in the low anxiety group (*M* = -16.00, 95%CI [-25.77, -6.23], Welch's *t*(13.57) = -3.52, *p* = .004, *d* = -1.19) and moderate anxiety group (*M* = -12.82, 95%CI [-17.75, -7.90], *t*(41) = -5.26, $p < .001$, *d* = -1.61), but not in the high anxiety group (*M* = -7.20, 95%CI [-15.43, 1.03], *t*(33) = -1.78, *p* = .084, *d* = -0.61; Figure 5-2). However, when comparing the VAS-stress scores, there were statistical significant differences in all the three anxiety groups (low anxiety: *M* = -35.24, 95%CI [-54.00, -16.49], *t*(38) = -3.80, *p* = .001, *d* = -1.28; moderate anxiety: *M* = -44.21, 95%CI [-56.30, -32.13], *t*(41) = -7.39, $p < .001$, *d* = -2.27; high anxiety: *M* = -39.87, 95%CI [-54.77, -24.96], Welch's *t*(21.48) = -5.55, $p < .001$, *d* = -1.90). Note that Welch's *t*-test is used when the assumption of homogeneity of variances has been violated, as assessed by Levene's test for equality of variances.

5.3.3 Performance on Stress Blocks

The majority (67.3-88.5%) of participants in the stress group scored 3 correct responses or less in stress blocks A, B and C. This means that those participants received negative feedback after completing those blocks of questions. One participant was able to score 7 of 8 questions correctly in Block C, and no one scored 8 of 8 questions correctly (see Table 5-2).

Table 5-2: Frequency and cumulative percentages of correct responses in Stress Blocks A to C.

Correct Response	**Stress Block A**		**Stress Block B**		**Stress Block C**	
N	*N*	%	*N*	%	*N*	%
0	10	19.2	7	13.5	5	9.6
1	21	59.6	19	50.0	5	19.2
2	10	78.8	10	69.2	13	44.2
3	4	86.5	10	88.5	12	67.3
4	3	92.3	5	98.1	7	80.8
5	2	96.2	1	100	6	92.3
6	2	100	0	100	3	98.1
7	0	100	0	100	1	100
8	0	100	0	100	0	100

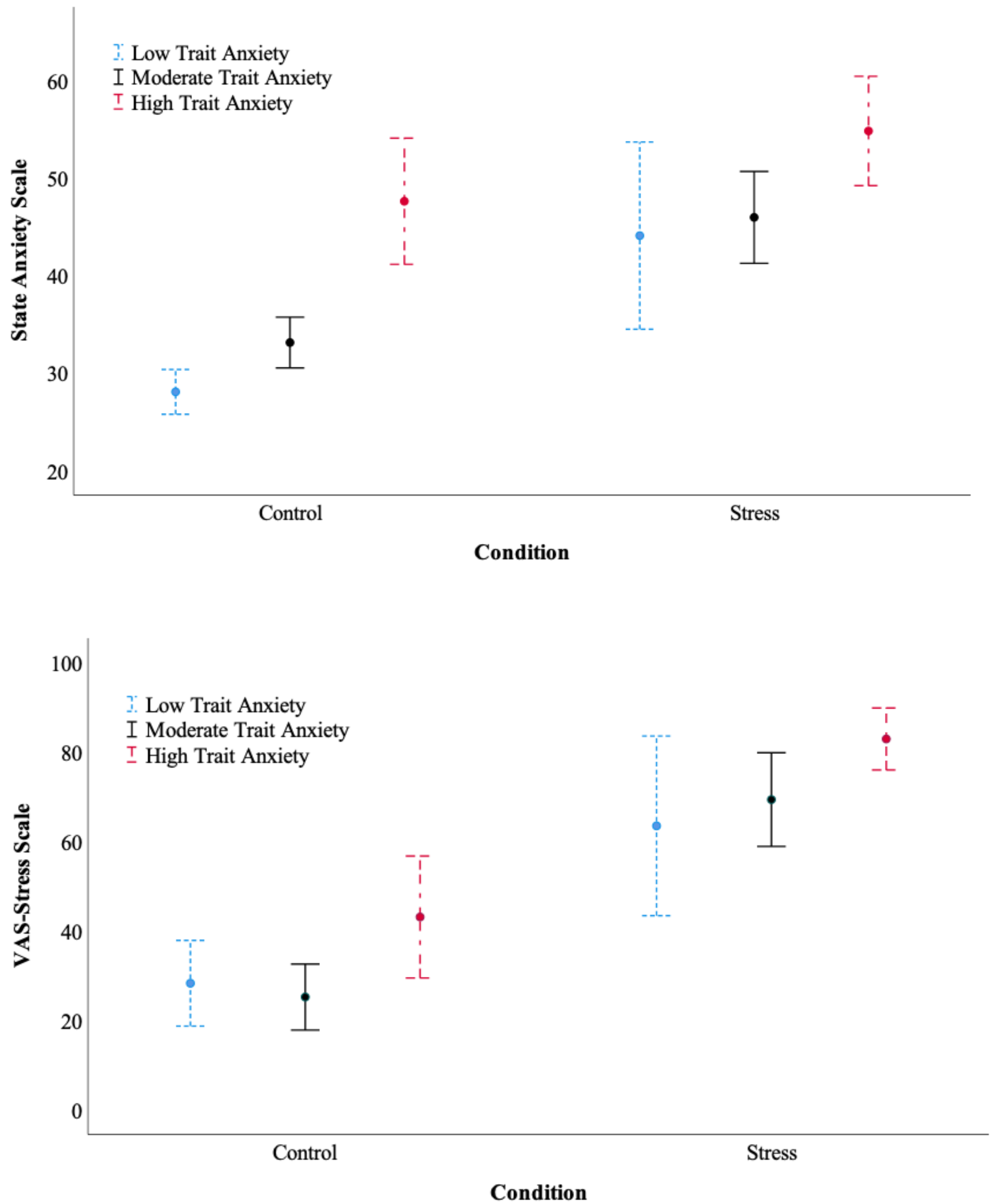

Figure 5-2: Mean state anxiety (top) and VAS-stress scores (bottom) for low, moderate and high trait anxiety participant groups. Error bars reflect 95% confidence intervals.

5.4 Discussion

The stress manipulation was found to be effective in the sample who participated in this study. The state anxiety and VAS-stress scores were significantly higher for the stress group than the control group, with and without accounting for trait anxiety as a moderator. The exception was the state anxiety levels in the high trait anxiety group. Here, the state anxiety levels in the stress condition were still higher than the non-stress condition, although the difference was not statistically significant. One possible explanation is that the online stress method was not effective enough to induce momentary stress to already highly anxious participants –a clear sign of a ceiling effect.

Directly comparing the findings of this study with published studies on stress-inducing methods can be limited (Narvaez Linares et al., 2020), especially that the online stressors are by their very nature less powerful than classical in-person stress tasks. Variations of TSST in previous research were able to cause elevations in state anxiety and VAS-stress levels comparable to the current stressor, but with smaller sample sizes. For instance, Guez et al. (2016) and Le et al. (2020) reported large effect sizes of their stressors on state anxiety ($\eta^2_p = 0.23$, $N = 46$) and VAS on stress ($d = 1.74$, $N = 76$), respectively. This difference in magnitude is likely to be due to a number of factors that may include the absence of researchers during the stress inducing period. Notably, however, findings of the current study appear to be more in line with the impact of established stressors that had minimal interactions of investigators during the stress manipulation (Dedovic et al., 2005; see Discussion in p. 325).

The stress stimuli selected for this study appear to be challenging since most participants scored 3 or less questions correctly. Thus, the selected stress stimuli made it possible to give negative and potentially stressful feedback to participants in all three stress blocks. It may also be inferred from the data that engagement of some participants in answering the questions in the stress blocks may have been sustained (e.g., some participants were able to score four, five, six or even seven questions correctly in a block, all of which were above the pre-set criterion score of three (see Table 5-2). However, the possibility that this procedure might lead to reduced engagement in some participants cannot be ruled out. Future studies should incorporate a consideration of whether low engagement/motivation might influence scores if, for example, a cognitive task was used after the stress induction.

The higher drop-out rate in the stress condition compared with the control condition could be due to a number of factors, namely the stress manipulation effectively causing stress and thus reduced motivation to complete the difficult tasks. The drop-out rate in this study appears to be higher than other validated stress methods. For instance, in a recent TSST method that was delivered by Zoom, one of 72 participants discontinued the study during the stress period (although it is worth noting that a total of 31 participants dropped out by the end of the experiment for other reasons, such as not showing up in scheduled sessions (Eagle et al., 2021)).

Participants recruited through crowdsourcing platforms, as in the current study, appear to have a higher dropout rate than in-person/offline studies (Stewart et al., 2017; Zhou & Fishbach, 2016). This may be due a range of factors including participants having the ability to preview the study (Stewart et al., 2017), and potentially returning the study before completing the tasks and without affecting their reputation score on the crowdsourcing platforms (Palan & Schitter, 2018). Furthermore, there may be fewer barriers to dropping out of an online study due to the anonymity afforded by the online setting in comparison to dropping out of a live study (in person, or online but with a video connection with the researchers). In addition, researchers may not be aware of participants who have dropped out as they do not count towards the quota allocated in a crowdsourcing platform, and thus researchers under-report them in published papers (Zhou & Fishbach, 2016).

Importantly, drop-outs can be condition-dependent, for reasons such as experiencing more mental fatigue in one condition compared to the other (Zhou & Fishbach, 2016). Though selective attrition can potentially influence internal validity, it is not likely that this caused a meaningful impact on the findings of this study, because the remaining randomised sample sizes in each condition for the method validation were reasonably comparable (i.e., 56% in comparison to 44%). Nevertheless, it may be beneficial for studies that use crowdsourcing platforms to include proactive countermeasure strategies (e.g., telling participants upfront that dropping out could affect the quality of data (Reips, 2000; Zhou & Fishbach, 2016)).

A number of limitations do exist in regard to using this online stress method that should be addressed in future studies. First, the findings from this study are based on the assessment of stress from self-reported measures (Nisbett & Wilson, 1977). Future research can include additional physiological measures, such as the approach taken by Harvie et al. (2021) who had participants measure their own heart rate.

Another limitation is that the baseline stress (e.g., via VAS) was not balanced for both groups. It was of concern that placing a VAS before the stress manipulation (so it could be balanced across conditions) could impact feelings and expectations of the participants, and hence impact their performance (e.g., Christensen-Szalanski & Willham, 1991).

Furthermore, as with any remote online study, there is no control over what participants do during the exercise. Despite the effort made by the researchers to control experimental stimuli and set explicit instructions for the exercise, participants are not monitored and may be carrying out other activities while taking part in the study (such as doing the exercise while relaxing on the sofa compared to a desk). Such variations in behaviour in completing the exercise may have the potential to influence the stress levels of participants, as opposed to being solely induced by the stress stimuli themselves.

Nevertheless, this is the first method that has been designed and used to induce stress in human participants effectively online without the presence of the researchers. It offers a cost-effective and easy-to-use method to induce momentary stress to human subjects in a controlled manner in an online setting. In addition, by not requiring the researchers to be agents of stress, the online method also enables quick access to large participant samples globally through crowdsourcing platforms (Peer et al., 2017). The method includes unpredictable social evaluative threats common in everyday life, including those in professional domains, which means it is a method that can offer a degree of ecological validity.

5.5 Conclusion

This chapter presents a new method to stress human subjects in an online setting without the presence of researchers. This method offers a cost-effective way to collect data from a diverse range of participant cohorts, which is particularly useful in situations where there is a need to carry out research in online environments. The building blocks of this method (such as having specific measures to enhance data quality collected) could be useful for in a wide range of studies that aim to collect quality psychological data online.

Having developed an online method that has been found to be effective in inducing stress in human participants, the next step is to evaluate the impact of this stressor on forensic science decisions. Chapter 6 utilises the developed stressor to study its possible influence on fingerprinting tasks (fingerprint comparisons were chosen as an example of pattern recognition tasks, like handwriting analysis and bullet comparisons), to further addresses objective (b) of the thesis.

Chapter 6 The Impact of Stress on Fingerprint Assessments: Novices *vs*. Experts (Study 4)

6.1 Introduction

Workplace stress has an impact on the quality of decisions made by professionals in a variety of expert domains, from healthcare (Arora et al., 2010) to policing (Akinola & Mendes, 2012). However, research discussion on the potential impact of stress on *decision-making* in forensic science has only recently been considered (e.g., Almazrouei et al., 2021; Jeanguenat & Dror, 2018). Stress has "clear implications for professions that are characterised by high levels of work pressure and intense cognitive demands" (Deligkaris et al., 2014, p. 118), so a consideration of the implications of stress upon forensic examiners is timely (Almazrouei et al., 2020; Helsloot & Groenendaal, 2011). Hence, the aim of Chapter 6 is to present a study on the impact of stress on forensic decision-making.

Several studies have investigated the influence of biasing task-irrelevant information (e.g., Dror & Charlton, 2006; Earwaker et al., 2015; Smalarz et al., 2016) or motivational and emotional factors (e.g., Charlton et al., 2010; Dror et al., 2005; Hall & Player, 2008; Osborne et al., 2014) on decisions in a forensic science context (for a review, see, Kukucka & Dror, 2022). However, there is a lack of research that investigates the impact of stress on forensic decision-making. Since fingerprint evidence is widely used and can carry significant weight in court proceedings (Mustonen et al., 2015), the research reported here considered the impact of stress on a fingerprint decision-making task. The trends that have been identified may well apply and reflect the impact of stress across other forensic domains where pattern recognition tasks are important (e.g., handwriting, toolmarks, etc).

The few studies on fingerprint decision-making that included a stress factor were limited in a number of ways. Some assessed the impact of a stressor (predominantly time pressure) on fingerprint decision-making in approaches that may make the findings not be ecologically valid (Kellman et al., 2014; Stevenage & Bennett, 2017; Thompson & Tangen, 2014; Zou et al., 2021). For example in some studies, the time provided to make a decision was unrealistically short (e.g., two seconds to reach a conclusion (Stevenage & Bennett, 2017)). In forensic settings, there is not often such time pressures (Kellman et al., 2014). Additionally, some of these studies used a two-alternative forced choice experimental paradigm (TAFC; see Bogacz et al., 2006) whereby the participants had to either decide an identification or an exclusion, but were not allowed to reach an inconclusive decision (e.g., Stevenage & Bennett, 2017; Thompson & Tangen, 2014; Zou et al., 2021). In casework, inconclusive decisions are allowed (Dror & Langenburg, 2019). Inconclusives are often considered to be less risky decisions compared to conclusive judgments (Dror & Langenburg, 2019), but they can have practical implications (e.g., potentially not identifying suspects) and should be considered when assessing expert performance (Dror & Scurich, 2020).

There are a number of different approaches used to induce stress on human subjects in research. One approach includes elements of social–evaluative threats, when one is judged negatively by others, such as receiving negative feedback. Another approach is uncontrollability, when nothing can be done to avoid negative consequences or change a situation, such as having a time limit for completing a task (Allen et al., 2017; Dickerson & Kemeny, 2004).

A meta-analysis of 208 laboratory-based stress studies found that stressors that combine the social–evaluative threats and uncontrollability approaches produced the greatest stress response in human subjects (Dickerson & Kemeny, 2004). In addition, it has been suggested that stressors that contain uncontrollable threats to the social self, such as public speaking, can have ecological validity (Allen et al., 2017) as they can occur in daily life (Lehman et al., 2015). Furthermore, they are common across cultures (Dickerson & Kemeny, 2004) and can be unpredictable or uncontrollable, even in professional domains (Akinola & Mendes, 2012; Arora et al., 2010; Schuetz et al., 2008). An example from the medical domain would be to have unexpected external visitors observing the progress of a surgical procedure (Schuetz et al., 2008).

Therefore, the aim of Chapter 6 was to collect data that offer insights into the impact of uncontrollable social evaluative stressors on fingerprint decision-making tasks. The study was comprised of two experiments: the first with novice participants, and the second with fingerprint expert participants. The first experiment acted as a pilot study to test the experimental design with novices before launching the second experiment with fingerprint experts, as well as serving as a comparison to consider the impact of stress on experts relative to novices.

6.2 Method

6.2.1 Fingerprint Stimuli

Prior to the study, 23 fingerprint pairs were chosen from a database of fingerprint pairs where the ground truth was known (i.e., same-source or different-source). The fingerprint pairs were assessed for difficulty by nine fingerprint experts (mean experience, 13.8 years; range, 3-34), in order to choose pairs of varying difficulty for inclusion in the experiments (see Appendices C-1 and C-2). For the difficulty assessment, the fingerprint pairs were presented side by side with a 5-point difficulty scale. A mean rating among the experts of 3.5–5.0 was considered "difficult"; 2.5–3.5 "medium"; and 1.0–2.5 "easy."

Of the 23 piloted pairs, 12 pairs were chosen for the study; six difficult pairs of which three were same-source and three were different-source pairs, and six easy pairs of which three were same-source and three were different-source pairs. The fingerprint pairs were randomly distributed and counterbalanced within Qualtrics by condition and by difficulty, so each participant made assessments of six pairs: three were difficult and three were easy. The aim was to account for the range of difficulty that fingerprint experts encounter in real casework (Kukucka et al., 2020).

Overall, the novice participants in the first experiment made 690 decisions (115 participants × 6 pairs each), half were different-source and half same-source. The control group made 366 decisions and the stress group 324 decisions. In the second experiment, expert participants made 204 decisions (34 participants × 6 pair of prints; 104 different-source and 100 same-source). The control group made 96 decisions and the stress group made 108 decisions (see Appendix C-3).

6.2.2 Stress Manipulation

The stress manipulation involved asking participants to answer 24 general knowledge and mathematical questions under a time limit, and feedback was given (e.g., "**WRONG!"** or "**TIME OUT!"**). Participants in the control group answered a comparable number of general knowledge questions, but without time limits and with no feedback. Furthermore, the questions in the stress condition were selected to be more difficult and prone to error than those in the control group in order to increase the level of stress (by increasing the probability of participants making mistakes and receiving negative feedback). Hence, this experimental design included both social evaluative threat (such as feedback messages after answering each question) and uncontrollability stress elements (such as time pressure for answering the questions), as outlined in Chapter 5.

6.2.3 Stress Manipulation Check

The effectiveness of the stress manipulation was assessed using the state anxiety scale (Spielberger et al., 1983). This established scale captures the situational anxiety levels of participants (i.e., the anxiety feelings at the present moment). This scale consists of 20 statements (e.g., I feel nervous) for which users indicate their degree of agreement on a 4-point scale, in regard to how they feel "right now". The scores range from a minimum of 20 to a maximum of 80 (Spielberger et al., 1983). While this is a self-reporting assessment, the scale has been validated and is commonly used to measure the effectiveness of stress manipulations (Arora et al., 2010; LeBlanc et al., 2005; Spielberger et al., 1983; Tanida et al., 2007).

6.2.4 Attention Check Screeners

Four attention check screeners were used to check that participants paid attention to the study tasks (Oppenheimer et al., 2009). Two of the four attention checks were related to a video on how to make a fingerprint assessment (in the first experiment with novices). Here, participants were asked to summarize the content of the video in two to three sentences. Additionally, the time they spent watching this 5-min, 43-s video was also assessed. The other two attention check screeners were related to completing the state anxiety scale. An additional item was embedded within the questionnaire as an attention check, requesting participants to "please tick somewhat." Furthermore, the pattern of answering the state anxiety questionnaire was checked (e.g., whether a participant consistently stating the same response of "Not at All" in an arbitrary fashion).

6.2.5 Measures

Participants were asked to report a conclusion (identification, exclusion or inconclusive) and the confidence level in their conclusion for each fingerprint pair. To understand the impact of stress on these decisions, the proportions (%) of each category of conclusion was calculated (see Appendix C-3). In addition, response times were recorded in Qualtrics without the knowledge of participants.

6.2.6 *Procedure*

Both experiments followed a between-subjects design, with participants randomly allocated via Qualtrics software into either a stress or a control condition. In each condition, there were three blocks. In each block, the participants made decisions on two pairs of fingerprints after which they answered eight general knowledge and mathematical questions. In the stress condition, these eight questions were difficult, presented with time limits, and feedback was given to participants. In the control condition, the questions were relatively easy, presented with no time limit and no feedback was provided to participants. In total, each participant answered 24 general knowledge/ mathematical questions and made decisions on six pairs of fingerprint. The six pairs of fingerprint varied in difficulty and the ground truth.

After the three blocks of general knowledge and mathematical questions and the pairs of prints, the effectiveness of the stress manipulation was measured with the state scale of State–Trait Anxiety Inventory (see Spielberger et al., 1983). Participants were then asked to provide their demographic information. At the end of the experiment, participants were debriefed and told that this study specifically aimed to induce momentary stress. In the debriefing, participants were given the opportunity to withdraw their data without giving a reason and without it affecting their rights and benefits. Ethical approval was granted by UCL Research Ethics Committee (#15395/003).

Novices in the first experiment received a short training on how to make fingerprint assessments before starting the actual study. This consisted of a five-minute online video tutorial and three exercises on fingerprint assessments in which feedback was given (one identification, one exclusion and one for inconclusive).

6.2.7 *Participants*

In the first experiment with novice participants, the participant selection criteria were 25–60 years of age with a minimum level of high school (or equivalent) education. These parameters were chosen to ensure that the cohort were comparable with that of expert fingerprint examiners and comparable to other studies with forensic experts. For example, Holt et al. (2017) reported the mean age for the 670 forensic examiners they surveyed was 39 years (median = 37, range =

23–66), and a few (6% of sample, $n = 40$) had an education level equivalent to that of a two-year degree or less.

Data were collected from 120 novice participants using the Prolific Academic platform. Five participants were excluded from the analysis (withdrew their data, did not meet the inclusion criteria as they were under the age of 25; or failed most of the attention checks). This left a final sample of 115 novice participants of whom 54.8% were males ($n = 63$; prefer not to disclose the sex: $n = 1$, 0.9%). The mean age of participants was 35 ($SD = 8$; range = 25-60). There were 61 (53%) participants in the control condition and 54 (47%) in the stress condition.

In the second experiment, data were collected from 34 fingerprint experts of whom 38.2% were males ($n = 13$) and 58.8% were females ($n = 20$; prefer not to disclose the sex: $n = 1$, 2.9%). The experts were based in five different countries: Bahrain, Saudi Arabia, the United Arab Emirates, the United Kingdom, and the United States. The mean experience of participants in fingerprint assessments was 17.4 years ($SD = 11.0$; range = 1-35). The mean age of participants was 43 ($SD = 10$; range = 25-57).

It is of note that initially there were 43 expert participants, but nine dropped out, all from the stress condition. This was perhaps a sign that the stress condition was indeed stressful. As a result of the drop-out, more expert participants were assigned to the stress condition. In the end, 18 experts (52.9%) were in the stress condition while 16 experts (47.1%) were in the control condition.

6.3 Results

6.3.1 Stress Manipulation

In the first experiment, the mean stress levels, as measured by state anxiety scale, were higher for the stress group compared with the control group, Welch's $t(96.34) = -6.84$, $p < .000$, with a mean of 51.15 ($SD = 13.10$) compared with a mean of 36.33 ($SD = 9.63$), respectively. The Welch t-test was used when the assumption of homogeneity of variances was not met, as assessed by Levene's test for equality of variances.

Neither age nor sex moderated state anxiety levels. Specifically, there was no correlation between momentary stress levels and age in either the control group ($r(61) = -0.08$, $p = .547$) or the stress group ($r(54) = 0.004$, $p = .976$). Moreover, there was no main effect of sex, $t(112) = -0.18$, $p = .857$, with a mean of 43.53 ($SD = 13.63$) for females and a mean of 43.06, ($SD = 13.71$) for males.

In the second experiment, the mean state anxiety score was higher for the stress group ($M = 40.22$, $SD = 10.77$) compared with the control group ($M = 36.94$, $SD = 12.07$). However, this was not statistically significant, $t(32) = -0.84$, $p = .408$. State anxiety levels were not moderated by age ($r(34) = -0.26$, $p = .145$), years of experience ($r(34) = -0.30$, $p = .090$) or sex ($t(31) = -1.48$, $p = .148$).

6.3.2 Decisions for Same-Source Evidence

The findings suggest that stress improved fingerprint expert assessments for same-source specimens (see Figure 6-1). Specifically, stress resulted in an observable increase in identification decisions (47% *vs.* 55%) and a decrease in exclusion decisions (20% *vs.* 12%) made by the experts—both changes could be categorised as improvement in performance. It appears that the difficulty of the fingerprint evidence moderated these findings, since increased identifications and decreased exclusions were most noticeable in the easy pairs.

Overall, stress did not influence expert risk-taking for same-source evidence (i.e., inconclusive decisions). However, stressed experts were evidently more risk averse when the fingerprint pairs were difficult (54% *vs.* 71%). Interestingly, stress resulted in similar changes to decisions for novices and experts, but only for the overall changes (see top chart of Figure 6-1).

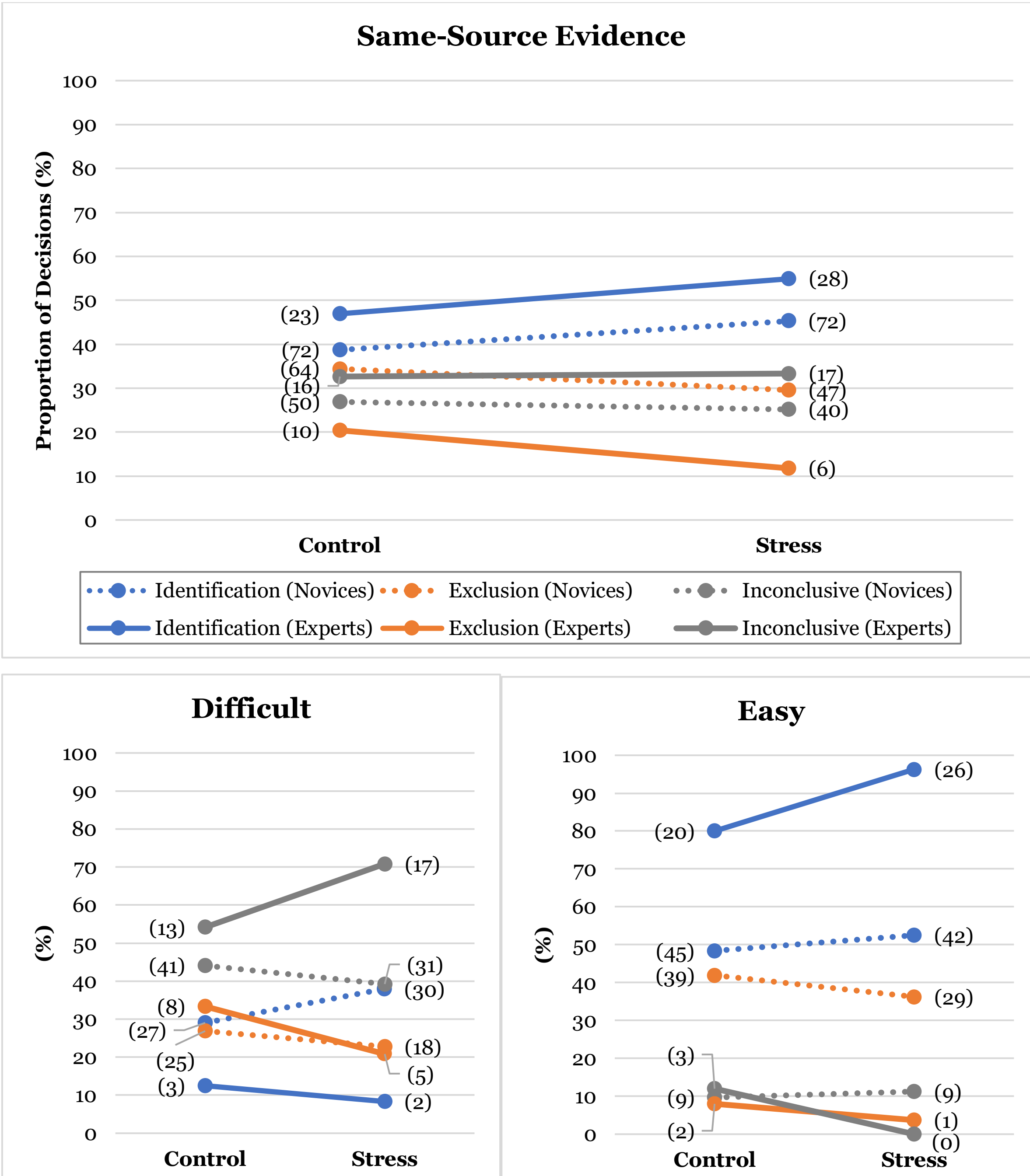

Figure 6-1: Proportions of decisions on same-source evidence for all fingerprint pairs (top), difficult pairs (bottom left) and easy pairs (bottom right). Number of decisions is shown in brackets; lines represent directionality of change between control and stress conditions.

6.3.3 Decisions for Different-Source Evidence

Overall, stress did not result in noticeable decision-making changes in either the expert or novice cohorts (see Figure 6-2). However, the difficulty of the fingerprint assessments played a role in this negligible effect of stress (consistent with same-source findings). For instance, for easy pairs, stress did not influence expert decisions or risk-taking at all—a possible sign of ceiling effect. However, for difficult pairs, stress resulted in minor changes that can be categorised as improved performance (i.e., increase of exclusions: 50% *vs*. 57% and decrease in identifications: 4% *vs*. 0%). No clear pattern was noticed for the decisions reached by novices.

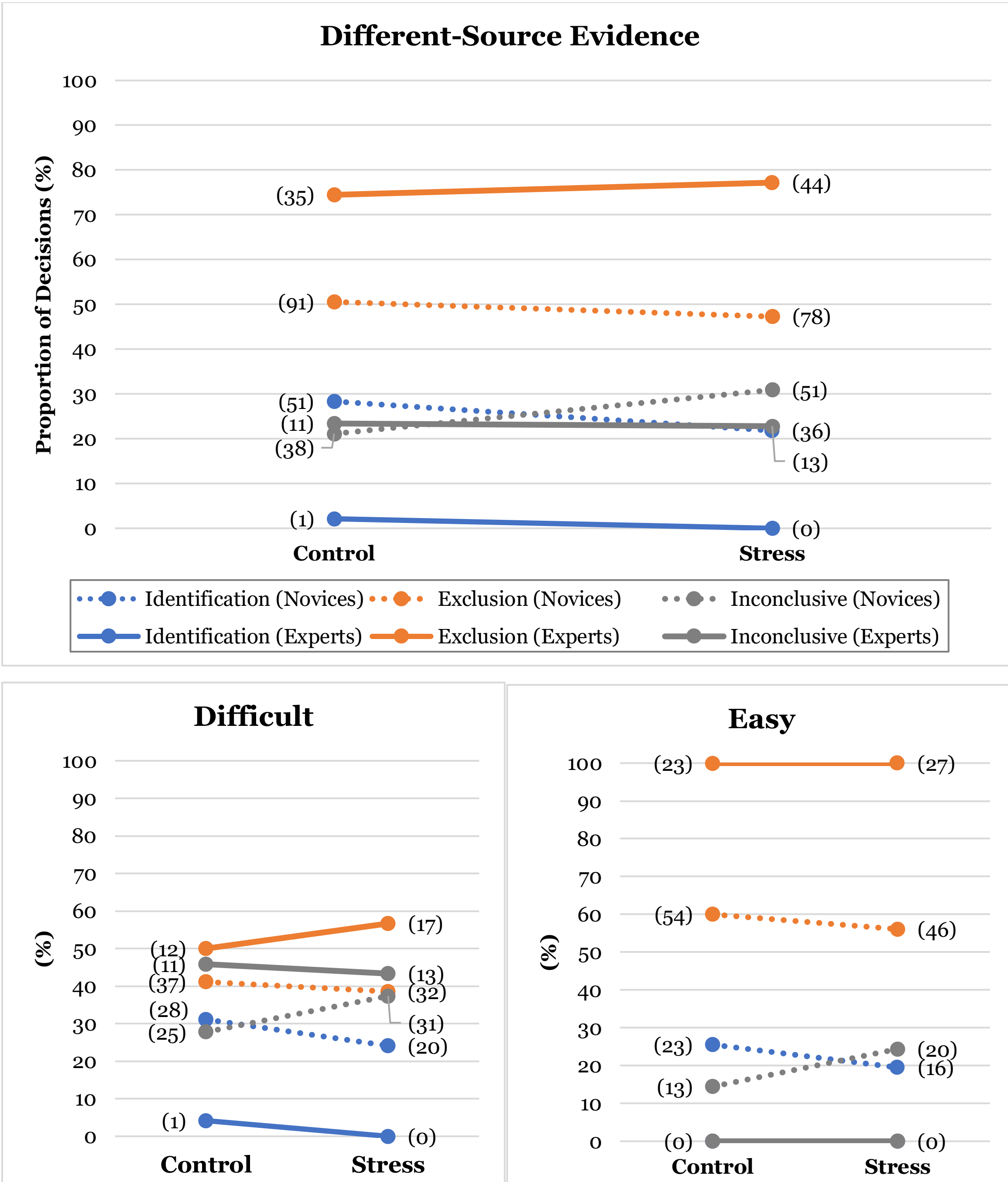

Figure 6-2: Proportions of decisions on different-source evidence for all fingerprint pairs (top), difficult pairs (bottom left) and easy pairs (bottom right). Number of decisions is shown in brackets; lines represent directionality of change between control and stress conditions.

6.3.4 Confidence Levels and Response Times

On average, nonexperts in the first experiment had moderate confidence in making their decisions (*M* = 59.60; *SD* = 23.56). In comparison, fingerprint experts had high confidence in making their decisions in the second experiment (*M* = 89.35; *SD* = 15.94). Table 6-1 summarises the findings on the impact of stress on confidence levels as well as response times. An additional targeted significance test was made on inconclusive decisions made by experts on difficult same-source evidence, since it was desired to understand the observable change in these decisions further (see bottom left chart of Figure 6-1).

The response time for each decision was recorded in seconds. In the first experiment, an outlier was identified and excluded, whose score was more than 30 IQRs above Q3 (i.e., the 75^{th} percentile). Similarly, one outlier was also identified and excluded in the second experiment. With a single score excluded, novices spent an average of 26.67 seconds (*SD* = 26.27; *Med* = 18.30) on each judgment and experts spent considerably longer with an average of 128.59 seconds (*SD* = 177.40; *Med* = 68.66) on each decision. Nevertheless, the response times remained skewed as assessed via the histograms and Q-Q plots. Hence, Mann-Whitney U was used to compare response times across the stress and no-stress conditions (see Table 6-1).

Table 6-1: The impact of stress on confidence Levels (CL) and response times (RT).

	Control	**Stress**	**Significance Testing**	
Novices				
CL (%, mean (*SD*))	61.34 (22.26)	57.70 (24.84)	$t(653.35) = 2.02$	$p = .044$
RT (sec, mean rank)	362.26	325.44	$U = 52790.50$, $z = -2.42$	$p = .015$
Experts				
CL (%, mean (*SD*))				
Overall	87.06 (18.37)	91.38 (13.19)	$t(170.31) = -1.91$	$p = .058$
Inconclusives for difficult same-source evidence	72.77 (23.76)	89.71 (13.84)	$t(18.12) = -2.29$	$p = .034$
RT (sec, mean rank)	110.02	94.94	$U = 4368.00$, $z = -1.83$	$p = .068$

6.4 General Discussion

The findings indicate that stress, as it was induced in this experiment, can improve fingerprint decision-making for both novices and experts, but mainly for same-source evidence. These findings are consistent with the published literature on the impact of moderate stressors on performance. Specifically, when stress is moderate, it can improve human performance (Epel et al., 2018; Yerkes & Dodson, 1908), including for experts in professional domains (e.g., in policing (Akinola & Mendes, 2012)). This could be due to alertness and improved attention (Kowalski-Trakofler et al., 2003; Paton & Flin, 1999)—a cognitive function that is mediated by working memory (Deligkaris et al., 2014). The induced stress level in the current study is considered 'moderate' because the mean state anxiety scores in the stress group (i.e., 51.15 for the experiment with novices, and 40.22 for the experiment with experts) were mid-way between the minimum score of 20 and maximum score of 80.

A key implication of this finding is that it may be useful to induce or maintain moderate levels of stress on forensic experts in general, and specifically on fingerprinting. It is acknowledged that forensic experts already operate in stressful situations, including potentially uncontrollable social evaluative threats. It may also be beneficial for the task to be challenging, thus enhancing the sense of stress during the performance of the task. Previously published research suggests that underload, boredom and repetitive tasks can impair performance of individuals (Driskell et al., 2014). Supervisors could play an important role in optimising expert performance through effectively communicating with the experts on their task loads and other needs .

However, it must be emphasised that the stressor used in this study did not induce a sharp improvement in decision-making, especially in different source-evidence contexts where changes were negligible. Several explanations could account for this observation. It could be that stress, does not impact decisions in the different-source evidence in the same way as same-source evidence. Another explanation could be that the nine experts who dropped out resulted in a different pattern of results than might otherwise have been if they had completed the study. This latter point could also explain the nonsignificant differences in stress manipulation for experts, as those who were possibly most impacted by stress simply dropped out of the study. Yet, since the dropped out experts were all from the stress group, this by itself may reflect the effectiveness of the stress manipulation.

Stress did not have a noticeable effect on fingerprint expert risk-taking, since the proportions of inconclusive decisions was comparable in the stress and control condition for both the same-source and different-source evidence. However, when the same-source prints were challenging, stressed experts were more conservative than non-stressed experts. Specifically, stressed experts reported more inconclusive decisions than non-stressed expert participants (a 17% difference), and with higher confidence levels. It is also worth noting that most of these difficult decisions were reported as inconclusives for both the stress and control conditions (more than 50%). This makes interpreting the impact of stress on expert performance challenging, especially given that inconclusives are already complex to interpret (Dror & Langenburg, 2019). On the one hand, reaching an inconclusive decision can be justifiable, given the difficulty level of the fingerprint pairs and that experts may be motivated to avoid erroneous identifications. On the other hand, it has been contended that reporting a large rate of inconclusives can result in a practical trade-off in potentially having fewer crimes resolved (e.g., see a discussion on “false inconclusives” (Dror, 2020b)).

As expected, fingerprint experts performed better than novices under stress and under no-stress (see Figure 6-1 and Figure 6-2). Experts took more time in making their judgments—on average, they spent approximately five times longer time in making their judgments compared with novices. Moreover, the stressor in this study had a noticeable impact on the fingerprint decision-making process of novices, but not as much on experts. Specifically, stressed novices made their fingerprint decisions faster and with lower confidence levels than the control group of novices (see Table 6-1).

Nevertheless, it is interesting that novices performed reasonably well despite the minimal training on fingerprint assessments they received in this experiment (e.g., identification decisions were reported in about 40% of decisions for same-source evidence, and the trend of improved performance for novices was similar to experts for same-source evidence; see Figure 6-1). Indeed, previous research found that even minimal training in fingerprinting received by naïve participants was effective (Stevenage & Pitfield, 2016). Specifically, this study reported that trained naïve participants performed significantly better on fingerprint assessments than untrained students, but their performance remained substantially lower than fingerprint experts (Stevenage & Pitfield, 2016).

Whilst it appears that novices outperformed experts in making more identifications for difficult same-source evidence, this could be due to experts taking less risk when reaching conclusive judgments. Experts consistently reported more inconclusive decisions than novices for both the difficult same-source and different-source fingerprint pairs, regardless of the stress condition. Indeed, previous empirical research has found that fingerprint experts were more risk averse than members of the general public (Mannering et al., 2021).

In the expert cohort, only one erroneous identification was made. However, the experts made a total of 16 erroneous exclusions. When examined closely, most (N = 13) of these errors arose from same-source fingerprint pairs that were determined to be difficult matching pairs in the pilot of fingerprint stimuli (see Section 6.2.1). Hence, on these occasions, it appears that the difficulty of the matching process played a more important role than the induced stress. Koehler and Liu (2021) suggested that experts can be prone to high error rates (up to 28.1%) when the fingerprints are difficult to assess. In such difficult assessments, opting for inconclusive decisions is what examiners tend to do –depending how these are scored (Dror & Scurich, 2020). The scoring of the inconclusive decisions is tricky, as there is no criteria to assess when these decisions are correct and when they are erroneous (Dror & Langenburg, 2019).

The limitations of the study design are acknowledged. This study was conducted online and therefore naturally induced less stress than real-life stress at work (see discussion in Chapter 5). In addition, this study did not investigate the individual differences in stress responses. It is important to remember that there are individual differences in responses to work related stress, both because different people may have different stress factors in their life outside of their work, as well as different people responding differently to stress.

6.5 Conclusion

The data produced from this study indicated a complex relationship between stress and forensic expert decision-making. Specifically:

- Stress *improved* the performance of both novices and experts on fingerprint assessments, but mainly for same-source evidence.
- Stress did not have an overall observable effect on the risk-taking of experts, measured through inconclusive decisions. However, when the same-source prints were difficult,

experts under stress exhibited less risk-taking by reporting more inconclusives than the control group.

- Fingerprint experts performed better than novices under stress and under no-stress.
- The stressor utilised in this study had a significant impact on the overall confidence levels and response times of novices, but not as much on experts.

This study demonstrates that stress can improve the performance of individuals in making decisions in a fingerprint comparison task. This study draws attention to the potentially positive impact of stress, and opens up avenues for both research to explore the drivers and mechanisms of this in order to inform practice. With additional insights, it may be that there is value in momentary stress on forensic experts, and adjusting the working environment to create challenge and variability of forensic tasks performed by experts. This study also highlights the importance of considering the risk-taking of experts, measured through inconclusive decisions, when assessing performance in stressful situations. Additional research should investigate the impact of stress on forensic decision-making, including in-person experimental stressors that maybe more reflective of stressors within the workplace of forensic experts.

Chapter 7 Discussion

7.1 Introduction

This thesis had two primary objectives. Studies 1 and 2 (Chapters 3 and 4) explored possible sources of stress that forensic experts experience in the workplace. Then, Studies 3 and 4 (Chapters 5 and 6) investigated though an experimental approach the possible impacts of stress on actual decisions about forensic evidence. This chapter presents a synthesised overview of three key themes that have emerged from these empirical studies, and their implications to research and practice:

Theme 1 Addressing common stress factors is a priority;
Theme 2 The positive impact of stress on decision-making; and
Theme 3 Evaluating stress and its impact is not simple or straightforward.

7.2 Theme 1: Addressing Common Stress Factors is a Priority

Stressors can be classified either as common across occupations, such as workload or lack of advancement, or forensic science-specific, such as being exposed to distressing crime scenes (Jeanguenat & Dror, 2018). Forensic experts seem to perceive much of their stress to come from sources that are common. For example, in Studies 1 and 2, two common sources (i.e., stress from management/supervisors, and stress from backlogs and the need to conduct many cases) stood out from forensic-specific stressors that were explored in this thesis (e.g., working in high-profile cases that may involve media coverage), and from extra-organisational stressors (e.g., personal issues). Previous research also highlighted the importance of common stressors, like case backlogs (Busey et al., 2021; National Institute of Justice, 2019).

This is an important observation because it could potentially drive stress-optimising efforts in a direction that is meaningful in practice (i.e., towards addressing common sources of stress). In addition, given that these sources of stress are common across occupations, there may be lessons from other domains that could be beneficial for achieving optimised stress levels in forensic science. Indeed, a recent thematic review identified that supportive supervisors are one of three key organisational stress optimisation factors in the forensic science workplace (Kelty et al., 2021). It should be clarified that this does not mean that forensic science-specific stressors are not important to address. Not all forensic science-specific stressors and their possible impacts on expert decision-making were explored in this thesis (e.g., cross-examination of experts, and other adversarial legal challenges (Jeanguenat & Dror, 2018)). It should also be noted that the term stress-optimising is used in Chapter 7, rather than stress-mitigation, because stress can also have a positive impact on performance.

7.3 Theme 2: The Positive Impact of Stress on Decision-Making

Stress is often considered as negative—for example its role in impacting well-being and decision-making. However, it has emerged from the findings of this thesis that there can be a *positive* impact of stress. Specifically, the stress method (Study 3) induced moderate stress, which in turn resulted in improving some of the expert fingerprint assessments (Study 4). It might be that moderate stress improved decision-making through increased attention, a cognitive function that is mediated by working memory (Deligkaris et al., 2014). These findings are consistent with studies in other domains which found that moderate stress can enhance expert performance (Akinola & Mendes, 2012; Epel et al., 2018; Kowalski-Trakofler et al., 2003; Yerkes & Dodson, 1908).

This is important so that stakeholders in forensic service providers do not solely think or treat stress as negative. Otherwise, the other side of the coin—the positive impact—could be neglected. This negative connotation about stress is already evidenced in the literature. For example, using the terms stress-mitigation or stress-minimisation strategies might imply that workplace stress needs to be uniformly reduced (e.g., "management can take steps to mitigate this stress" (Busey et al., 2021, p.4); "providing multiple types of interventions to mitigate occupational stress" (Goldstein & Alesbury, 2021, p.4)).

Furthermore, a common understanding is that stress is about how people perceive demands (e.g., tasks at work) as greater than their ability to respond to them (Lazarus & Folkman, 1984). However, there is also the aspect of lower demands and boredom in doing the workplace tasks (Driskell et al., 2014), which could cause issues of vigilance. Hence, the possible positive impact of stressors suggests that there might be occasions when stress at the forensic science workplace should be enhanced, not mitigated, so that expert performance is enhanced.

7.4 Theme 3: Evaluating Stress and its Impact is not Simple or Straightforward

This thesis has highlighted that understanding the role of stress in forensic science is a complex issue. For instance, findings from Studies 1 and 2 indicated that forensic experts who have more experience doing their job, reported higher levels of stress. Yet, this evaluation needs to be considered with caution, since experience of experts could be confounded with age (Patterson, 2003) or job role, or other factors that could also moderate the stress levels. Furthermore, the task of inducing stress to human subjects is challenging. This is because stress research requires to generate feelings of stress effectively, but at the same time, the research should carried in an ethical manner to mitigate possible long-term effects of stress (Ferreira, 2019).

The three themes that emerged from the thesis, including that evaluating stress within the forensic science context is complex, has driven the development of *Stressor-Stress Response in Expert Decision-Making* model. This model, presented in Figure 7-1 (see section 7.5.4), aims to capture this complexity and its different contributing elements that could play a role in expert decision-making under stress.

7.5 Synthesis

7.5.1 Common Stressors in Forensic Science

To understand the role of common stressors (Theme 1*)*, there is value in considering the factors that might motivate forensic science professionals, such as the role of popular media (Cole and Dioso-Villa 2009) and the sense of contributing to society and justice through identifying perpetrators and solving crimes (Charlton et al. 2010). These factors, in combination with the reported low feelings of stress identified (e.g., questions 7 and 8 in Study 2), might suggest that some forensic science-specific sources of stress are positively motivating rather than negatively stressful. Some researchers classify the stressors that may offer personal development and achievement for the individual as *challenge stressors*, and the stressors that may negatively impact one's performance as *hindrance stressors* (Wood & Michaelides, 2016). It seems that common stressors, such as supervisory roles or managing caseloads, act more as hindrance stressors (see Studies 1 and 2). In other words, it might be that the working environment and organisational culture surrounding the forensic experts doing their tasks (Dror, 2020a; Gochhayat et al., 2017), not the forensic tasks themselves, have more negative impacts.

Two common stressors were identified as potential main sources in this thesis: stress from management/supervisors, and stress from backlogs and the need to conduct many cases (Studies 1 and 2). Managers and supervisors are key stakeholders that forensic experts communicate with during forensic casework and with whom they build relationships (Almazrouei et al., 2019). Such relationships could be a source of stress to forensic experts. For example, one in four forensic experts strongly felt that their stress originated from their management/supervisors (Study 1, Part One). Importantly, stress from managers/supervisors may also have an influence on the decision-making of forensic experts, such as implicit pressures to reach certain forensic conclusions (Study 1, Part Two). In cases of implicit/explicit pressures, 'transparency' of such contexts may well be useful for stakeholders that rely upon forensic science (Almazrouei, 2020; see also Principle 7 of the Sydney Declaration (Roux et al., 2022)).

Relationships with managers/supervisors could, paradoxically, be supportive so as to moderate the stress felt by forensic experts (Harper, 2022; Holt et al., 2017). Supportive supervisors was identified as one of three key organisational stress-optimisation strategies in the forensic science workplace (Kelty et al., 2021). Kelty et al. (2021) reported that that supervisors (and managers) could buffer stress by being proactive in necessary actions, such as becoming an integral part of the team, being flexible, not micromanaging, being knowledgeable on specific tasks and building mutual trust. It has also been suggested that managers should develop skills and abilities to be emotionally intelligent where they can manage emotions and stress at an individual level with forensic experts (termed *The Emotion-Regulation Skills-Abilities Model* (Harper, 2022)). In such supportive environments, managers could adapt a human-centred approach rather than a task-centred one. Enhancing emotional intelligence may enable managers to actively listen and build confidence with forensic experts so as to manage their stress and well-being (Harper, 2022), which may result in enhanced performance at the forensic science tasks.

The second potential main source of stress, as identified in the thesis, originates from case backlogs. Backlogs and the need to work on many cases is not a new issue in forensic science (Houck, 2020; Roux & Weyermann, 2020), and could be considered a common source of stress (Jeanguenat & Dror, 2018). Therefore, it was suggested that backlogs should be considered a typical issue for most forensic science providers, unless they are extreme (Busey et al., 2021). However, it is acknowledged that how case backlogs are measured and assessed could be complex. For instance, the measure of backlogs (e.g., turnaround time to complete cases) may differ across forensic science providers (Kobus et al., 2011). Adding to this complexity is how experts perceive stress from case backlogs. For example, the experts in Studies 1 and 2 might have interpreted the survey questions about backlogs in a way that they were constantly on work-mode, even outside the workplace. With technologies (e.g., smartphones) there may be an expectation for availability and connectivity beyond the working hours (Beer & Mulder, 2020), thus enhancing online vigilance (Johannes et al., 2021), and potentially affecting employees' stress and well-being—an area that needs further investigation in future forensic science research.

Busey et al. (2021) indicated that direct measures and quotas of case backlogs at the workplace can be counterproductive. It was suggested that experts may change their behaviour and cut corners to meet the quotas, thus potentially affecting the quality of forensic science judgments. In addition, adding extra overtime may cause fatigue that can influence the performance of experts (Busey et al., 2021). Hence, novel approaches that view backlogs as normal (unless extreme) and not as static metrics, may be necessary (Busey et al., 2021). One approach suggests that forensic science providers could address supply-demand imbalances by being dynamic (Kobus et al., 2011). This may mean continual process improvement, by including stakeholders (such as legal and police stakeholders), for effective triaging and reduced backlogs (Houck, 2020). Another approach suggests having a two-way dialog between supervisors/managers and forensic experts on aspects, such as what motivates them to work and how to best optimise laboratory demands with resources to address backlogs (Busey et al., 2021), that is working towards stress-optimisation (Theme 2).

7.5.2 Re-thinking of Stress as a Human Factor in Forensic Science

Supervisors, managers and experts in the forensic science workplace need to first recognise that stress is an important human factor issue, because it could negatively or positively (e.g., Study 4) affect their own well-being and performance (Jeanguenat & Dror, 2018). It might be useful to rethink how stress is perceived in forensic science organisations—from an issue that may be considered as part of the job (e.g., Webb et al., 2002), or when mental health crises like Covid-19 occur (De Kock et al., 2021; Fournier et al., 2022; Puzzo et al., 2022), to a human factor matter on its own that needs continuous evaluation. For instance, private corporations, such as Google, were proactively creating workplace environments that can address the negative impacts of stress on their experts, even before the Covid-19 pandemic (Schaufenbuel, 2015). Some researchers even argue that practitioners and researchers in forensic science should learn from lessons learnt in the pandemic to prepare for challenges (Roux & Weyermann, 2020), such as creating workplace environments to address the negative of impact of stress in forensic science.

The working cultures and stress factors vary across forensic science organisations, and even across forensic fields and individuals within the same organisation. Hence, it is not possible to recommend a one-size-fits-all solution. Therefore, it might be useful for the forensic service providers to take steps to measure stressors and their impact within their own context. This might also enhance the effectiveness of stress-optimisation strategies that the forensic service providers would implement.

Forensic service providers should consider the possible negative or positive impacts of stress on forensic expert decision-making. For instance, a key finding of Study 4 is that moderate stress enhanced forensic expert performance in some forensic tasks, which is consistent with the established knowledge (Benson & Casey, 2013; Yerkes & Dodson, 1908). Hence, it might be beneficial for forensic science providers to take steps to 'nudge' experts (Thaler & Sunstein, 2009) in ways to increase or sustain optimal stress levels if/when appropriate. Since a stress response is an individual issue (Epel et al., 2018; Lazarus & Folkman, 1984), it might be valuable to first identify experts who are underloaded or have been doing the same forensic tasks for a long time, which may cause boredom, and then introduce challenging tasks (for example, by rotating roles).

In highly stressful situations, targeted training could be an approach that may enhance responses to these contexts. For instance, e-training on crime scene scenarios using technologies, such as virtual reality, may prepare experts for handling stressful elements in real scenes (Dror, 2007). In addition, introducing training on mindfulness techniques at the workplace can provide necessary skills required to effectively respond to challenging scenarios, rather than reacting in a fight-or-flight, auto pilot response (Jeanguenat & Dror, 2018). Ideally, any approach that aims to optimise stress, or enhance performance under stress, needs to be empirically informed for its effectiveness in the workplace. Moreover, empirical testing of stress-optimising strategies could be particularly useful, given the complexity of stressor-stress response relationship (Theme 3).

7.5.3 Complexity of Stress Evaluation in Forensic Science

The most challenging part of this thesis was designing a stressor that could be utilised to induce stress feelings in human subjects online in an ethical manner (Study 3; see also Almazrouei et al.

(2022)). The online stress method that was developed for this thesis is based on ecological stressors of social-evaluative threats and uncontrollably (Dickerson & Kemeny, 2004). It will hopefully make a significant contribution to future work in this field, and beyond, by creating a means of being able to carry out studies of this nature in an online setting, which can increase the diversity of participants able to take part (Upadhyay & Lipkovich, 2020), and enable research to continue when in-person research may not be possible (Wigginton et al., 2020). However, it is acknowledged that an online stress method without the presence of stressing agents (e.g., the researchers) is naturally less effective that an in-person stress paradigm or real stressors in the workplace, so these latter types of stressors could be potential avenues for future research.

On a practical level, collecting data to understand forensic expert stress, or to test stress-optimising strategies might be challenging, because experts are typically busy with forensic casework. A potential solution would be to collect data from non-experts when/if applicable. Evidence from Studies 3 and 4 highlighted that novices can offer valuable insights on human behaviour under stress, even in an online environment where researchers were not present. For instance, with minimal training, novices performed reasonably well and followed a pattern consistent with experts when assessing same-source fingerprint pairs (Study 4). However, whilst data from novices could be valuable, insights from their data are limited and may not necessarily fully translate to expert decision-making. This is partially because experts are more reliant on shortcuts or schemas from their accumulated experience to enable processing information more efficiently (Gigerenzer & Gaissmaier, 2011).

7.5.4 Stressor-Stress Response Model in Expert Decision-Making

Evaluating stress and its impact is highly complex (Theme 3). A Stressor-Stress Response Model in Expert Decision-Making is presented in Figure 7-1, which aims to offer a holistic overview of stressors and stress responses in professional disciplines. Here, the model is illustrated with examples within the forensic science discipline, yet its insights could be applicable to any professional disciplines where experts may operate under stress (e.g., policing, medical and legal professionals).

Considering the forensic science context, the model highlights the cognitive factors affecting forensic expert decision-making under stress at the crime scene, in the laboratory and/or in the court. The model is derived from theories and concepts in the published literature, such as cognitive factors in decision-making (Dror, 2007), models of stress (Cooper & Marshall, 1976; Epel et al., 2018), and the integrated forensic science process (Morgan, 2017a). In addition, it is supported by empirical evidence from the published literature and the current thesis. The introduced model (Figure 7-1) may help in highlighting current gaps in knowledge in stress research within the forensic science contexts, and beyond. It is also hoped that this model will assist managers with the identification of possible sources of stress and the impact of stressors to contribute to the design of stress-optimising strategies in the workplace.

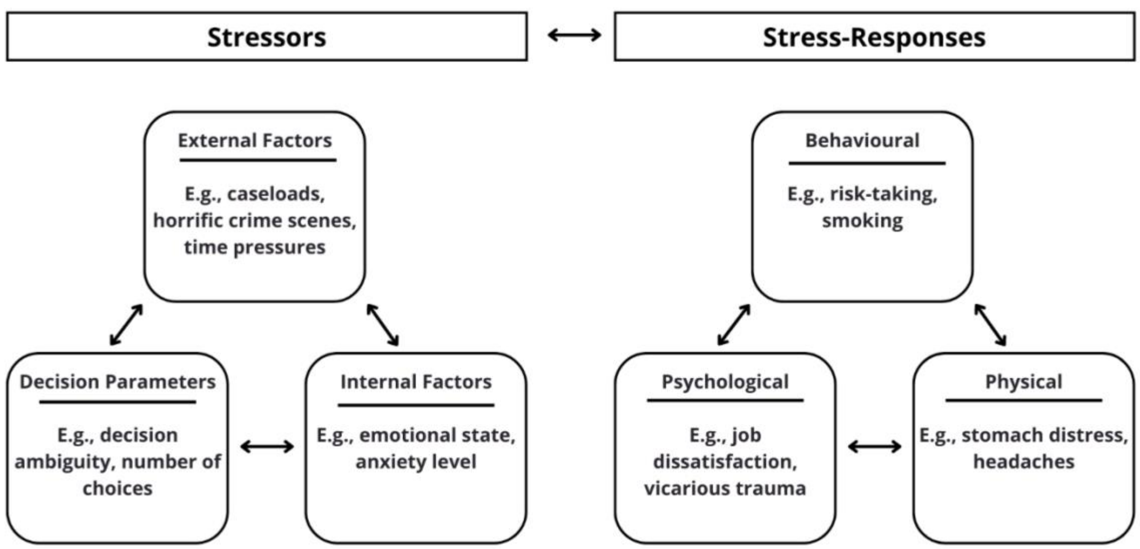

Figure 7-1: Stressor-Stress Response Model in Expert Decision-Making. Here, forensic experts are at the core of the model. Stressors could be pertinent to the decision parameters that the experts are making, internal to the individual experts themselves, and/or external factors. Responses to these stressors could be in the form of behavioural, psychological, and/or physical responses. These stressors and stress responses are interconnected with each other, and may cognitively affect the whole forensic science process (crime scene to court). Two-way arrows show the connectivity between the elements of the model.

7.5.4.1 Stressors

At the core of this model is the decision maker—the forensic experts. It is the experts who conduct forensic casework and analysis. Three cognitive factors could be involved in expert decisions under stress: decision parameters, internal and external factors (Dror, 2007). Decision parameters pertain to how complex the decision is. Examples may include the complexity of the decision and the number different alternative choices (e.g., type and number of available minutiae to consider in fingerprint assessments). Forensic experts—as is the case for all experts—have limited cognitive resources to examine different alternatives (Kahneman, 2003), and make either a definitive or a non-definitive decision. This is an important factor that needs consideration when evaluating the impact of stress.

To illustrate the aforementioned point, there has been recent criticism that non-definitive decisions (such as inconclusive conclusions in the fingerprint field (Dror & Langenburg, 2019), or an undetermined manner of death in forensic pathology field (Dror et al., 2021)) may be a preferred option to take (see a discussion on inconclusive decisions in Part Two of Study 1 and also in Study 4). Opting for non-definitive decisions, like inconclusives, could be a way to avoid risky definitive decisions that maybe challenged in court (Dror & Langenburg, 2019). Added to this complexity is the interpretation of inconclusive judgments in forensic science contexts (Study 4). For instance, an increase in inconclusive decisions could be the result of decision fatigue (i.e., inability to make decisions as a result of casework pressures and stress; see Busey et al., (2015)), or an increase in inconclusive decisions could be the result of experts being more conservative when stressed.

Individual characteristics of the decision maker play an important role in stress responses (See Model of Stress at Work (Cooper & Marshall, 1976)). These relate to internal factors that can impact forensic experts, not the decision itself or external circumstances (Dror, 2007). Examples may include the expert's level of anxiety, aversion to ambiguity, emotional state, or confidence levels (Dror, 2007; Raptis et al., 2017; Saposnik et al., 2017). For instance, it was found in Study 3 of this thesis that non-expert individuals with high trait anxiety were not influenced by the online stressor, in contrast to individuals with low and moderate trait anxiety levels. Despite the insight that decision parameters and internal factors are distinct, they could also be interrelated and affect one another (Dror, 2007). For instance, forensic examiners who are highly averse to ambiguity (i.e., those who have extreme dislike of events of unknown probabilities (Levy et al., 2010)) may be influenced when faced with complex and time-pressured decisions (Helsloot & Groenendaal, 2011), such as collecting traces from ambiguous crime scenes (de Gruijter et al., 2016; van den Eeden et al., 2019).

Making decisions in stressful situations could also involve external factors, such as time pressure that police officers may face when deciding to shoot a suspect (Dror, 2007), or that forensic experts may also face (Helsloot & Groenendaal, 2011; Zou et al., 2021)—but arguably to a lesser degree. External factors are not about the decision or the decision-maker, but they are pertinent to the circumstances surrounding forensic expert decision-making. Such factors may include forensic-specific stressors (e.g., being exposed to horrific case details), or stressors that are not specific to forensic science (e.g., lack of advancement at work; see Theme 1). Notably, it is contended that external factors could be connected with both decision and internal factors (Dror, 2007). For instance, time-pressure was found to raise cognitive demands on individuals (Dror et al., 1999). In turn, individuals could make a choice among a complex decision of different choices given the available cognitive resources (Dror, 2007), such as triaging which items collected from the crime scene to send for forensic analysis (US Bureau of Justice Assistance, 2019).

7.5.4.2 Responses to Stress

As demonstrated in Figure 7-1, responses to stress could be broadly categorised as: behavioural, psychological, or physical (Kelty & Gordon, 2015; Spector, 2012). When faced with stressors, forensic experts may behave differently, affecting their decision-making (Jeanguenat & Dror, 2018). For instance, fingerprint examiners were more likely to find a 'match' when they in a highly emotional context (Dror et al., 2005). Importantly, stress response is an individual-level issue (Lazarus & Folkman, 1984). Hence, the same stressor could be perceived differently (Epel et al., 2018; Lazarus & Folkman, 1984), which can result in different expert decision-making processing or conclusions. In addition, responses to stress could be either positive (i.e., result in improvement of expert decision-making performance; see *Theme 2*), or negative on experts' decision-making, or even their coping mechanisms with stress. Some may use positive or adaptive coping mechanisms to manage stress (e.g., use of humour in the workplace (Kelty & Gordon, 2015)). Alternatively, forensic examiners may go for negative or maladaptive coping mechanisms (e.g., smoking (Holt & Blevins, 2011)).

The published literature reported different forms of psychological responses that forensic experts may experience. For example, these responses could be momentary (e.g., anger), or over a period of time (e.g., job satisfaction); they could also be direct responses (e.g., post-traumatic stress disorder (Yoo et al., 2013)), or indirect responses (e.g., secondary traumatic stress from exposure to contents that involve crimes against children (Burruss et al., 2018; Busey et al., 2021)). The final category of stress responses pertains to physiological reactions of participants (e.g., headaches, dizziness, stomach distress (Spector, 2012)). These physical reactions could result from forensic-specific or common stressors, including the working environment of experts. For instance, Kelty et al. (2021) discusses that the working environment, such as working space with excessive heats or furniture inappropriate for long hours of sedentary tasks, is a key stressor to digital examiners (Kelty et al., 2021).

Stressor responses could be interconnected and influence one another. For instance, working in a high-profile case for a prolonged period of time could result in fatigue (physical), which can affect expert decision-making (behavioural) (for example, see the discussion on the effect of fatigue on working memory of fingerprint experts (Busey et al., 2015)). Another example that links psychological and physical response is where secondary traumatic stress may include physical signs, such as headaches and muscle tension (e.g., see Table 1 in Kelty et al. (2021)).

Forensic experts make decisions and are exposed to stressors throughout the forensic science process: crime scene, laboratory, and court (Morgan, 2017a, Morgan et al., 2018; see also Section 7.5.4.1). Morgan et al., (2018) clarify that forensic expert decision-making is a connected process. That is, decisions influenced by context, risk-taking or other human factors happening early at the crime scene may well cascade to the laboratory, which then may affect the legal outcome of the forensic case (Dror et al., 2017; Nakhaeizadeh et al., 2017). Adapting this understanding, the model highlights that those stressors affecting crime scene decisions may well extend beyond the crime scene, up until the stage where decision-making is taking place in a court setting.

It is hoped that this model can help in identifying gaps to be addressed in future research. For instance, this thesis did not address the impact of stress at the crime scene or in the court stages of the forensic science process. The model may also offer a systems-level view to forensic service providers to develop stress-optimising strategies that are specific to the stressors or stress responses outlined (see also the discussion on forensic science as a ‘system’ within a broader context and interactions in the criminal justice system (Houck, 2020)).

However, this model is limited and could be improved in future research. For instance, it might be useful to expand on the stress model by including the specific relationships between forensic expert and other stakeholders in forensic science. Specifically, forensic experts could communicate with multiple stakeholders about their work (e.g., judges, lawyers, ISO auditors, regulators, family, victims, etc (Almazrouei et al., 2019; Dror & Pierce, 2020)). These interactions could generate different types of stressors to experts (e.g., adversarial allegiance with lawyers (Murrie et al., 2009)) that may require context specific solutions. Approaches, such as System Thinking (Houck, 2020; Midgley & Lindhult, 2021) or Social Network Analysis (Campana, 2016) might be helpful in better understanding such interactions. This is because these approaches suggest assessing the individuals (e.g., the forensic expert) in connection within the broader context, system or network that they operate at, rather than in isolation.

Chapter 8 Conclusion

The main aim of this PhD research was to gain insights into the possible sources of stress and their impact on decision-making in forensic science. Deeper understanding of stress in the forensic science workplace has the potential to contribute to best practices and to enhance the quality of forensic expert decision-making, particularly under conditions of stress. To address this aim, the thesis addressed two main research questions through a holistic approach that brought together self-reporting and experimental data approaches:

- Research Question 1: What degree of workplace stress (and feedback) do forensic experts experience? (Self-reported data)
- Research Question 2: Does stress impact forensic science decision-making? Specifically, does stress influence fingerprint assessments made by novices and fingerprint experts? (Experimental data)

In this chapter, the key findings of this thesis that have addressed the possible sources and impact of stress on forensic experts' well-being and decision-making are presented. Ten practical implications of these findings are then outlined.

8.1 Key Findings

8.1.1 Research Question (1): What degree of workplace stress (and feedback) do forensic experts experience?

In Study 1 (Chapter 3), 150 practicing forensic experts from the same laboratory were surveyed about their experiences of workplace stress, and the explicit and implicit feedback they receive. Forensic examiners reported that their high stress levels originated more from workplace related factors (management and/ or supervision, backlogs and the pressure to do many cases) than from personal related factors (family, medical and/ or financial). The findings showed that a small proportion (8%) of the forensic examiners sometimes felt strong implicit feedback about what conclusions were expected from them, and that some (14%) also felt strongly that they were more appreciated when they helped to solve a case (e.g., by reaching a 'match' as opposed to an 'inconclusive' conclusion). Differences were found when comparing workplace stress and feedback levels across three core forensic science fields (forensic biology, chemistry and latent prints), and across career stages (early, mid, and late).

Study 2 followed-up the insights gained from the work presented in Chapter 3. In this study, a new sample of 41 forensic experts from two laboratories were surveyed about the sources of their stress, the support they receive, and the potential influence of stress on their decisions (see Chapter 4). Stress from managers, supervisors and case backlogs were identified as significant factors that contributed to stress in the workplace. Neither the type of case nor working in high-profile cases were reported to be major sources of stress. Crime scene examiners reported feeling higher levels of stress from personal reasons and from the nature of their cases than analytical examiners. Female examiners reported feeling more stressed than male examiners from both general stressors and workplace stressors. Examiners in the 11–15 years of experience group felt more pressure as a result of circumstances at work than examiners in the 7–10 years group. The level of management support was not associated with either the feelings of general stress or stress in the workplace. Examiners varied in their perceptions of whether stress affected their judgements: 39% felt that their judgments were influenced by stress, while 22% did not and 39% were unsure.

Therefore, Chapters 3 and 4 contribute data to address research question 1. Taking the findings together, many of the sampled forensic examiners felt that they operate under pressure, and that the level of pressure varies by field, sex and experience. These stress feelings could extend to pressure from feedback during casework and reporting conclusions.

8.1.2 Research Question (2): Does stress impact forensic science decision-making? Specifically, does stress influence fingerprint assessments made by novices and fingerprint experts?

To address this question, a new method was developed to induce stress in human subjects during online participation in a research study, without the presence of researchers (see Chapter 5). In working towards the validation of this method, participants in the stress inducing condition (N = 52, 44%) were asked to answer general knowledge and mathematical questions which people often get wrong, and did so under time pressure as well as receiving feedback. In contrast, participants in the control condition (N = 66, 56%) did not have time pressure or receive feedback. The stress manipulation was found to be effective, as the reported state anxiety and visual analog scale on stress scores were higher for the stress group than for the non-stress group (both findings, p <.001). Consistent findings were found when accounting for trait anxiety as a moderator, with the exception of the state anxiety levels in high trait anxiety group. This stressing method combines the established stress conditions of uncontrollability (such as time pressures) and social evaluative threats (such as negative feedback). In addition, the method contains specific measures (such as a commitment statement and attention check questions) to enhance the internal validity by preventing and detecting cheating or random responses. This method can be deployed through any commonly available online software. It offers a simple and cost-effective way to collect data online –which fits the increasing need to carry out research in virtual and online environments (Kirschbaum, 2021; Upadhyay & Lipkovich, 2020).

Following the development of the stress-inducing method presented in Chapter 5, an experimental study was conducted to examine the impact of stress on forensic science decision-making contexts, where experts can face various levels of stress. This study examines fingerprint decisions made under stress, by novices (N = 115) and fingerprint experts (N = 34). Findings suggested a complex relationship between stress and expert performance. On the one hand, stress improved the performance of both novices and experts on fingerprint assessments, but mainly for same-source evidence. On the other hand, stress had an impact on risk-taking. When the same-source prints were difficult, stressed experts were less risk-taking and reported more inconclusive conclusions with higher confidence than the control group. Furthermore, stress had a significant impact on the overall confidence levels and response times of novices, but not experts. Stress and decision-making tasks are important factors that should be considered when creating optimal working environments for increasing decision quality.

Therefore, Chapters 5 and 6 contribute data to address research question 2. The stressor used in this thesis had a positive impact on fingerprint assessments by both novices and experts. However, this relationship is a complex one, given the secondary impact of stress on risk-taking.

8.2 Practical Implications

The findings from this thesis have the potential to inform and contribute to policies to enhance the well-being and decision-making of forensic experts. Practical implications may include:

1. Going forward, it is important to recognise that stress is an important human factor that may negatively or positively impact the well-being and decision-making of professionals in the workplace. Evidence from Study 2 demonstrated that experts were divided and did not have the same perception on the potential impact of stress on their judgments. Such a recognition by the experts and their supervisors/managers could be a first step to putting forward stress-optimising strategies. It has been suggested that a focus on managing the stress and mental health of employees has increased in the last two years due to the Covid-19 pandemic (De Kock et al., 2021; Fournier et al., 2022; Puzzo et al., 2022). However, it might be useful to re-consider stress as an integral organisational factor in expert decision-making (Dror, 2020a) that needs continuous assessments, not just when crises happen. Lessons from private organisations could be useful here. For instance, Google, Target and other private corporations have been proactive in creating workplace environments that can address the negative impacts of stress on their experts, even before the Covid-19 pandemic (Schaufenbuel, 2015).

2. Designing stress-optimising strategies to specifically address stressors that are common across occupations, such as stress from supervisors/managers in forensic science (Jeanguenat & Dror, 2018). For instance, the findings in Studies 1 and 2 suggest that supervisors/ managers play an important role in why forensic experts feel stressed at work. It has also been identified in the literature that supportive supervisors could moderate stress (Kelty et al., 2021). Hence, a possible strategy would be to have targeted training for supervisors/ managers on emotional intelligence (Harper, 2022; Lidén, 2020). Emotional intelligence is described as a competence that may enable a person to alleviate stress of another through acquired skills, such as active listening and understanding expectations of another (Harper, 2022). In a working environment that has been characterised as time-pressured (Helsloot & Groenendaal, 2011) and with case backlogs (National Institute of Justice, 2019), it might be beneficial for supervisors/ managers to acquire skills needed to 'pause', when necessary, to manage the emotions and needs of the forensic experts.

3. Approaches to increase the transparency of how expert decisions are reached (Almazrouei et al., 2019; Almazrouei, 2020; Earwaker et al., 2020) may need to be integrated with standard operating procedures of forensic service providers (Dror & Pierce, 2020; Heavey et al., 2022). Evidence from Study 1 suggest that experts may feel implicit pressures that come from feedback with stakeholders (such as police investigators), to reach certain conclusions (e.g., by reaching a 'match' as opposed to an inconclusive conclusion). If such pressures occurred in a case, they may play a role in the decision-making of experts. Hence, in these situations, a transparency approach, such as the forensic disclosure model, may be useful in demonstrating the context of experts' decisions (Almazrouei et al., 2019; Almazrouei, 2020). The model suggests that it might be beneficial to document these communications in the casework files, and be transparent about them by disclosing them to stakeholders (e.g., judges), when/if appropriate (Almazrouei et al., 2019).

4. Stress-optimisation strategies should be context-specific (e.g., considering the demographical background of experts). For instance, Study 2 revealed that experts who operate at the crime scene experienced higher stress from personal factors (such as family and financial issues) than those who operate mainly in the laboratory. Hence, it would be valuable to understand better the context of crime scene work in order to offer effective solutions, such as addressing stress from work shifts and commitments of crime scene experts with their families (Kelty & Gordon, 2015).

5. There may need for a reconsideration of how case backlogs contribute to the stress and performance of experts. The findings from this thesis identified that case backlogs and the need to work on many cases were perceived to be key stressors by forensic experts (Studies 1 and 2), even more than other workplace stress factors, such as working in high profile cases or circumstances at work from investigation/prosecution pressure (Study 2). However, as the pressure to work on many cases could be a common issue (Jeanguenat & Dror, 2018), this factor might need to be considered as a typical challenge unless it is extreme (Busey et al., 2021). Hence, a valuable recommendation to optimise experts' stress originating from case backlogs might be to re-assess how backlogs are measured (e.g., possibly by abandoning direct quotas, as experts may cut corners to meet such quotas (Busey et al., 2021)).

6. Stress does not always have a negative impact (e.g., see Study 4), so forensic service providers may need to consider opportunities to optimise stress levels to enhance expert performance in specific situations or scenarios. This could be achieved in practice through different approaches, since the response to stress is dependent on individuals (Epel et al., 2018; Lazarus & Folkman, 1984), and the stress originating from the forensic science tasks. In the case of low stress, it might be valuable to first identify experts who are underloaded or have been doing the same forensic science tasks for a long time, which may cause boredom (Driskell et al., 2014), and then introduce challenging tasks (for example, by rotating roles). In the case of high stress, targeted training is an approach that may enhance responses in stressful situations. For instance, e-training on crime scene scenarios using technologies, such as virtual reality, may prepare experts for handling stressful elements in real scenes (Dror, 2007).

7. Inconclusive decisions should be considered when evaluating expert performance in casework. If an expert completes many cases quickly by unjustifiably reporting too many inconclusives, then one might argue that their performance is high, if inconclusive decisions are not considered. However, taking low risks through *not* making conclusive decisions could negatively impact casework (e.g., not identifying suspects) (Dror & Langenburg, 2019). Study 4 has taken a step forward towards this direction, by assessing inconclusives decisions. This evaluation revealed insights on the complex impact of stress on fingerprint assessments in that expert may justifiably not take risks when the fingerprint decisions are difficult (Study 4). It is acknowledged that there is still no consensus among researchers on how inconclusive conclusions should be incorporated in assessing expert performance (Dror & Scurich, 2020; Scurich & Dror, 2020; Weller & Morris, 2020), but these disagreements should not hinder efforts towards this direction.

8. Being able to carry out online studies has a number of advantages, such as being able to carry the studies when in-person research may not be possible (Wigginton et al., 2020) and to increase the diversity of participants (Upadhyay & Lipkovich, 2020). Hence, researchers and practitioners in forensic science, and beyond, may benefit from using measures to enhance quality of online studies or interventions (e.g., piloting a stress-optimisation strategy). Specifically, Study 3 of this thesis contained specific measures

(such as adding a commitment statement) to enhance the internal validity of online studies by preventing and detecting cheating or random responses (see Appendix B-1).

9. Collaborative research is encouraged, particularly for addressing difficult research questions and real-life challenges (e.g., see a recent Manifesto on Collaborative Research (Barker et al., In Press)). For instance, the development of the online stress-inducing method was not a simple task (Study 3), and it required insights from multiple disciplines (e.g., political science for enhancing data quality on knowledge-based questions (Clifford & Jerit, 2016)). Furthermore, informal 'collaborative' insights were gained from supervisors, colleagues, friends and family throughout the development of this study. The outcome was a novel stress-inducing method, which may enable advancements in addressing stress challenges within the workplace and life generally (Almazrouei et al., 2022).

10. Researchers as well as practitioners in forensic science fields (and other disciplines) may benefit from the Stressor-Stress Response Model in Expert Decision-Making (Chapter 7). This model offers an overview of stressors and stress-responses within the forensic science context, and the possible relationships that may exist between them. Hence, the model might help in identifying gaps in knowledge on stress research (e.g., stress situations during forensic expert testimony at courts), and might help informing the design of stress-optimising strategies in the forensic science workplace.

This thesis has investigated some of the possible sources and impacts of stress in forensic science contexts. Specifically, research undertaken here differs from other stress research in forensic science in that it offered a discussion and data focused on forensic expert decision-making, as opposed to solely their well-being. In addition, the experimental approach was a step forward to better understand the relationship between stress and expert performance, which is critical because it is the forensic experts themselves who make decisions that are relied upon by other stakeholders (e.g., police for investigations).

References

Adderley, R., Smith, L. L., Bond, J. W., & Smith, M. (2012). Physiological measurement of crime scene investigator stress. *International Journal of Police Science & Management*, *14*(2), 166–176. https://doi.org/10.1350/ijps.2012.14.2.274

Akinola, M., & Mendes, W. B. (2012). Stress-induced cortisol facilitates threat-related decision making among police officers. *Behavioral Neuroscience*, *126*(1), 167–174. https://doi.org/10.1037/a0026657

Allen, A. P., Kennedy, P. J., Dockray, S., Cryan, J. F., Dinan, T. G., & Clarke, G. (2017). The Trier Social Stress Test: Principles and practice. *Neurobiology of Stress*, *6*, 113–126. https://doi.org/10.1016/j.ynstr.2016.11.001

Allendorfer, J. B., Heyse, H., Mendoza, L., Nelson, E. B., Eliassen, J. C., Storrs, J. M., & Szaflarski, J. P. (2014). Physiologic and cortical response to acute psychosocial stress in left temporal lobe epilepsy—A pilot cross-sectional fMRI study. *Epilepsy & Behavior*, *36*, 115–123. https://doi.org/10.1016/j.yebeh.2014.05.003

Allendorfer, J. B., Nenert, R., Hernando, K. A., DeWolfe, J. L., Pati, S., Thomas, A. E., Billeaud, N., Martin, R. C., & Szaflarski, J. P. (2019). FMRI response to acute psychological stress differentiates patients with psychogenic non-epileptic seizures from healthy controls – A biochemical and neuroimaging biomarker study. *NeuroImage: Clinical*, *24*, 101967. https://doi.org/10.1016/j.nicl.2019.101967

Almazrouei, M. A. (2020). Comment on "cognitive and human factors in expert decision making: Six fallacies and the eight sources of bias." *Analytical Chemistry*, *92*(18), 12725–12726. https://doi.org/10.1021/acs.analchem.0c03002

Almazrouei, M. A., Dror, I. E., & Morgan, R. M. (2019). The forensic disclosure model: What should be disclosed to, and by, forensic experts? *International Journal of Law, Crime and Justice*, *59*, 100330. https://doi.org/10.1016/j.ijlcj.2019.05.003

Almazrouei, M. A., Dror, I. E., & Morgan, R. M. (2020). Organizational and human factors affecting forensic decision-making: Workplace stress and feedback. *Journal of Forensic Sciences*, *65*(6), 1968–1977. https://doi.org/10.1111/1556-4029.14542

Almazrouei, M. A., Morgan, R. M., & Dror, I. E. (2021). Stress and support in the workplace: The perspective of forensic examiners. *Forensic Science International: Mind and Law*, 100059. https://doi.org/10.1016/j.fsiml.2021.100059

Almazrouei, M. A., Morgan, R. M., & Dror, I. E. (2022). A method to induce stress in human subjects in online research environments. *Behavior Research Methods*. https://doi.org/10.3758/s13428-022-01915-3

American Society of Crime Laboratory Directors. (2019). *Trauma and stress in the field of forensi science*. https://www.ascld.org/wp-content/uploads/2019/02/Trauma-and-Stress-in-the-Field-of-Forensic-Science.pdf

Anshel, M. H. (2000). A conceptual model and implications for coping with stressful events in police work. *Criminal Justice and Behavior*, *27*(3), 375–400.

Arora, S., Sevdalis, N., Nestel, D., Woloshynowych, M., Darzi, A., & Kneebone, R. (2010). The impact of stress on surgical performance: A systematic review of the literature. *Surgery*, *147*(3), 318–330. https://doi.org/10.1016/j.surg.2009.10.007

Arscott, E., Morgan, R., Meakin, G., & French, J. (2017). Understanding forensic expert evaluative evidence: A study of the perception of verbal expressions of the strength of evidence. *Science & Justice*, *57*(3), 221–227. https://doi.org/10.1016/j.scijus.2017.02.002

Ask, K., & Granhag, P. A. (2005). Motivational sources of confirmation bias in criminal investigations: The need for cognitive closure. *Journal of Investigative Psychology and Offender Profiling*, *2*(1), 43–63. https://doi.org/10.1002/jip.19

Baddeley, A. D., & Logie, R. H. (1999). The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.004

Barker, N., Pervez, A., Wahome, M., McKinlay, A., Al Haj Sleiman, N., Harniess, P., Puskás, N., Mac, D., Almazrouei, M. A., Ezenwajiaku, C., Isiwele, A., Tan, N., D'Aprix, M., Petsou, A., & Love Soper, J. (In Press). A collaborative research manifesto! An early career response to uncertainties. *International Journal of Social Research Methodology*.

Becker, W. S., Dale, W. M., Lambert, A., & Mangus, D. (2005). Letter to the editor—Forensic lab directors' perceptions of staffing issues. *Journal of Forensic Sciences*, *50*(5), 1–3. https://doi.org/10.1520/JFS2005201

Beer, P., & Mulder, R. H. (2020). The effects of technological developments on work and their implications for continuous vocational education and training: A systematic review. *Frontiers in Psychology*, *11*, 918. https://doi.org/10.3389/fpsyg.2020.00918

Bell, A. S., Rajendran, D., & Theiler, S. (2012). Job stress, wellbeing, work-life balance and work-life conflict among Australian academics. *Electronic Journal of Applied Psychology*, *8*(1), 25–37. https://doi.org/10.7790/ejap.v8i1.320

Benson, H., & Casey, A. (Eds.). (2013). *Stress management: Approaches for preventing and reducing stress*. Harvard Medical School.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765. https://doi.org/10.1037/0033-295X.113.4.700

Bonach, K., & Heckert, A. (2012). Predictors of secondary traumatic stress among children's advocacy center forensic interviewers. *Journal of Child Sexual Abuse*, *21*(3), 295–314. https://doi.org/10.1080/10538712.2012.647263

Brady, P. Q., Fansher, A. K., & Zedaker, S. B. (2019). Are parents at a higher risk for secondary traumatic stress?: How interviewing child victims impacts relationships with forensic interviewer's friends and family. *Child Abuse & Neglect*, *88*, 275–287. https://doi.org/10.1016/j.chiabu.2018.11.017

Burke, R. J. (1994). Stressful events, work-family conflict, coping, psychological burnout, and well-being among police officers. *Psychological Reports*, *75*(2), 787–800. https://doi.org/10.2466/pr0.1994.75.2.787

Burruss, G. W., Holt, T. J., & Wall-Parker, A. (2018). The hazards of investigating internet crimes against children: Digital evidence handlers' experiences with vicarious trauma and coping behaviors. *American Journal of Criminal Justice*, *43*(3), 433–447. https://doi.org/10.1007/s12103-017-9417-3

Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences*, *23*(3), 255–282. https://doi.org/10.1016/0165-4896(92)90043-5

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432–459. https://doi.org/10.1037/0033-295X.100.3.432

Busey, & Dror, I. E. (2011). Special abilities and vulnerabilities in forensic expertise. In A. McRoberts (Ed.), *Fingerprint Sourcebook (ch. 15, pp. 1-23)*. NIJ Press.

Busey, Sudkamp, L., Taylor, M. K., & White, A. (2021). Stressors in forensic organizations: Risks and solutions. *Forensic Science International: Synergy*, 100198. https://doi.org/10.1016/j.fsisyn.2021.100198

Busey, Swofford, H. J., Vanderkolk, J., & Emerick, B. (2015). The impact of fatigue on latent print examinations as revealed by behavioral and eye gaze testing. *Forensic Science International*, *251*, 202–208. https://doi.org/10.1016/j.forsciint.2015.03.028

Campana, P. (2016). Explaining criminal networks: Strategies and potential pitfalls. *Methodological Innovations*, *9*, 205979911562274. https://doi.org/10.1177/2059799115622748

Champod, C. (2014). Research focused mainly on bias will paralyse forensic science. *Science & Justice*, *54*(2), 107–109. https://doi.org/10.1016/j.scijus.2014.02.004

Charlton, D., Fraser-Mackenzie, P. A. F., & Dror, I. E. (2010). Emotional experiences and motivating factors associated with fingerprint analysis. *Journal of Forensic Sciences*, *55*(2), 385–393. https://doi.org/10.1111/j.1556-4029.2009.01295.x

Choi, E., Johnson, D. A., Moon, K., & Oah, S. (2018). Effects of positive and negative feedback sequence on work performance and emotional responses. *Journal of Organizational Behavior Management*, *38*(2–3), 97–115. https://doi.org/10.1080/01608061.2017.1423151

Christensen-Szalanski, J. J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *48*(1), 147–168. https://doi.org/10.1016/0749-5978(91)90010-Q

Chung, J., & Monroe, G. S. (2003). Exploring social desirability bias. *Journal of Business Ethics*, *44*(4), 291–302. https://doi.org/10.1023/A:1023648703356

Clifford, S., & Jerit, J. (2016). Cheating on political knowledge questions in online surveys: An assessment of the problem and solutions. *Public Opinion Quarterly*, *80*(4), 858–887. https://doi.org/10.1093/poq/nfw030

Cole, S. A. (2006). Is fingerprint identification valid? Rhetorics of reliability in fingerprint proponents' discourse. *Law & Policy*, *28*(1), 109–135. https://doi.org/10.1111/j.1467-9930.2005.00219.x

Cooper, G. S., & Meterko, V. (2019). Cognitive bias research in forensic science: A systematic review. *Forensic Science International*, *297*, 35–46. https://doi.org/10.1016/j.forsciint.2019.01.016

Cooper, & Marshall, J. (1976). Occupational sources of stress: A review of the literature relating to coronary heart disease and mental ill health. *Journal of Occupational and Organizational Psychology*, *49*(1), 11–28. https://doi.org/10.1111/j.2044-8325.1976.tb00325.x

Corner, E., & Gill, P. (2019). Psychological distress, terrorist involvement and disengagement from terrorism: A sequence analysis approach. *Journal of Quantitative Criminology*, *36*, 499–526. https://doi.org/10.1007/s10940-019-09420-1

Craven, H. P., Hallmark, M., Holland, F., & Maratos, F. A. (2022). Factors influencing successful coping among crime scene investigation (CSI) personnel: Recruiting for resilience – A mixed methods study. *Journal of Police and Criminal Psychology*, *37*, 549–568. https://doi.org/10.1007/s11896-022-09521-x

Criminal Procedure Rules. (2014). *Rule 33.2*. https://www.justice.gov.uk/courts/procedure-rules/criminal/docs/crim-proc-rules-2014-part-33.pdf

Cullen, F. T., Lemming, T., Link, B. G., & Wozniak, J. F. (1985). The impact of social supports on police stress. *Criminology*, *23*(3), 503–522. https://doi.org/10.1111/j.1745-9125.1985.tb00351.x

Curley, L. J., Munro, J., Lages, M., MacLean, R., & Murray, J. (2020). Assessing cognitive bias in forensic decisions: A review and outlook. *Journal of Forensic Sciences*, *65*(2), 353–360. https://doi.org/10.1111/1556-4029.14220

de Gruijter, M., de Poot, C. J., & Elffers, H. (2016). The influence of new technologies on the visual attention of CSIs performing a crime scene investigation. *Journal of Forensic Sciences*, *61*(1), 43–51. https://doi.org/10.1111/1556-4029.12904

De Kock, J. H., Latham, H. A., Leslie, S. J., Grindle, M., Munoz, S.-A., Ellis, L., Polson, R., & O'Malley, C. M. (2021). A rapid review of the impact of COVID-19 on the mental health of healthcare workers: Implications for supporting psychological well-being. *BMC Public Health*, *21*(1), 104. https://doi.org/10.1186/s12889-020-10070-3

Dedovic, K., Renwick, R., Mahani, N. K., Engert, V., Lupien, S. J., & Pruessner, J. C. (2005). The Montreal Imaging Stress Task: Using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain. *Journal of Psychiatry & Neuroscience*, *30*(5), 319–325. https://doi.org/PMID: 16151536

Deligkaris, P., Panagopoulou, E., Montgomery, A., & Masoura, E. (2014). Job burnout and cognitive functioning: A systematic review. *Work & Stress: An International Journal of Work, Health & Organisations*, *28*(2), 107–123. https://doi.org/10.1080/02678373.2014.909545

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Driskell, T., Driskell, J. E., & Salas, E. (2014). Stress, performance and decision making in organizations. In S. Highhouse, R. S. Dalal, & E. Salas (Eds.), *Judgment and decision making at work* (1st ed., pp. 251–276). Routledge.

Dror, I. E. (2007). Perception of risk and the decision to use force. *Policing: A Journal of Policy and Practice*, *1*(3), 265–272. https://doi.org/10.1093/police/pam041

Dror, I. E. (2009). On proper research and understanding of the interplay between bias and decision outcomes [Letter to the Editor]. *Forensic Science International*, *191*(1–3), e17–e18. https://doi.org/10.1016/j.forsciint.2009.03.012

Dror, I. E. (2015). Cognitive neuroscience in forensic science: Understanding and utilizing the human element. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1674), 20140255. https://doi.org/10.1098/rstb.2014.0255

Dror, I. E. (2016). A Hierarchy of expert performance. *Journal of Applied Research in Memory and Cognition*, *5*(2), 121–127. https://doi.org/10.1016/j.jarmac.2016.03.001

Dror, I. E. (2018). Biases in forensic experts. *Science*, *360*(6386), 243. https://doi.org/10.1126/science.aat8443

Dror, I. E. (2020a). Cognitive and human factors in expert decision making: Six fallacies and the eight sources of bias. *Analytical Chemistry*, *92*(12), 7998–8004. https://doi.org/10.1021/acs.analchem.0c00704

Dror, I. E. (2020b). The error in "error rate": Why error rates are so needed, yet so elusive. *Journal of Forensic Sciences*, *65*(4), 1034–1039. https://doi.org/10.1111/1556-4029.14435

Dror, I. E., & Bucht, R. (2012). Psychological perspectives on problems with forensic science evidence. In B. L. Cutler (Ed.), *Conviction of the innocent: Lessons from psychological research* (pp. 257–276). American Psychological Association. https://doi.org/10.1037/13085-012

Dror, I. E., Busemeyer, J. R., & Basola, B. (1999). Decision making under time pressure: An independent test of sequential sampling models. *Memory & Cognition*, *27*(4), 713–725. https://doi.org/10.3758/BF03211564

Dror, I. E., & Charlton, D. (2006). Why experts make errors. *Journal of Forensic Identification*, *56*(4), 600–616.

Dror, I. E., Charlton, D., & Péron, A. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International*, *156*, 74–78. https://doi.org/10.1016/j.forsciint.2005.10.017

Dror, I. E., & Cole, S. A. (2010). The vision in "blind" justice: Expert perception, judgment, and visual cognition in forensic pattern recognition. *Psychonomic Bulletin & Review*, *17*(2), 161–167. https://doi.org/10.3758/PBR.17.2.161

Dror, I. E., & Hampikian, G. (2011). Subjectivity and bias in forensic DNA mixture interpretation. *Science & Justice*, *51*(4), 204–208. https://doi.org/10.1016/j.scijus.2011.08.004

Dror, I. E., & Langenburg, G. (2019). "Cannot decide": The fine line between appropriate inconclusive determinations versus unjustifiably deciding not to decide. *Journal of Forensic Sciences*, *63*, 1–6. https://doi.org/10.1111/1556-4029.13854

Dror, I. E., McCormack, B., & Epstein, J. (2015). Cognitive bias and its impact on expert witnesses and the court. *The Judges' Journal*, *54*(4), 8–14.

Dror, I. E., Morgan, R. M., Rando, C., & Nakhaeizadeh, S. (2017). Letter to the editor-The bias snowball and the bias cascade effects: Two distinct biases that may impact forensic decision making. *Journal of Forensic Sciences*, *62*(3), 832–833. https://doi.org/10.1111/1556-4029.13496

Dror, I. E., Péron, A. E., Hind, S.-L., & Charlton, D. (2005). When emotions get the better of us: The effect of contextual top-down processing on matching fingerprints. *Applied Cognitive Psychology*, *19*(6), 799–809. https://doi.org/10.1002/acp.1130

Dror, I. E., & Pierce, M. L. (2020). ISO standards addressing issues of bias and impartiality in forensic work. *Journal of Forensic Sciences*, *65*(3), 800–808. https://doi.org/10.1111/1556-4029.14265

Dror, I. E., & Scurich, N. (2020). (Mis)use of scientific measurements in forensic science. *Forensic Science International: Synergy*, *2*, 333–338. https://doi.org/10.1016/j.fsisyn.2020.08.006

Dror, I. E., & Stoel, R. D. (2014). Cognitive forensics: Human cognition, contextual information, and bias. In *Encyclopaedia of Criminology and Criminal Justice* (pp. 353–363). Springer.

Dror, I. E., Wertheim, K., Fraser-Mackenzie, P., & Walajtys, J. (2012). The impact of human-technology cooperation and distributed cognition in forensic science: Biasing effects of AFIS contextual information on human experts. *Journal of Forensic Sciences*, *57*(2), 343–352. https://doi.org/10.1111/j.1556-4029.2011.02013.x

Du, J., Huang, J., An, Y., & Xu, W. (2018). The relationship between stress and negative emotion: The mediating role of rumination. *Clinical Research and Trials*, *4*(1), 1–5. https://doi.org/10.15761/CRT.1000208

Eagle, D. E., Rash, J. A., Tice, L., & Proeschold-Bell, R. J. (2021). Evaluation of a remote, internet-delivered version of the Trier Social Stress Test. *International Journal of Psychophysiology*, *165*, 137–144. https://doi.org/10.1016/j.ijpsycho.2021.03.009

Earwaker, H., Morgan, R. M., Harris, A. J. L., & Hall, L. J. (2015). Fingermark submission decision-making within a UK fingerprint laboratory: Do experts get the marks that they need? *Science & Justice*, *55*(4), 239–247. https://doi.org/10.1016/j.scijus.2015.01.007

Earwaker, H., Nakhaeizadeh, S., Smit, N. M., & Morgan, R. M. (2020). A cultural change to enable improved decision-making in forensic science: A six phased approach. *Science & Justice*, *60*(1), S1355030619300218. https://doi.org/10.1016/j.scijus.2019.08.006

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammer. *Studies in Second Language Acquisition*, *28*(02), 339–368. https://doi.org/10.1017/S0272263106060141

Epel, E. S., Crosswell, A. D., Mayer, S. E., Prather, A. A., Slavich, G. M., Puterman, E., & Mendes, W. B. (2018). More than a feeling: A unified view of stress measurement for population science. *Frontiers in Neuroendocrinology*, *49*, 146–169. https://doi.org/10.1016/j.yfrne.2018.03.001

Evans, J., & Stanovich, K. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psycological Science*, *8*(3), 223–241.

Executive Office of the President's Council of Advisors on Science and Technology. (2016). *Forensic science in criminal Courts: Ensuring scientific validity of feature-comparison methods*. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf

Federal Rules of Evidence. (2020). *Federal Rule 702—Testimony by expert witnesses*. https://www.rulesofevidence.org/article-vii/rule-702/

Ferreira, S. O. (2019). Emotional activation in human beings: Procedures for experimental stress induction. *Psicologia USP*, *30*, e180176. https://doi.org/10.1590/0103-6564e20180176

Field, A. (2018). *Discovering statistics using IBM SPSS statistics*. SAGE Publications.

Forensic Science Regulator. (2015). *Cognitive bias effects relevant to forensic science examinations*. https://www.gov.uk/government/publications/cognitive-bias-effects-relevant-to-forensic-science-examinations

Found, B. (2015). Deciphering the human condition: The rise of cognitive forensics. *Australian Journal of Forensic Sciences*, *47*(4), 386–401. https://doi.org/10.1080/00450618.2014.965204

Fournier, A., Laurent, A., Lheureux, F., Ribeiro-Marthoud, M. A., Ecarnot, F., Binquet, C., & Quenot, J.-P. (2022). Impact of the COVID-19 pandemic on the mental health of professionals in 77 hospitals in France. *PLOS ONE*, *17*(2), e0263666. https://doi.org/10.1371/journal.pone.0263666

Frank, J., Lambert, E. G., & Qureshi, H. (2017). Examining police officer work stress using the job demands–resources model. *Journal of Contemporary Criminal Justice*, *33*(4), 348–367. https://doi.org/10.1177/1043986217724248

Gardner, B. O., Kelley, S., Murrie, D. C., & Dror, I. E. (2019). What do forensic analysts consider relevant to their decision making? *Science & Justice*, S1355030618302867. https://doi.org/10.1016/j.scijus.2019.04.005

Georgiou, N., Morgan, R. M., & French, J. C. (2020). Conceptualising, evaluating and communicating uncertainty in forensic science: Identifying commonly used tools through an interdisciplinary configurative review. *Science & Justice*, S1355030620300046. https://doi.org/10.1016/j.scijus.2020.04.002

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*(1), 451–482. https://doi.org/10.1146/annurev-psych-120709-145346

Gochhayat, J., Giri, V. N., & Suar, D. (2017). Influence of organizational culture on organizational effectiveness: The mediating role of organizational communication. *Global Business Review*, *18*(3), 691–702. https://doi.org/10.1177/0972150917692185

Gok, K., & Atsan, N. (2016). Decision-making under stress and its implications for managerial decision-making: A review of literature. *International Journal of Business and Social Research*, *6*(3), 38–47. https://doi.org/10.18533/ijbsr.v6i3.936

Goldstein, J. Z., & Alesbury, H. S. (2021). Self-reported levels of occupational stress and wellness in forensic practitioners: Implications for the education and training of the forensic workforce. *Journal of Forensic Sciences*, 1556-4029.14699. https://doi.org/10.1111/1556-4029.14699

Government Chief Scientific Adviser. (2015). *Annual report: Forensic science and beyond: Authenticity, provenance and assurance, evidence and case studies*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/506461/gs-15- 37a-forensic-science-beyond-report.pdf

Guez, J., Saar-Ashkenazy, R., Keha, E., & Tiferet-Dweck, C. (2016). The Effect of Trier Social Stress Test (TSST) on item and associative recognition of words and pictures in healthy participants. *Frontiers in Psychology*, *7*, 507. https://doi.org/10.3389/fpsyg.2016.00507

Gunnar, M. R., Reid, B. M., Donzella, B., Miller, Z. R., Gardow, S., Tsakonas, N. C., Thomas, K. M., DeJoseph, M., & Bendezú, J. J. (2021). Validation of an online version of the Trier Social Stress Test in a study of adolescents. *Psychoneuroendocrinology*, *125*, 105111. https://doi.org/10.1016/j.psyneuen.2020.105111

Gutshall, C. L., Hampton, D. P., Sebetan, I. M., Stein, P. C., & Broxtermann, T. J. (2017). The effects of occupational stress on cognitive performance in police officers. *Police Practice and Research*, *18*(5), 463–477. https://doi.org/10.1080/15614263.2017.1288120

Hall, Dollard, M. F., Tuckey, M. R., Winefield, A. H., & Thompson, B. M. (2010). Job demands, work-family conflict, and emotional exhaustion in police officers: A longitudinal test of competing theories. *Journal of Occupational and Organizational Psychology*, *83*(1), 237–250. https://doi.org/10.1348/096317908X401723

Hall, L. J., & Player, E. (2008). Will the introduction of an emotional context affect fingerprint analysis and decision-making? *Forensic Science International*, *181*(1–3), 36–39. https://doi.org/10.1016/j.forsciint.2008.08.008

Hamnett, H. J., & Jack, R. E. (2019). The use of contextual information in forensic toxicology: An international survey of toxicologists' experiences. *Science & Justice*, *59*(4), 380–389. https://doi.org/10.1016/j.scijus.2019.02.004

Harper, D. S. (2022). How to influence positive change? Managers' involvement as emotional architects in the solution for relieving forensic examiners' workplace stress. *Forensic Science International: Synergy*, *4*, 100214. https://doi.org/10.1016/j.fsisyn.2021.100214

Harvie, H. M. K., Jain, B., Nelson, B. W., Knight, E. L., Roos, L. E., & Giuliano, R. J. (2021). Induction of acute stress through an internet-delivered Trier Social Stress Test as assessed by photoplethysmography on a smartphone. *Stress*, *24*(6), 1023–1032. https://doi.org/10.1080/10253890.2021.1995714

Heavey, A. L., Turbett, G. R., Houck, M. M., & Lewis, S. W. (2022). Toward a common language for quality issues in forensic science. *WIREs Forensic Science*, *4*(4), 1–9. https://doi.org/10.1002/wfs2.1452

Helsloot, I., & Groenendaal, J. (2011). Naturalistic decision making in forensic science: Toward a better understanding of decision making by forensic team leaders. *Journal of Forensic Sciences*, *56*(4), 890–897. https://doi.org/10.1111/j.1556-4029.2011.01714.x

Holt, T. J., & Blevins, K. R. (2011). Examining job stress and satisfaction among digital forensic examiners. *Journal of Contemporary Criminal Justice*, *27*(2), 230–250. https://doi.org/10.1177/1043986211405899

Holt, T. J., Blevins, K. R., & Smith, R. W. (2017). Examining the impact of organizational and individual characteristics on forensic scientists' job stress and satisfaction. *Journal of Crime and Justice*, *40*(1), 34–49. https://doi.org/10.1080/0735648X.2016.1216731

Horikawa, M., & Yagi, A. (2012). The relationships among trait anxiety, state anxiety and the goal performance of penalty shoot-Out by university soccer players. *PLoS ONE*, *7*(4), e35727. https://doi.org/10.1371/journal.pone.0035727

Houck, M. M. (2009). Is forensic science a gateway for women in science? *Forensic Science Policy & Management*, *1*(1), 65–69. https://doi.org/10.1080/19409040802629744

Houck, M. M. (2020). Backlogs are a dynamic system, not a warehousing problem. *Forensic Science International: Synergy*, *2*, 317–324. https://doi.org/10.1016/j.fsisyn.2020.10.003

Houck, M. M., & Speaker, P. (2020). Project FORESIGHT: A ten-year retrospective. *Forensic Science International: Synergy*, *2*, 275–281. https://doi.org/10.1016/j.fsisyn.2020.08.005

House of Lords Science and Technology Select Committee. (2019). *Forensic science and the criminal justice system: A blueprint for change*. https://publications.parliament.uk/pa/ld201719/ldselect/ldsctech/333/333.pdf

Howes, L. M. (2015). The communication of forensic science in the criminal justice system: A review of theory and proposed directions for research. *Science and Justice*, *55*(2), 145–154. https://doi.org/10.1016/j.scijus.2014.11.002

Innocence Project. (2022). *DNA exonerations in the United States*. https://innocenceproject.org/overturning-wrongful-convictions-involving-flawed-forensics/

Iorga, M., Soponaru, C., Hanganu, B., & Ioan, B.-G. (2016). The burnout syndrome of forensic pathologists. The influences of personality traits, job satisfaction and environmental factors. *Romanian Journal of Legal Medicine*, *24*(4), 325–332. https://doi.org/10.4323/rjlm.2016.325

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, *38*(12), 1217–1218. https://doi.org/10.1111/j.1365-2929.2004.02012.x

Jeanguenat, A. M., Budowle, B., & Dror, I. E. (2017). Strengthening forensic DNA decision making through a better understanding of the influence of cognitive bias. *Science & Justice*, *57*(6), 415–420. https://doi.org/10.1016/j.scijus.2017.07.005

Jeanguenat, A. M., & Dror, I. E. (2018). Human factors effecting forensic decision making: Workplace stress and well-being. *Journal of Forensic Sciences*, *63*(1), 258–261. https://doi.org/10.1111/1556-4029.13533

Johannes, N., Meier, A., Reinecke, L., Ehlert, S., Setiawan, D. N., Walasek, N., Dienlin, T., Buijzen, M., & Veling, H. (2021). The relationship between online vigilance and affective well-being in everyday life: Combining smartphone logging with experience sampling. *Media Psychology*, *24*(5), 581–605. https://doi.org/10.1080/15213269.2020.1768122

Johnson, S., Cooper, C., Cartwright, S., Donald, I., Taylor, P., & Millet, C. (2005). The experience of work-related stress across occupations. *Journal of Managerial Psychology*, *20*(2), 178–187. https://doi.org/10.1108/02683940510579803

Julian, R., & Kelty, S. F. (2015). Forensic science as "risky business": Identifying key risk factors in the forensic process from crime scene to court. *Journal of Criminological Research, Policy and Practice*, *1*(4), 195–206. https://doi.org/10.1108/JCRPP-09-2015-0044

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. https://doi.org/10.1037/0003-066X.58.9.697

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. https://doi.org/10.1016/j.jarmac.2013.01.001

Kavanagh, J. (2005). *Stress and performance: A review of the literature and its applicability to the military*. https://www.rand.org/pubs/technical_reports/TR192.html

Kellman, P. J., Mnookin, J. L., Erlikhman, G., Garrigan, P., Ghose, T., Mettler, E., Charlton, D., & Dror, I. E. (2014). Forensic comparison and matching of fingerprints: Using quantitative image measures for estimating error rates through understanding and predicting difficulty. *PLoS ONE*, *9*(5), e94617. https://doi.org/10.1371/journal.pone.0094617

Kelty, S. F., & Gordon, H. (2015). No burnout at this coal-face: Managing occupational stress in forensic personnel and the implications for forensic and criminal justice agencies. *Psychiatry, Psychology and Law*, *22*(2), 273–290. https://doi.org/10.1080/13218719.2014.941092

Kelty, S. F., Julian, R., Bruenisholz, E., & Wilson-Wilde, L. (2018). Dismantling the justice silos: Flowcharting the role and expertise of forensic science, forensic medicine and allied health in adult sexual assault investigations. *Forensic Science International*, *285*, 21–28. https://doi.org/10.1016/j.forsciint.2018.01.015

Kelty, S. F., McQueen, E., Pymont, C., & Green, N. (2021). Avoiding burnout at the digital forensics coalface: Targeted strategies for forensic agencies in the management of job-related stress. *Forensic Science International: Digital Investigation*, *38*, 301127. https://doi.org/10.1016/j.fsidi.2021.301127

Keogh, E., & French, C. C. (2001). Test anxiety, evaluative stress, and susceptibility to distraction from threat. *European Journal of Personality*, *15*(2), 123–141. https://doi.org/10.1002/per.400

Kerstholt, J. H. (1994). The effect of time pressure on decision-making behaviour in a dynamic task environment. *Acta Psychologica*, *86*(1), 89–104. https://doi.org/10.1016/0001-6918(94)90013-2

Kerstholt, J. H., Eikelboom, A., Dijkman, T., Stoel, R., Hermsen, R., & van Leuven, B. (2010). Does suggestive information cause a confirmation bias in bullet comparisons? *Forensic Science International*, *198*, 138–142. https://doi.org/10.1016/j.forsciint.2010.02.007

Kerstholt, J. H., Paashuis, R., & Sjerps, M. (2007). Shoe print examinations: Effects of expectation, complexity and experience. *Forensic Science International*, *165*(1), 30–34. https://doi.org/10.1016/j.forsciint.2006.02.039

Kirschbaum, C. (2021). Why we need an online version of the Trier Social Stress Test. *Psychoneuroendocrinology*, *125*, 105129. https://doi.org/10.1016/j.psyneuen.2021.105129

Kirschbaum, C., Diedrich, O., Gehrke, J., Wüst, S., & Hellhammer, D. H. (1991). Cortisol and behavior: The "Trier Mental Challenge Test" (TMCT)—First evaluation of a new psychological stress test. In A. Ehlers, W. Fiegenbaum, I. Florin, & J. Margraf (Eds.), *Perspectives and promises of clinical psychology* (pp. 67–78). Springer.

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test' – A Tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*(1–2), 76–81. https://doi.org/10.1159/000119004

Kobus, H., Houck, M., Speaker, P., Riley, R., & Witt, T. (2011). Managing performance in the forensic sciences: Expectations in light of limited budgets. *Forensic Science Policy & Management: An International Journal*, *2*(1), 36–43. https://doi.org/10.1080/19409044.2011.564271

Koehler, J. J., & Liu, S. (2021). Fingerprint error rate on close non-matches. *Journal of Forensic Sciences*, *66*(1), 129–134. https://doi.org/10.1111/1556-4029.14580

Kornbrot, D. E. (1988). Random walk models of binary choice: The effect of deadlines in the presence of asymmetric payoffs. *Acta Psychologica*, *69*, 109–127.

Kowalski-Trakofler, K. M., Vaught, C., & Scharf, T. (2003). Judgment and decision making under stress: An overview for emergency managers. *International Journal of Emergency Management*, *1*(3), 278–289. https://doi.org/10.1504/IJEM.2003.003297

Kukucka, J., & Dror, I. E. (2022). Human factors in forensic science: Psychological causes of bias and error. In *In D. DeMatteo & K. Scherr (Eds.), The Oxford Handbook of Psychology and Law*. Oxford University Press. https://psyarxiv.com/8pqyt/

Kukucka, J., Dror, I. E., Yu, M., Hall, L., & Morgan, R. M. (2020). The impact of evidence lineups on fingerprint expert decisions. *Applied Cognitive Psychology*, *34*, 1143–1153. https://doi.org/10.1002/acp.3703

Kukucka, J., & Kassin, S. M. (2014). Do confessions taint perceptions of handwriting evidence? An empirical test of the forensic confirmation bias. *Law and Human Behavior*, *38*(3), 256–270. https://doi.org/10.1037/lhb0000066

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. https://doi.org/10.1016/j.jarmac.2017.09.001

Langenburg, G. (2011). Chapter 14: Scientific research supporting the foundations of friction ridge examinations. In *Fingerprint Sourcebook*. National Institute of Justice.

Langenburg, G., Champod, C., & Wertheim, P. (2009). Testing for potential contextual bias effects during the verification stage of the ACE-V methodology when conducting fingerprint comparisons. *Journal of Forensic Sciences*, *54*(3), 571–582. https://doi.org/10.1111/j.1556-4029.2009.01025.x

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer.

Le, J. T., Watson, P., Begg, D., Albertella, L., & Le Pelley, M. E. (2020). Physiological and subjective validation of a novel stress procedure: The Simple Singing Stress Procedure. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01505-1

Leadbetter, M. (2007). Letter to the editor. *Fingerprint Whorld*, *33*, 231.

LeBlanc, V. R., MacDonald, R. D., McArthur, B., King, K., & Lepine, T. (2005). Paramedic performance in calculating drug dosages following stressful scenarios in a human patient simulator. *Prehospital Emergency Care*, *9*(4), 439–444. https://doi.org/10.1080/10903120500255255

Lehman, B. J., Cane, A. C., Tallon, S. J., & Smith, S. F. (2015). Physiological and emotional responses to subjective social evaluative threat in daily life. *Anxiety, Stress, & Coping*, *28*(3), 321–339. https://doi.org/10.1080/10615806.2014.968563

Lesciotto, K. M. (2015). The impact of Daubert on the admissibility of forensic anthropology expert testimony. *Journal of Forensic Sciences*, *60*(3), 549–555. https://doi.org/10.1111/1556-4029.12740

Levy, I., Snell, J., Nelson, A. J., Rustichini, A., & Glimcher, P. W. (2010). Neural representation of subjective value under risk and ambiguity. *Journal of Neurophysiology*, *103*(2), 1036–1047. https://doi.org/10.1152/jn.00853.2009

Lidén, M. (2020). Emotions and cognition in international criminal justice: An exploration from cognitive biases to emotional intelligence. *Forensic Science International: Mind and Law*, *1*, 100037. https://doi.org/10.1016/j.fsiml.2020.100037

Mannering, W. M., Vogelsang, M. D., Busey, T. A., & Mannering, F. L. (2021). Are forensic scientists too risk averse? *Journal of Forensic Sciences*, *66*(4), 1377–1400. https://doi.org/10.1111/1556-4029.14700

Matud, M. P. (2004). Gender differences in stress and coping styles. *Personality and Individual Differences*, *37*(7), 1401–1415. https://doi.org/10.1016/j.paid.2004.01.010

Maule, A. J., Hockey, G. R. J., & Bdzola, L. (2000). Effects of time-pressure on decision-making under uncertainty: Changes in affective state and information processing strategy. *Acta Psychologica*, *104*(3), 283–301. https://doi.org/10.1016/S0001-6918(00)00033-0

Midgley, G., & Lindhult, E. (2021). A systems perspective on systemic innovation. *Systems Research and Behavioral Science*, *38*(5), 635–670. https://doi.org/10.1002/sres.2819

Miller, L. S. (1984). Bias among forensic document examiners: A need for procedural changes. *Journal of Police Science and Administration*, *12*(4), 407–411.

Mnookin, J. L. (2008). The validity of latent fingerprint identification: Confessions of a fingerprinting moderate. *Law, Probability and Risk*, *7*(2), 127–141. https://doi.org/10.1093/lpr/mgm022

Morgan, R. M. (2017a). Conceptualising forensic science and forensic reconstruction. Part I: A conceptual model. *Science & Justice*, *57*(6), 455–459. https://doi.org/10.1016/j.scijus.2017.06.002

Morgan, R. M. (2017b). Conceptualising forensic science and forensic reconstruction. Part II: The critical interaction between research, policy/law and practice. *Science & Justice*, *57*(6), 460–467. https://doi.org/10.1016/j.scijus.2017.06.003

Morgan, R. M., & Bull, P. A. (2007). Forensic geoscience and crime detection: Identification, interpretation and presentation in forensic geoscience. *Minerva Medicolegale*, *127*(2), 73–89.

Morgan, R. M., Earwaker, H., Nakhaeizadeh, S., Harris, A. J. L., Rando, C., & Dror, I. E. (2018). Interpretation of forensic evidence at every step of the forensic science process: Decision-making under uncertainty. In R. Wortley, A. Sidebottom, N. Tilley, & G. Laycock (Eds.), *Routledge Handbook of Crime Science* (pp. 408–420). Routledge.

Morgan, R. M., & Levin, E. A. (2019). A crisis for the future of forensic science: Lessons from the UK of the importance of epistemology for funding research and development. *Forensic Science International: Synergy*, *1*, 243–252. https://doi.org/10.1016/j.fsisyn.2019.09.002

Morgan, R. M., Meakin, G. E., French, J. C., & Nakhaeizadeh, S. (2020). Crime reconstruction and the role of trace materials from crime scene to court. *WIREs Forensic Science*, *2*(1), 1–18. https://doi.org/10.1002/wfs2.1364

Murrie, D. C., Boccaccini, M. T., Guarnera, L. A., & Rufino, K. A. (2013). Are forensic experts biased by the side that retained them? *Psychological Science*, *24*, 1889–1897. https://doi.org/10.1177/0956797613481812

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? *Psychology, Public Policy, and Law*, *15*(1), 19–53.

Murrie, D. C., Gardner, B. O., Kelley, S., & Dror, I. E. (2019). Perceptions and estimates of error rates in forensic science: A survey of forensic analysts. *Forensic Science International*, *302*, 109887. https://doi.org/10.1016/j.forsciint.2019.109887

Mustonen, V., Hakkarainen, K., Tuunainen, J., & Pohjola, P. (2015). Discrepancies in expert decision-making in forensic fingerprint examination. *Forensic Science International*, *254*, 215–226. https://doi.org/10.1016/j.forsciint.2015.07.031

Nakhaeizadeh, S., Morgan, R. M., Rando, C., & Dror, I. E. (2017). Cascading bias of initial exposure to information at the crime scene to the subsequent evaluation of skeletal remains. *Journal of Forensic Sciences*, *63*(2), 403–411. https://doi.org/10.1111/1556-4029.13569

Narvaez Linares, N. F., Charron, V., Ouimet, A. J., Labelle, P. R., & Plamondon, H. (2020). A systematic review of the Trier Social Stress Test methodology: Issues in promoting study comparison and replicable research. *Neurobiology of Stress*, *13*, 100235. https://doi.org/10.1016/j.ynstr.2020.100235

National Academy of Sciences. (2009). *Strengthening forensic science in the United States: A path forward*. The National Academies Press.

National Commision on Forensic Science. (2015). *Ensuring that forensic analysis is based upon task-relevant information*. https://www.justice.gov/archives/ncfs/file/818196/download

National Institute of Justice. (2019). *Report to Congress: Needs assessment of forensic laboratories and medical examiner/coroner offices*. U.S. Department of Justice. https://nij.ojp.gov/library/publications/report-congress-needs-assessment-forensic-laboratories-and-medical

National Institute of Justice. (2020). *Research and evaluation in safety, health, and wellness in the criminal justice system, fiscal year 2020*. U.S. Department of Justice. https://nij.ojp.gov/funding/opportunities/nij-2020-17296

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Office of the Inspector General. (2006). *A review of the FBI's handling of the Brandon Mayfield case*. U.S. Department of Justice. https://oig.justice.gov/special/s0601/final.pdf

Oliver, W. R. (2018). Comment on Kukucka, Kassin, Zapf, and Dror (2017), "Cognitive bias and blindness: A global survey of forensic science examiners." *Journal of Applied Research in Memory and Cognition*, *7*(1), 161. https://doi.org/10.1016/j.jarmac.2018.01.008

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872. https://doi.org/10.1016/j.jesp.2009.03.009

Osborne, N. K. P., Woods, S., Kieser, J., & Zajac, R. (2014). Does contextual information bias bitemark comparisons? *Science & Justice*, *54*(4), 267–273. https://doi.org/10.1016/j.scijus.2013.12.005

Osborne, N. K. P., & Zajac, R. (2015). An imperfect match? Crime-related context influences fingerprint decisions: Contextual bias in fingerprint decisions. *Applied Cognitive Psychology*, *30*(1), 126–134. https://doi.org/10.1002/acp.3180

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, *11*(6), 988–1010. https://doi.org/10.3758/BF03196730

Page, M., Taylor, J., & Blenkin, M. (2012). Context effects and observer bias-implications for forensic odontology. *Journal of Forensic Sciences*, *57*(1), 108–112. https://doi.org/10.1111/j.1556-4029.2011.01903.x

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Paton, D., & Flin, R. (1999). Disaster stress: An emergency management perspective. *Disaster Prevention and Management: An International Journal*, *8*(4), 261–267. https://doi.org/10.1108/09653569910283897

Patterson, G. T. (2003). Examining the effects of coping and social support on work and life stress among police officers. *Journal of Criminal Justice*, *31*(3), 215–226. https://doi.org/10.1016/S0047-2352(03)00003-5

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Peterson, J. L., Hickman, M. J., Strom, K. J., & Johnson, D. J. (2013). Effect of forensic evidence on criminal justice case processing. *Journal of Forensic Sciences*, *58*(S1), S78–S90. https://doi.org/10.1111/1556-4029.12020

Puzzo, I., Aldridge-Waddon, L., Stokes, N., Rainbird, J., & Kumari, V. (2022). The impact of the COVID-19 pandemic on forensic mental health services and clinical outcomes: A longitudinal study. *Frontiers in Psychiatry*, *12*, 780236. https://doi.org/10.3389/fpsyt.2021.780236

Quick, J., & Henderson, D. (2016). Occupational stress: Preventing suffering, enhancing wellbeing. *International Journal of Environmental Research and Public Health*, *13*(5), 459. https://doi.org/10.3390/ijerph13050459

Raptis, S., Chen, J. N., Saposnik, F., Pelyavskyy, R., Liuni, A., & Saposnik, G. (2017). Aversion to ambiguity and willingness to take risks affect therapeutic decisions in managing atrial fibrillation for stroke prevention: Results of a pilot study in family physicians. *Patient Preference and Adherence*, *11*, 1533–1539. https://doi.org/10.2147/PPA.S143958

Raymond, T., & Julian, R. (2015). Forensic intelligence in policing: Organisational and cultural change. *Australian Journal of Forensic Sciences*, *47*(4), 371–385. https://doi.org/10.1080/00450618.2015.1052759

Reips, U.-D. (2000). The Web experiment method: Advantages, disadvantages, and solutions. *Psychological Experiments on the Internet*, 89–117. https://doi.org/10.1016/B978-012099980-4/50005-8

Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science*, *13*(2), 60–66. https://doi.org/10.1111/j.0963-7214.2004.00275.x

Riedl, R. (2022). On the stress potential of videoconferencing: Definition and root causes of Zoom fatigue. *Electronic Markets*, *32*, 153–177. https://doi.org/10.1007/s12525-021-00501-3

Risinger, D. M., Saks, M. J., Thompson, W. C., & Rosenthal, R. (2002). The Daubert/ Kumho implications of observer effects in forensic science: Hidden problems of expectation and suggestion. *California Law Review*, *90*(1), 56.

Roux, C., Bucht, R., Crispino, F., De Forest, P., Lennard, C., Margot, P., Miranda, M. D., NicDaeid, N., Ribaux, O., Ross, A., & Willis, S. (2022). The Sydney declaration – Revisiting the essence of forensic science through its fundamental principles. *Forensic Science International*, *332*, 111182. https://doi.org/10.1016/j.forsciint.2022.111182

Roux, C., Crispino, F., & Ribaux, O. (2012). From forensics to forensic science. *Current Issues in Criminal Justice*, *24*(1), 7–24. https://doi.org/10.1080/10345329.2012.12035941

Roux, C., & Weyermann, C. (2020). Can forensic science learn from the COVID-19 crisis? *Forensic Science International*, *316*, 110503. https://doi.org/10.1016/j.forsciint.2020.110503

Roux, C., Willis, S., & Weyermann, C. (2021). Shifting forensic science focus from means to purpose: A path forward for the discipline? *Science & Justice*, *61*(6), 678–686. https://doi.org/10.1016/j.scijus.2021.08.005

Saks, M. J. (2009). Concerning L.J. Hall, E. Player, "Will the introduction of an emotional context affect fingerprint analysis and decision-making?" [Forensic Sci. Int. 181 (2008) 36–39]. *Forensic Science International*, *191*(1–3), e19. https://doi.org/10.1016/j.forsciint.2009.06.011

Saks, M. J., & Koehler, J. (2005). The coming paradigm shift in forensic identification science. *Science*, *309*(5736), 892–895. https://doi.org/10.1126/science.1111565

Saks, M. J., & Koehler, J. (2008). The individualization fallacy in forensic science evidence. *Vanderbilt Law Review*, *61*(1), 199–219.

Saks, M. J., Risinger, D. M., Rosenthal, R., & Thompson, W. C. (2003). Context effects in forensic science: A review and application of the science of science to crime laboratory practice in the United States. *Science & Justice*, *43*(2), 77–90. https://doi.org/10.1016/S1355-0306(03)71747-X

Salinas, C. R., & Webb, H. E. (2018). Occupational stress and coping mechanisms in crime scene personnel. *Occupational Medicine*, *68*(4), 239–245. https://doi.org/10.1093/occmed/kqy030

Saposnik, G., Sempere, A. P., Prefasi, D., Selchen, D., Ruff, C. C., Maurino, J., & Tobler, P. N. (2017). Decision-making in multiple sclerosis: The role of aversion to ambiguity for therapeutic inertia among neurologists (DIScUTIR MS). *Frontiers in Neurology*, *8*(65), 1–8. https://doi.org/10.3389/fneur.2017.00065

Schaufenbuel, K. (2015). *Why Google, Target, and General Mills are investing in mindfulness*. https://hbr.org/2015/12/why-google-target-and-general-mills-are-investing-in-mindfulness#:~:text=Google%20believes%20that%20these%20mindfulness,stress%20and%20improve%20mental%20focus.

Schuetz, M., Gockel, I., Beardi, J., Hakman, P., Dunschede, F., Moenk, S., Heinrichs, W., & Junginger, Th. (2008). Three different types of surgeon-specific stress reactions identified by laparoscopic simulation in a virtual scenario. *Surgical Endoscopy*, *22*(5), 1263–1267. https://doi.org/10.1007/s00464-007-9605-1

Scurich, N., & Dror, I. E. (2020). Continued confusion about inconclusives and error rates: Reply to Weller and Morris. *Forensic Science International: Synergy*, *2*, 703–704. https://doi.org/10.1016/j.fsisyn.2020.10.005

Seigfried-Spellar, K. C. (2018). Assessing the psychological well-being and coping mechanisms of law enforcement investigators vs. Digital forensic examiners of child pornography investigations. *Journal of Police and Criminal Psychology*, *33*(3), 215–226. https://doi.org/10.1007/s11896-017-9248-7

Sharma, N., Chakrabarti, S., & Grover, S. (2016). Gender differences in caregiving among family—Caregivers of people with mental illnesses. *World Journal of Psychiatry*, *6*(1), 7–17. https://doi.org/10.5498/wjp.v6.i1.7

Smalarz, L., Madon, S., Yang, Y., Guyll, M., & Buck, S. (2016). The perfect match: Do criminal stereotypes bias forensic evidence analysis? *Law and Human Behavior*, *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. *Science & Justice*, *58*(2), 128–137. https://doi.org/10.1016/j.scijus.2017.09.005

Söderkvist, S., Ohlén, K., & Dimberg, U. (2018). How the experience of emotion is modulated by facial feedback. *Journal of Nonverbal Behavior*, *42*(1), 129–151. https://doi.org/10.1007/s10919-017-0264-1

Sok, J., Blomme, R., & Tromp, D. (2014). Positive and negative spillover from work to home: The role of organizational culture and supportive arrangements. *British Journal of Management*, *25*(3), 456–472. https://doi.org/10.1111/1467-8551.12058

Spector, P. E. (2012). *Industrial and organizational psychology: Research and practice* (6th ed.). Wiley.

Spellman, B. A., Eldridge, H., & Bieber, P. (2022). Challenges to reasoning in forensic science decisions. *Forensic Science International: Synergy*, *4*, 100200. https://doi.org/10.1016/j.fsisyn.2021.100200

Spielberger, C. D., Gorsuch, R., Lushene, R., Vagg, P., & Jacobs, G. (1983). *Manual for the state–trait anxiety inventory* [Data set]. Consulting Psychologists Press. https://doi.org/10.1037/t06496-000

Stacey, R. (2004). A report on the erroneous fingerprint individialization in the Madrid train bombing case. *Journal of Forensic Identification*, *54*(6), 706–718.

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 Science International*, *276*, 93–106. https://doi.org/10.1016/j.forsciint.2017.04.009

Stevenage, S. V., & Pitfield, C. (2016). Fact or friction: Examination of the transparency, reliability and sufficiency of the ACE-V method of fingerprint analysis. *Forensic Science International*, *267*, 145–156. https://doi.org/10.1016/j.forsciint.2016.08.026

Stewart, N., Chandler, J., & Paolacci, G. (2017). Crowdsourcing samples in cognitive science. *Trends in Cognitive Sciences*, *21*(10), 736–748. https://doi.org/10.1016/j.tics.2017.06.007

Sunde, N., & Dror, I. E. (2019). Cognitive and human factors in digital forensics: Problems, challenges, and the way forward. *Digital Investigation*, *29*, 101–108. https://doi.org/10.1016/j.diin.2019.03.011

Tanida, M., Katsuyama, M., & Sakatani, K. (2007). Relation between mental stress-induced prefrontal cortex activity and skin conditions: A near-infrared spectroscopy study. *Brain Research*, *1184*, 210–216. https://doi.org/10.1016/j.brainres.2007.09.058

Taylor, M. C., Laber, T. L., Kish, P. E., Owens, G., & Osborne, N. K. P. (2016a). The reliability of pattern classification in bloodstain pattern analysis, Part 1: Bloodstain patterns on rigid non-absorbent surfaces. *Journal of Forensic Sciences*, *61*(4), 922–927. https://doi.org/10.1111/1556-4029.13091

Taylor, M. C., Laber, T. L., Kish, P. E., Owens, G., & Osborne, N. K. P. (2016b). The reliability of pattern classification in bloodstain pattern analysis-PART 2: Bloodstain patterns on fabric surfaces. *Journal of Forensic Sciences*, *61*(6), 1461–1466. https://doi.org/10.1111/1556-4029.13191

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving Decisions About Health, Wealth, and Happiness* (1st ed.). Penguin.

The Fingerprint Inquiry. (2011). https://www.webarchive.org.uk/wayback/archive/20170104175208/http://www.webarchive.org.uk/wayback/archive/20150428160012/http:/www.thefingerprintinquiryscotland.org.uk/inquiry/21.html

Thompson, M. B., & Tangen, J. M. (2014). The nature of expertise in fingerprint matching: Experts can do a lot with a little. *PLoS ONE*, *9*(12), e114759. https://doi.org/10.1371/journal.pone.0114759

Towler, A., White, D., Ballantyne, K., Searston, R. A., Martire, K. A., & Kemp, R. I. (2018). Are forensic scientists experts? *Journal of Applied Research in Memory and Cognition*, *7*(2), 199–208. https://doi.org/10.1016/j.jarmac.2018.03.010

Ulery, B. T., Hicklin, R. A., Buscaglia, J., & Roberts, M. A. (2011). Accuracy and reliability of forensic latent fingerprint decisions. *Proceedings of the National Academy of Sciences*, *108*(19), 7733–7738. https://doi.org/10.1073/pnas.1018707108

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), e32800. https://doi.org/10.1371/journal.pone.0032800

Ulery, B. T., Hicklin, R. A., Roberts, M. A., & Buscaglia, J. (2017). Factors associated with latent fingerprint exclusion determinations. *Forensic Science International*, *275*, 65–75. https://doi.org/10.1016/j.forsciint.2017.02.011

Upadhyay, U. D., & Lipkovich, H. (2020). Using online technologies to improve diversity and inclusion in cognitive interviews with young people. *BMC Medical Research Methodology*, *20*(1), 159. https://doi.org/10.1186/s12874-020-01024-9

US Bureau of Justice Assistance. (2019). *Triage of forensic evidence testing: A guide for prosecutors* (No. 254513). https://bja.ojp.gov/library/publications/triage-forensic-evidence-testing-guide-prosecutors

van den Eeden, C. A. J., de Poot, C. J., & van Koppen, P. J. (2019). The forensic confirmation bias: A comparison between experts and novices. *Journal of Forensic Sciences*, *64*(1). https://doi.org/10.1111/1556-4029.13817

VanderKolk, J. R. (2011). Chapter 9: Examination process. In *Fingerprint Sourcebook*. National Institute of Justice.

Webb, D. A., Sweet, D., & Pretty, I. A. (2002). The emotional and psychological impact of mass casualty incidents on forensic odontologists. *Journal of Forensic Sciences*, *47*(3), 539–541. https://doi.org/10.1520/JFS2001330

Weller, T. J., & Morris, M. D. (2020). Commentary on: I. Dror, N Scurich "(Mis)use of scientific measurements in forensic science" Forensic Science International: Synergy 2020 https://doi.org/10.1016/j.fsisyn.2020.08.006. *Forensic Science International: Synergy*, *2*, 701–702. https://doi.org/10.1016/j.fsisyn.2020.10.004

Wigginton, N. S., Cunningham, R. M., Katz, R. H., Lidstrom, M. E., Moler, K. A., Wirtz, D., & Zuber, M. T. (2020). Moving academic research forward during COVID-19. *Science*, *368*(6496), 1190–1192. https://doi.org/10.1126/science.abc5599

Wood, S. J., & Michaelides, G. (2016). Challenge and hindrance stressors and wellbeing-based work–nonwork interference: A diary study of portfolio workers. *Human Relations*, *69*(1), 111–138. https://doi.org/10.1177/0018726715580866

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1002/cne.920180503

Yoo, Y.-S., Cho, O.-H., Cha, K.-S., & Boo, Y.-J. (2013). Factors influencing post-traumatic stress in Korean forensic science investigators. *Asian Nursing Research*, *7*(3), 136–141. https://doi.org/10.1016/j.anr.2013.07.002

Yu, R., Yang, L., Guo, X., & Zhang, Y. (2015). Effect of time pressure on dynamic visual search performance. *Procedia Manufacturing*, *3*, 4658–4664. https://doi.org/10.1016/j.promfg.2015.07.556

Zapf, P. A., & Dror, I. E. (2017). Understanding and mitigating bias in forensic evaluation: Lessons from forensic science. *International Journal of Forensic Mental Health*, *16*(3), 227–238. https://doi.org/10.1080/14999013.2017.1317302

Zavala, A. M., Day, G. E., Plummer, D., & Bamford-Wade, A. (2018). Decision-making under pressure: Medical errors in uncertain and dynamic environments. *Australian Health Review*, *42*(4), 395. https://doi.org/10.1071/AH16088

Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, *111*(4), 493–504. https://doi.org/10.1037/pspa0000056

Zimmer, P., Buttlar, B., Halbeisen, G., Walther, E., & Domes, G. (2019). Virtually stressed? A refined virtual reality adaptation of the Trier Social Stress Test (TSST) induces robust endocrine responses. *Psychoneuroendocrinology*, *101*, 186–192. https://doi.org/10.1016/j.psyneuen.2018.11.010

Zou, J., Liu, S., Sun, Q., Wang, C., & Liu, Y. (2021). How time pressure and fingerprint complexity affect fingerprint examiner performance in an Eye Tracking Study. *Forensic Science International*, *328*, 111007. https://doi.org/10.1016/j.forsciint.2021.111007

Appendix A: Questionnaire for Chapter 3

Please rate the following statements. It is totally anonymous, so please be honest.

	1 (Low)	2	3	4	5	6	7 (High)
1. In the past year, I often felt stressed while at work.							
2. The stress I felt originated from personal reasons (e.g., family, medical and/ or financial).							
3. The stress I felt originated from management and/or supervisors.							
4. The stress I felt originated from backlogs and pressure to do many cases.							
5. I get feedback about my work (e.g., from management, supervisors, police investigators and/or legal advocates.							
6. I feel management, supervisors, police investigators and/ or legal advocates appreciated it more when I help to solve a case (e.g., when I find a 'match' rather than 'inconclusive').							
7. Sometimes I feel I know what management, supervisors, police investigators and/ or legal advocates want or expect my conclusion to be.							

Which section do you work at (e.g., DNA, firearms, latent prints, etc)?____________
Years of experience:___________
Thank you.

Appendix B: Supplementary Materials for Chapter 5

Appendix B-1: The three phases in developing the online stress method

This study presents a method developed for inducing stress online, but without the presence of researchers. The development of this stress method—adapted from the Trier Mental Challenge Test Stress Protocol originally developed by Kirschbaum et al. (1991)—had three phases. In Phase I, stress stimuli that consisted of mathematical and general knowledge questions was selected. In Phase II, the stress stimuli were piloted to determine the mean response time for each question and to determine the accuracy rate for each question. In Phase III, the stress elements, such as calculating time deadlines for answering the questions based on the mean response times in Phase II, were identified.

Phase I – Selection of Stress Stimuli

Two main categories of questions were selected for the online method: mathematical questions (N = 30) and general knowledge questions (N = 30). Because the study was conducted online and without the presence of researchers, it was critical to have a mix of questions for inducing stress. This way, participants were prevented from anticipating what they might be asked, to minimise the possibility of participants finding the answers through an online search, and as a mechanism for sustaining the motivation of the participants to complete the exercise (Yip, 2004).

Three genres of mathematical questions were selected: “tricky” (10 questions), numerical reasoning (8 questions), and basic arithmetic (12 questions; see Table 1 for example questions and Appendix B-2 for the full list of questions). The mathematics questions were carefully selected so that they would have varying levels of difficulties and could be solved mentally (i.e., without a calculator and without the use of a paper and pencil).

The ‘tricky’ genre consisted of questions that seem straightforward and intuitive but are rarely answered correctly (Gardner, 1986; Kahneman, 2011). The numerical reasoning questions required basic knowledge of mathematics and have been adapted from numerical reasoning tests used to test job applicants (Smith, 2017). These two genres of questions can be mentally challenging because they require more than merely a calculator (given that participants can have access to calculators). The third genre— the basic arithmetic questions— involves fundamental arithmetical operations, such as addition and multiplication, and resemble the mathematics questions used in the Mental Challenge Test (Allendorfer et al., 2019; Dedovic et al., 2005; Kirschbaum et al., 1991). To increase the difficulty level, no parentheses were added within the equations.

All the mathematical questions required free-text entry answers, and the answers are intended to be whole numbers (without fractions) to minimise the opportunity for participant confusion and to simplify coding for the answers (for example, if fractions were used, then both 2.5 and 2 1/2 would need to be coded as correct answers). Participants were instructed to input only the numerical answer for each question without a description (for example, "bananas" in question 7) or a unit (for example, "pounds" in question 18). A content validation was programmed for all the mathematical questions to ensure that participants only provided numeric answers.

The general knowledge questions were all multiple-choice questions derived from two sources. Sixteen questions came from a bank of questions used in professional workshops to which participants often provide wrong answers (see questions 31–46 in Appendix B-2). The other 14 questions were identified through an online web search for 'easy general knowledge questions' (see questions 47–60 in Appendix B-2) in order to include a set of general knowledge questions of varying difficulty.

Phase II – Piloting the Stress Stimuli

Overview

A pilot study was undertaken in which participants were asked to answer all the selected questions in Phase I with no time limit and no feedback. The pilot study had two aims. First to assess the mean response time of each question in order to calculate a preset time deadline for the stress group. The second to select questions to allocate for the stress and control conditions of the experiment: the stress group was to be provided with questions that were often answered incorrectly. This would increase the likelihood of the stress group receiving negative feedback (such as "wrong").

Recruitment of Participants

The participants were recruited using the Prolific crowdsourcing platform widely used for academic research (Clemmow et al., 2020; Palan & Schitter, 2018; Peer et al., 2017). The participant selection criteria were 25–60 years of age with a minimum level of high school (or equivalent) education. The online experiments were developed and deployed using Qualtrics.

Table 1: Examples of mathematical and general knowledge questions.

Genre	Question	Answer	Source
'Tricky' Math Questions	Q.1. A pen and pad cost one dollar and ten cents. The pen costs one dollar more than the pad. How much does the pad cost (in cents)? (*please type the number only*).	5 cents	(Kahneman, 2011)
	Q.5 A person was born on May 6, 30 B.C. He died on May 6, 30 A.D. How old was he when he died? (*please type the number only*).	59 years	(Gardner, 1986)
Numerical Reasoning	Q12. The Arsenal football club "games won" to "games lost" record last season was 2:3. How many games did they play last season if all the games were either won or lost and Arsenal won 6 games? (*please type the number only*).	15 games	(Smith, 2017)
	Q15. A plant grows by 5% each year. Its height was 90 cm when it was planted. In which year will the plant exceed 1 m in height. (*please type the number only*).	Year 3	(Smith, 2017)

Basic Arithmetic	Q19. What is the answer? $3 \times 4 - 9 =$ Q28. What is the answer? $4 \times 3^2 \div 1^4 + 4 =$	3 40	These questions follow the paradigm of the Mental Challenge Stress Protocol (Kirschbaum et al., 1991).
General Knowledge	Q32. Which ocean goes to the deepest depths? A. Pacific B. Arctic C. Atlantic D. Indian E. Southern	A. Pacific	Taken from a bank of questions that have already been used in professional workshops.
	Q49. Which planet is known as the Red Planet? A. Neptune B. Jupiter C. Mars D. Mercury E. Earth	C. Mars	Found in Edsys (2020) by searching the internet for "easy general knowledge questions"

Strategy for Response Quality

Participants recruited from online crowdsourcing platforms generally have high-quality responses and low "cheating" rates (Berinsky et al., 2012; Clifford & Jerit, 2016; Motta et al., 2016). For example, it has been identified that students were more inclined to look up the answers in web-based political knowledge surveys than crowdsourced users of Mechanical Turk, potentially due to crowdsourced participants having financial incentives to finish quickly rather than spending time looking up answers (Clifford & Jerit, 2016). As subsequent phases in developing the online stress method presented here depended on the responses from this pilot study, measures were included to prevent and detect low-quality responses. These measures aimed to improve the internal validity by preventing and detecting cheating or random responses.

Commitment Statement

Clifford and Jerit (2016) compared different methods for reducing cheating in online surveys (such as directly asking participants not to cheat or asking them to willingly commit to not use outside help). They found that a commitment request to participants with "yes" and "no" answers resulted in the lowest degrees of cheating and the greatest predictive validity. In addition, a commitment statement can avoid the unpredicted disadvantage regarding participants' goodwill, such as when asking the participants directly not to cheat. Hence, a commitment statement using language adapted from Clifford and Jerit (2016) was used as a prevention measure to cheating:

> *Commitment Statement:* Data of this scientific study may ultimately contribute to improving the decision-making in forensic science. Hence, it is important to us that you do NOT use outside sources, like a calculator or the internet. Will you answer the following questions without help from outside sources? (Yes / No)

Question Wording Alternations

Previous research suggested that "Googling cheaters" may look for factual answers online while completing the survey (Jensen & Thomsen, 2014). Such behaviour is prevented or reduced in face-to-face surveys (Heerwegh & Loosveldt, 2008). Yet, because this study was conducted completely online, a new solution that has the potential to minimise online search activities was warranted.

The solution utilised for this study entailed simple alterations to the question wording. Specifically, nonessential information or words in the question (such as names or dates) were amended, when possible, whereas the essential information was retained to increase the difficulty of looking for answers through an internet search.

In the example in Table 2, the name and title "Larygitis, a Greek orator" were nonessential to answering the question; hence, they were removed from the original question. Moreover, the day and month of birth were amended. The answer would still be the same—59 years old. In this example, when copying the original question for an online search, the question is found in the first search line as well as the source of the original question, but the amended question and its source do not appear in the entire first page of the online search.

Amending the nonessential information in the original questions was only done for the numeric reasoning and tricky questions. It was not possible to amend the words in the general knowledge questions because these questions did not typically contain nonessential information.

Table 2: Example for amending the nonessential words of tricky question number 5.

Original Question	Amended Question
Larygitis, a Greek orator, was born on July 4, 30 B.C. He died on July 4, 30 A.D. How old was he when he died?	A person was born on May 6, 30 B.C. He died on May 6, 30 A.D. How old was he when he died?

Image Not Text

Researchers in the political science literature have discussed that image-based knowledge questions on individuals—as opposed to text-only questions—can be a useful way to minimise looking up answers online (Motta et al., 2017; Prior, 2014). Taking this approach, in the current study, the basic arithmetic questions were coded in Qualtrics as images, not text. This prevention measure did not allow participants to simply copy and paste the formulas online to find the answers.

Randomisation of the Questions

Evidence suggests that some crowdsourced users may share experimental tasks on unofficial online forums where they can interact with other users and provide information on specific tasks (Schmidt, 2015). Randomisation of questions has been suggested as a potential way to reduce cheating through collaborations with others (McLeod et al., 2003) and can help in minimising the order effect (Oldendick 2008). Therefore, in this method, the order of questions was randomised using Qualtrics.

High Approval Rates

Prolific offers a prescreening tool called "approval rate" for participants, which is the number of studies approved by researchers divided by the number of total studies completed by the participant. Therefore, the approval rate can be an indicative measure regarding the quality of data a Prolific participant may offer. It has been found that restricting participation to participants with high approval rates (specifically, above 95% approval ratings) is an effective method to enhance data quality (Peer et al., 2014). Therefore, for this study, only participants with a high approval rate of 95% or above were chosen.

Attention Check Questions

Attention check screeners are questions that are typically used to check whether participants read the instructions. They are embedded within the exercise and look similar to experimental stimuli in terms of length and format (such as the number of multiple-choice options). However, the answers may not follow a normal format or expectations, so participants will need to read the instructions to answer them (Oppenheimer et al., 2009; see Table 3). Some researchers have argued that attention checks may affect the responses of subsequent questions because they are in essence 'trick questions' (Kane & Barabas, 2019; Hauser et al., 2018).

However, given the nature of the subject pool (see Peer et al. (2017) on attention check failures from croudsourced participants), they are still being used in empirical research as useful tools to detect participant attentiveness to the study tasks (Clemmow et al., 2020). In addition, some researchers recommend using more than one attention check question to improve data quality

(Berinsky et al., 2014). Therefore, two questions were used for Phase II to minimise potential impact, if any, on the overuse of attention checks to the responses of actual stress stimuli (see Table 3). One attention question resembled a general knowledge question, and one attention check question resembled a numerical reasoning/tricky question. An attention check question resembling basic arithmetic can be easily detected by the participants so no attention check question that looked similar to the basic mathematical questions was included.

Table 3: Attention check questions.

Attention Check Question	**Answer**	**Notes**
What color is the sky? Please select "purple" as an answer to make sure you are paying attention. A. White B. Blue C. Green D. Purple E. Black	D. Purple	Resembles general knowledge questions.
There are seven orange trees on a farm. These trees double in number every 6 months. Please type "orange" as an answer to this question to make sure that you are paying attention. What is the answer?	Orange	Resembles numerical reasoning and tricky questions.

Explicit Instructions

Providing clear instructions can enhance the understanding of the experimental tasks by reducing ambiguity which can enhance the quality of responses (Alekseev et al., 2017). In this method, the language of the instructions was carefully written to minimise ambiguity regarding the tasks the participants were being asked to complete (e.g., the instructions made it clear that the use of a calculator and paper/pencil were not allowed):

Instructions: Please read the instructions carefully before proceeding:

1. You are asked to answer a number of questions to the best of your ability.

2. The questions in the study may include general knowledge questions, numerical reasoning questions, and so on.

3. The study should take approximately 45 min but not more than 60 min.

4. The use of a calculator is NOT permitted.

5. The use of a paper and pencil is NOT permitted.

6. Please rely on your own personal knowledge WITHOUT using internet searches.

7. You are asked to be complete the study in a single time session.

8. Please switch off distractions (like emails) and find a quiet place to focus on this study.

9. Please only take part in this study if you are using a desktop/laptop computer.

Bot Reponses

Computer programs may be used in crowdsourcing platforms to complete the tasks automatically, known as a bot response (Chmielewski & Kucker, 2020). Some researchers recommend using Captcha Verification (or Completely Automated Public Turing Test to tell Computers and Humans Apart) to ensure that the responses come from real human participants, not programs (Chmielewski & Kucker, 2020). For this study, one captcha verification question from Qualtrics was included at the beginning of the survey. The participants were presented with a challenge consisting of an image of words or characters that they needed to correctly complete before proceeding.

Elimination of Low-Quality Responses

The responses of all Prolific participants were accepted, and the participants were paid for their time. However, the plan was to carefully check the data for each participant considering multiple factors to decide whether or not to keep their data for the analysis (see Table 4).

The specific criteria outlined in Table 4 helped "red-flag" potential low-quality responses. Red-flagged responses were investigated thoroughly to determine whether there was a consistent pattern of low-quality response across different questions. Thus, the aim of the elimination criteria was not to provide quantifiable measures to eliminate low-quality responses or cheaters (Berinsky et al., 2014), but to gain a more holistic view of the quality of responses for each participant.

Table 4: Criteria for eliminating low-quality data from analysis.

Attention check	The participant fails more than one attention check question. Evidence suggests that elimination of data based on a single screener may result in bias in research (Berinsky et al., 2014).
Length of time	Consistently very high response times. That is, response times are repetitively much higher than those of other participants across multiple mathematical/general knowledge questions. Such response behaviours raise concerns that that they search for answers on the internet. For instance, Clifford and Jerit (2016) found that self-reported cheaters spent significantly more time in answering the general knowledge questions than non-cheaters. Consistently very low response times. This response behaviour raises concerns that participants randomly answer the questions, without cognitive effort (Börger, 2016).
Pattern of subjects' responses	High number of correct answers on difficult questions (i.e., more than just chance).

	Consistently choosing the same responses. For example, choosing the same option in multiple choice questions (Clemmow et al., 2020).

Results of Phase II

Quality of Responses

In total, 35 participants were recruited via Prolific. The responses of five participants were excluded from the analysis, leaving a final sample of 30 participants (60% male, n = 18; age range = 25–47 years; M = 33.53 years, SD = 6.43 years). Data of the five participants were excluded due to their responses raising concerns; they showed a consistent pattern of low-quality responses based on the length of time they spent on the questions. The responses of two participants were consistently very fast, raising concerns over the cognitive effort made in answering the questions (Börger, 2016). The response times were even too short for potentially reading a question and thinking about the answer (for example, one participant spent 2.2 s, 2.7 s, and 2.8 s on questions 45, 41, and 42).

Conversely, the responses of the other three participants were consistently extremely slow with the response times of those participants being very high for a number of simple general knowledge questions (for example one participant spent around 3 min to answer question 49 on the longest river in the world). This increased response time raised concerns that these participants could be looking up answers via the web (Clifford & Jerit, 2016) and so they were excluded.

Only one participant failed one of the two attention check screeners. After careful assessment of the quality of the responses from that participant (see Table 4), their data were not excluded so that data were not eliminated based on a single screener as this may result in bias (Berinsky et al., 2014; Clemmow et al., 2020).

Elimination of Outliers

To calculate reasonable time limits for answering the questions, two established methods were compared. The first method sets a deadline by subtracting one standard deviation from the mean response time (Benson & Beach, 1996; Tsiga et al., 2013). The second method uses 70% of the established mean response times (Kellogg et al., 1999). It is critical that established mean response times are not inflated as having outliers can potentially make the means (or the standard deviations) higher or lower than they should be. For example, if the standard deviation method was applied for the current study without any data treatment, then some of the deadlines would be too short (for example, a 1- or 2-s deadline) or would be negative (when the standard deviation is higher than the mean). Thus, in this study the outliers were excluded.

Ratcliff (1993) developed simulations of response times to test the different methods used in removing response time outliers based on the central tendency approach—a common method of response time data treatment (Whelan, 2008). Therefore in this study response-time outliers were eliminated following Ratcliff (1993):

1. Cutoffs should be chosen as a function of proportions of responses eliminated. Based on the reaction time data simulations, it was recommended that a reasonable range to choose cutoffs should aim to eliminate 5–15% of data (i.e., keep the central 85–95% of the data). In the current study, only 7.6% response time data points were excluded. This exclusion rate falls within the lower end of the recommended range.

2. When there is a high variability among the subjects' response time means, as in the current study, then an elimination method using a standard deviation cutoff is recommended (Ratcliff, 1993). One of the cutoffs tested was 1.5 absolute standard deviation above the mean, which yielded acceptable results. This value (i.e., absolute z score = 1.5 *SD* or higher) was chosen for the current experiment because, upon preliminary assessment, this cutoff value allows for excluding 5–15% of the data (see recommendation 1, above).

In summary, 136 response time data points were at, or above, 1.5 standard deviation from the mean and thus were excluded (7.6%, $N = 1{,}800$). If the response times were considered an outlier using the method above, then, both the reaction times and the actual responses to the question were eliminated.

Accuracy and Mean Response Times

Appendix B-2 show the accuracies and mean response times of all questions tested (N = 60). Accuracy (%) refers to the number of correct responses by number of total responses (i.e., N after removing outliers). The questions were divided into either the stress condition (N = 30) or the control condition (N = 30) based on accuracy. Stress questions were of the lowest accuracy for each type of question (see questions in bold in Appendix B-2).

Phase III – Inclusion of Stress Elements

After the stress stimuli were tested in the pilot study, the following elements were included in the stress method:

Time Limits

Having time limits to answer the question increases the sense of uncontrollability (Allen et al., 2017; Dickerson & Kemeny, 2004). There are different approaches to setting time limits such as having a fixed deadline for the participants to answer a block of mathematical questions (Kirschbaum et al., 1991) or using a program that continuously increases/decreases time limits depending on the subject's performance (Dedovic et al., 2005). For this study, each question had a precalculated time limit that was the same across participants, regardless of performance. Using this approach, the participants would not be able to predict the deadlines for subsequent questions, potentially increasing the sense of uncontrollability and stress levels.

On assessing the two methods to set the time allowed for answering each question, a 70% mean method was more appropriate than a standard deviation method, as this latter method produced timeframes for a number of questions that were too short to read the question properly let alone answer it, even after eliminating the outliers. For example, in question 39, the deadline would be 3s using the standard deviation method and 9s using the 70% of mean method (see Appendix B-3). If time pressures were too tight, the participants could have decided to randomly answer the general knowledge and mathematical questions, thus invalidating the stress stimuli. Time limits were rounded to the nearest whole number.

Distribution of Stress Stimuli

The distribution of the mathematical/general knowledge questions in the stress and the control conditions was carefully considered. Forty-eight questions of the 60 questions used in the pilot study were selected with 12 questions being removed to ensure that the exercise was not too long and to ensure an even distribution of questions across the two groups (see Appendix B-3). For the even distribution, both groups were asked to answer the same number of questions (N = 24) and the same type of questions (see Table 5). Specifically, participants would be asked to answer eight questions in each block. When possible, questions that were comparable in terms of similarity of content for the stress and control group were used. For example, questions 16 and 11 were chosen for the control group and stress group, respectively, because they both involve distance or speed vector mathematical problems.

Table 5: Even distribution of 48 mathematical/general knowledge questions across three blocks.

Block	**Question Type**	**Control Group**	**Stress Group**
Block A	Numerical Reasoning	1	1
	Tricky Questions	1	1
	Basic Arithmetic	2	2
	General Knowledge Questions	4	4
	Total Questions	**8**	**8**
Block B	Numerical Reasoning	1	1
	Tricky Questions	1	1
	Basic Arithmetic	2	2
	General Knowledge Questions	4	4
	Total Questions	**8**	**8**
Block C	Numerical Reasoning	1	1

Tricky Questions	1	1
Basic Arithmetic	2	2
General Knowledge Questions	4	4
Total Questions	**8**	**8**

This approach was repeated in three blocks: eight questions in block A, eight questions in block B, and eight questions in block C. Block A was expected to generate feelings of stress in the participants, whereas blocks B and C were expected to restore the stress level so that it would not dissipate with time. After block C, the stress levels were reported (i.e., the stress manipulation check).

Feedback

The feedback given to the participants is an important element in affecting the feelings of a social evaluative threat and the stress levels in participants (Allen et al., 2017; Dickerson & Kemeny, 2004). Incorporating negative feedback (such as receiving a message of "WRONG") has been identified as a social–evaluative threat element that leads to elevated levels of stress (Dedovic et al., 2005). This study adapted similar feedback messages to those used in previous studies that utilised the Trier Mental Challenge Test (Allendorfer et al., 2014, 2019; Dedovic et al., 2005; Kirschbaum et al., 1991).

Feedback Before the Stress Stimuli Block

Feedback or messages were shown to participants at three stages: before, during, and after each stress stimuli block (see Table 6). Before the first block of mathematical/general knowledge questions, participants in the control group were told that their performance was not being monitored. This message was expected to reduce stress, thus increasing the difference in stress levels between the stress and nonstress groups.

Table 6: Feedback given to participants before, during and after each stress stimuli block.

Feedback	**Stress Group**	**Control Group**
Before Stress Stimuli Block	“**WARNING! PLEASE READ CAREFULLY:** You are reminded that your performance is being monitored by Prolific Academic. You are reminded that there is a required minimum performance when answering the questions. Your individual performance will be compared with the rest of the participants to determine whether your data will be used in the study. Note that some questions have a limited amount of time for you to answer them, whereas others don’t. The questions with the limited amount time will have a clock showing the remaining time.” Participants click on the following message to continue the study: “I understand that my performance is being monitored.”	**“Note that your performance is NOT being evaluated, so please attempt to answer the questions as accurately as possible.”** Participants click on the following message to continue the study: “I understand that my performance is NOT being monitored.”
During Stress Stimuli Block	**WRONG!** for incorrect answer. **TIME OUT!** when the allocated time runs out.	No feedback given.

	OK for correct answers.	
After Stress Stimuli Block	If the participant scores three correct answers or lower, the following message will appear: **Unfortunately, your individual score for the math and general knowledge questions you have just completed was lower than, or did not supersede, the average performance of participants.** If the participant scores four correct answers or higher, the following message will appear: **OK, you have completed this block of math and general knowledge questions, and you can now proceed to the next step.**	No feedback given.

Conversely, participants in the stress group were warned that their performance was being monitored by multiple stakeholders—Prolific and the researchers. In addition, it was clarified that a minimum individual performance was required and that their individual performance would be compared with that of other participants for inclusion in the study using statements adapted and modified from Dedovic et al. (2005). This was expected to increase the feelings of social evaluative threat and thus stress levels of the participants. Another element of this framing was that participants may have felt the risk of losing their monetary incentives from Prolific, which would potentially increase their stress levels (even though the consent form clearly stated that the participants would be paid for their time, even if they wished to withdraw their data). These statements were not deceptive, as opposed to some previous stress research studies (such as Allendorfer et al., 2014, 2019) but were provided within the context of Prolific stating that it monitors response quality to avoid bot-like responses (Bradley, 2018).

Feedback During the Stress Stimuli Block

During the stress block, if a participant answered a question incorrectly, a "**WRONG!**" message in red would appear immediately on the screen. Conversely, a neutral "**OK**" message appeared in grey if the question were answered correctly (see Kirschbaum et al. (1991)). While some studies have utilised positive feedback, such as "CORRECT" (Dedovic et al., 2005), neutral feedback was deemed to be more suitable in this study because it was expected to result in a greater difference in stress levels between the control and stress groups. If the time allocated to the question ran out, a "**TIME OUT!**" message appeared in red (i.e., the participant left the space empty or did not select any option in multiple choice questions). For the control condition, no feedback was given.

For this exercise, the color red was used for negative messages (such as "**WRONG!**") and the original Qualtrics color of grey was used for neutral messages (such as "**OK**"). This color-coding was used for the feedback messages before, during, and after the stress stimuli blocks as the color red is associated with negative emotions, such as anger and rage (Joosten et al., 2012; Plutchik, 2001), which can be related to stress levels (Du et al., 2018; Kutchma, 2003).

Feedback After the Stress Stimuli Block

After a block of eight questions, the nonstress group received no feedback messages. However, the stress group received one of two messages, as shown in Table 6. If the participant scored three correct answers or lower in this block, then a negative message would appear explicitly comparing the individual score with those of other participants. This had the potential to further increase the social evaluative threat component of stress (Dickerson & Kemeny, 2004; Kirschbaum et al., 1991). If the participant scored four or more questions correct in this block, a neutral message would appear that had no reference to individual or group performance.

Previous studies that took place with the researchers being present have provided feedback after the stress stimuli in different ways including increasing or decreasing the difficulty of questions to enforce a result of 20–45% correct answers with participants being told that the average performance was 80–90% (Dedovic et al. 2005). Other approaches have included asking participants to write their scores on a blackboard in front of the group (Kirschbaum et al., 1991). However, because the exercise in this study took place completely online and without the presence of the researchers, a different approach was needed. It was necessary to choose a fixed cut-off average performance for each block of mathematical/general knowledge questions. Hence, the questions were distributed across the three blocks in the stress condition so that their average accuracy percentages were comparable (i.e., 35.9% to 36.9%; see Appendix B-3).

There were eight questions in each block and all the questions had the same weight, so that there can be a consistent grade level requirement in each block. Individual score can be automatically calculated via Qualtrics by adding the number of correct responses. As a result, there were nine possible individual scores for each block (Table 7) and it was possible to approximate the *average performance* for all three blocks (i.e., 35.9% to 36.9%) to 37.5%. Therefore, the average score performance within each block of eight questions was considered as 37.5% (or three correct answers).

Table 7: Possible individual scores for each stress block.

Number of Correct Responses	0	1	2	3	4	5	6	7	8
Individual Accuracy Rate	0%	12.5%	25%	37.5%	50%	62.5%	75%	87.5%	100%

References

Alekseev, A., Charness, G., & Gneezy, U. (2017). Experimental methods: When and why contextual instructions are important. *Journal of Economic Behavior & Organization*, *134*, 48–59. https://doi.org/10.1016/j.jebo.2016.12.005

Allen, A. P., Kennedy, P. J., Dockray, S., Cryan, J. F., Dinan, T. G., & Clarke, G. (2017). The Trier Social Stress Test: Principles and practice. *Neurobiology of Stress*, *6*, 113–126. https://doi.org/10.1016/j.ynstr.2016.11.001

Allendorfer, J. B., Heyse, H., Mendoza, L., Nelson, E. B., Eliassen, J. C., Storrs, J. M., & Szaflarski, J. P. (2014). Physiologic and cortical response to acute psychosocial stress in left temporal lobe epilepsy—A pilot cross-sectional fMRI study. *Epilepsy & Behavior*, *36*, 115–123. https://doi.org/10.1016/j.yebeh.2014.05.003

Allendorfer, J. B., Nenert, R., Hernando, K. A., DeWolfe, J. L., Pati, S., Thomas, A. E., Billeaud, N., Martin, R. C., & Szaflarski, J. P. (2019). FMRI response to acute psychological stress differentiates patients with psychogenic non-epileptic seizures from healthy controls – A biochemical and neuroimaging biomarker study. *NeuroImage: Clinical*, *24*, 101967. https://doi.org/10.1016/j.nicl.2019.101967

Benson III, L., & Beach, L. R. (1996). The effects of time constraints on the prechoice screening of decision options. *Organizational Behavior and Human Decision Processes*, *67*(2), 222–228. https://doi.org/10.1006/obhd.1996.0075

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*(3), 351–368. https://doi.org/10.1093/pan/mpr057

Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2014). Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *American Journal of Political Science*, *58*(3), 739–753. https://doi.org/10.1111/ajps.12081

Börger, T. (2016). Are fast responses more random? Testing the effect of response time on scale in an online choice experiment. *Environmental and Resource Economics*, *65*(2), 389–413. https://doi.org/10.1007/s10640-015-9905-1

Bradley, P. (2018). *Bots and data quality on crowdsourcing platforms*. Prolific Academic. https://blog.prolific.co/bots-and-data-quality-on-crowdsourcing-platforms

Chmielewski, M., & Kucker, S. C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science*, *11*(4), 464–473. https://doi.org/10.1177/1948550619875149

Clemmow, C., Schumann, S., Salman, N. L., & Gill, P. (2020). The base rate study: Developing base rates for risk factors and indicators for engagement in violent extremism. *Journal of Forensic Sciences*, *65*(3), 865–881. https://doi.org/10.1111/1556-4029.14282

Clifford, S., & Jerit, J. (2016). Cheating on political knowledge questions in online surveys: An assessment of the problem and solutions. *Public Opinion Quarterly*, *80*(4), 858–887. https://doi.org/10.1093/poq/nfw030

Dedovic, K., Renwick, R., Mahani, N. K., Engert, V., Lupien, S. J., & Pruessner, J. C. (2005). The Montreal Imaging Stress Task: Using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain. *Journal of Psychiatry & Neuroscience*, *30*(5), 319–325.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Du, J., Huang, J., An, Y., & Xu, W. (2018). The relationship between stress and negative emotion: The mediating role of rumination. *Clinical Research and Trials*, *4*(1). https://doi.org/10.15761/CRT.1000208 edsys. (2020). *General knowledge for kids*. https://www.edsys.in/general-knowledge-for-kids-105-questions-and-answers

Gardner, M. (1986). Entertaining mathematical puzzles. http://www.arvindguptatoys.com/arvindgupta/martingardnerone.pdf

Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high-internet-coverage population: Differences in response quality. *Public Opinion Quarterly*, *72*(5), 836–846. https://doi.org/10.1093/poq/nfn045

Jensen, C., & Thomsen, J. P. F. (2014). Self-reported cheating in web surveys on political knowledge. *Quality & Quantity*, *48*(6), 3343–3354. https://doi.org/10.1007/s11135-013-9960-z

Joosten, E., van Lankveld, G., & Spronck, P. (2012). Influencing player emotions using colors. *Journal of Intelligent Computing*, *3*(2), 76–86.

Kahneman, D. (2011). *Thinking fast and slow*. Penguin Random House UK.

Kellogg, J. S., Hopko, D. R., & Ashcraft, M. H. (1999). The effects of time pressure on arithmetic performance. *Journal of Anxiety Disorders*, *13*(6), 591–600. https://doi.org/10.1016/S0887-6185(99)00025-0

Kirschbaum, C., Diedrich, O., Gehrke, J., Wüst, S., & Hellhammer, D. H. (1991). Cortisol and behavior: The "Trier Mental Challenge Test" (TMCT)—First evaluation of a new psychological stress test. In A. Ehlers, W. Fiegenbaum, I. Florin, & J. Margraf (Eds.), *Perspectives and Promises of Clinical Psychology* (pp. 67–78). Springer.

Kutchma, T. M. (2003). The effects of room color on stress perception: Red versus green environments, *3*(12). *Journal of Undergraduate Research at Minnesota State University, Mankato*. https://cornerstone.lib.mnsu.edu/jur/vol3/iss1/3/#:~:text=It%20was%20found%20that%20subjects,significant%20role%20in%20stress%20perception.

McLeod, I., Zhang, Y., & Yu, H. (2003). Multiple-choice randomization. *Journal of Statistics Education*, *11*(1), 4. https://doi.org/10.1080/10691898.2003.11910695

Motta, M. P., Callaghan, T. H., & Smith, B. (2017). Looking for answers: Identifying search behavior and improving knowledge-based data quality in online surveys. *International Journal of Public Opinion Research*, *29*(4), 575–603. https://doi.org/10.1093/ijpor/edw027

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872. https://doi.org/10.1016/j.jesp.2009.03.009

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, *46*(4), 1023–1031. https://doi.org/10.3758/s13428-013-0434-y

Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, *89*(4), 344–350. https://www.jstor.org/stable/27857503

Prior, M. (2014). Visual political knowledge: A different road to competence? *The Journal of Politics*, *76*(1), 41–57. https://doi.org/10.1017/S0022381613001096

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. https://doi.org/10.1037/0033-2909.114.3.510

Schmidt, G. B. (2015). Fifty days an MTurk worker: The social and motivational context for Amazon Mechanical Turk workers. *Industrial and Organizational Psychology*, *8*(2), 165–171. https://doi.org/10.1017/iop.2015.20

Smith, H. (2017). *How to pass numerical reasoning tests* (3rd ed.). Kogan Page Limited.

Tsiga, E., Panagopoulou, E., Sevdalis, N., Montgomery, A., & Benos, A. (2013). The influence of time pressure on adherence to guidelines in primary care: An experimental study. *BMJ Open*, *3*(4), e002700. https://doi.org/10.1136/bmjopen-2013-002700

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record*, *58*(3), 475–482. https://doi.org/10.1007/BF03395630

Yip, D. Y. (2004). Questioning skills for conceptual change in science instruction. *Journal of Biological Education*, *38*(2), 76–83. https://doi.org/10.1080/00219266.2004.9655905

Appendix B-2: Mean accuracy rate and response times of the stress stimuli

B-2-A: Tricky mathematical questions (numeric text-entry). Half the questions with the lower accuracy rates than the other half are in bold.

Question	Answer	*N*	Accuracy (%)	Mean (secs)	*SD* (secs)
1. A pen and pad cost one dollar and ten cents. The pen costs one dollar more than the pad. How much does the pad cost (in cents)? (*please type the number only*).	**5 cents**	**27**	**33.3**	**41.238**	**34.969**
2. There is a patch of flowers in a garden. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire garden, how many days would it take for the patch to cover half of the garden? (*please type the number only*).	47 days	29	44.8	33.788	22.227
3. What is the answer? $\mathbf{6 \div 2(1+2) =}$	**9**	**29**	**27.6**	**23.951**	**12.500**
4. In a juice factory, it takes 5 machines 5 seconds to produce 5 boxes of juice. How many seconds would it take 100 machines to make 100 boxes of juice? (*please type the number only*).	5 seconds	27	55.6	34.105	23.791
5. A person was born on May 6, 30 B.C. He died on May 6, 30 A.D. How old was he when he died? (*please type the number only*).	**59 years**	**29**	**0.0**	**32.839**	**24.286**
6. Divide 30 by ½ and add 10. What is the result?	**70**	**28**	**28.6**	**13.604**	**6.406**
7. A boy had five bananas and ate all but three. How many bananas were left? (*please type the number only*)	3 bananas	28	78.6	15.034	6.598
8. What two whole numbers—not fractions—make the unlucky number 13 when multiplied together? (*you may*	1 x 13	27	74.0	49.818	24.908

type the two numbers in any order as 'number 1 x number 2').					
9. If there are 12 one-cent stamps in a dozen, then how many two-cent stamps are in a dozen? (*please type the number only*).	12 stamps	29	48.3	21.022	8.716
10. If a clock takes five seconds to strike 6 o'clock, how long will it take to strike 12 o'clock? (*please type the number only*).	**11 seconds**	**27**	**0.0**	**30.491**	**17.226**

B-2-B: Numerical reasoning questions (numeric text-entry). Half the questions with the lower accuracy rates than the other half are in bold.

Question	Answer	*N*	Accuracy (%)	Mean (secs)	*SD* (secs)
11. Ruth goes walking to work at 8.10am. She stops to buy a tea and read her newspaper for 15 minutes and arrives at work at 9.55am. The distance between her house and work is 6 miles. What is her average walking speed in miles per hour? (*please type the number only*).	**4 miles per hour**	**28**	**50.0**	**79.835**	**37.960**
12. Arsenal football club 'games won' to 'games lost' record last season was 2 : 3. How many games did they play last season if all the games were either won or lost and Arsenal won 6 games? (*please type the number only*).	**15 games**	**26**	**46.2**	**58.236**	**28.477**
13. The total entrance price for a theatre show for 2 adults and 2 children is $24. The ticket price for an adult is twice the price for a child's ticket. How much does an adult's ticket cost? (*please type the number only*).	$8	28	75.0	61.711	38.434
14. While on vacation in Italy, William withdraws €200 from his bank account and receives a	**8 notes**	**26**	**65.4**	**91.728**	**51.193**

pile of €10 and €20 notes. How many €10 notes does William receive if he receives 14 notes in total? (*please type the number only*).					
15. A plant grows by 5% each year. Its height was 90 cm when it was planted. In which year will the plant exceed 1 m in height? (*please type the number only*).	**Year 3**	**28**	**39.3**	**45.878**	**24.254**
16. A car travels along a road at a rate of 40 mph for 4.5 hours; how far does the car travel in miles? (*please type the number only*).	180 miles	28	75.0	36.905	20.130
17. The price of a barrel of oil increased from £20 to £24 between May and August 1993. By what percentage did the price of oil increase during this period? (*please type the number only*).	20%	27	66.7	41.406	29.452
18. A rooster and a hen together weigh 27 pounds. If the rooster weighs twice as much as the hen does, how much does the hen weighs in pounds? (*please type the number only*).	9 pounds	27	77.8	42.773	21.990

B-2-C: Basic arithmetic (numeric text-entry). Half the questions with the lower accuracy rates than the other half are in bold.

Question	**Answer**	***N***	**Accuracy (%)**	**Mean (secs)**	***SD* (secs)**
19. What is the answer? $3 \times 4 - 9 =$	3	28	100.0	8.645	3.238
20. What is the answer? $\mathbf{2 - 6 \div 3 =}$	**0**	**26**	**88.5**	**10.585**	**5.585**
21. What is the answer? $\mathbf{8 \div 4 + 2 \times 5 =}$	**12**	**29**	**65.5**	**16.324**	**6.890**
22. What is the answer? $93 - 16 =$	77	26	96.2	14.621	6.796
23. What is the answer? $84 - 19 - 7 =$	58	28	96.4	22.104	8.713
24. What is the answer? $\mathbf{3 \times 3 - 2 \times 2 =}$	**5**	**29**	**58.6**	**14.997**	**11.297**
25. What is the answer?	67	26	96.2	9.197	3.139

$75 - 8 =$					
26. What is the answer? $\mathbf{20 - 5 + 2 \times 3 =}$	**21**	**28**	**57.1**	**19.347**	**9.466**
27. What is the answer? $5 \times 2^2 - 3 =$	17	27	88.9	11.754	6.028
28. What is the answer? $\mathbf{4 \times 3^2 \div 1^4 + 4 =}$	**40**	**26**	**61.5**	**27.165**	**13.569**
29. What is the answer? $\mathbf{3 - 3 \times 3 \div 3^2 =}$	**2**	**27**	**44.4**	**24.921**	**19.548**
30. What is the answer? $12 \div 6 \times 3 \div 2 =$	3	28	89.3	22.511	16.945

B-2-D: General knowledge questions (multiple choice). Half the questions with the lower accuracy rates than the other half are in bold.

Question	Answer	*N*	Accuracy (%)	Mean (secs)	*SD* (secs)
31. In the following sentence, there are two missing words. Which two words from the options (A–E) below fit the missing words best: The higher court's reversal of its previous ruling on the issue of suspected terrorists __________________ its reputation for____________. A. sustained, infallibility B. compromised, consistency C. bolstered, doggedness D. aggravated, inflexibility E. dispelled, vacillation	B	26	76.9	47.599	19.616
32. Which ocean goes to the deepest depths? A. Pacific B. Arctic C. Atlantic D. Indian E. Southern	A	27	59.3	10.212	6.118
33. What is the meaning of the musical term 'allegro'? **A. loud** **B. soft** **C. quick** **D. slow** **E. stop**	**C**	**28**	**57.1**	**13.062**	**6.789**
34. If $x/3 = x^2$, then value of x can be which of the following? I. -1/3; II. 0; III. 1/3 **A. I only** **B. II only** **C. III only** **D. II and III only** **E. I, II, and III**	**D**	**28**	**25.0**	**45.234**	**30.067**

35. If the population of a town doubles every 10 years, then population in the year (X + 100) years will be how many times the population in the year (X)? **A. 512** **B. 100** **C. 1,024** **D. 10** **E. 1,000**	**C**	**29**	**24.1**	**65.090**	**37.891**
36. Which one of the following categories is NOT awarded a Nobel Prize? **A. Physics** **B. Chemistry** **C. Biology** **D. Medicine** **E. Literature**	**C**	**28**	**39.3**	**20.515**	**14.047**
37. What percentage cocoa solids must chocolate contain to be legally called chocolate? **A. 99** **B. 50** **C. 15** **D. 25** **E. 35**	**E**	**29**	**20.7**	**15.811**	**10.073**
38. What is the major vitamin found in brown rice? **A. A** **B. B** **C. C** **D. D** **E. E**	**B**	**29**	**34.5**	**16.835**	**11.059**
39. Samite is a type of? **A. Cake** **B. Stone** **C. Dog** **D. Fabric** **E. Horse**	**D**	**29**	**17.2**	**12.369**	**9.612**

40. Who wrote 'don't count your chickens before they are hatched'? **A. Shakespeare** **B. Ben Franklin** **C. Chaucer** **D. Aesop** **E. Dickens**	**D**	**27**	**22.2**	**10.356**	**5.499**
41. What is the capital of Cambodia? **A. Luang Prabang** **B. Vientiane** **C. Ho Chi Minh** **D. Phnom Penh** **E. Hebei**	**D**	**28**	**39.3**	**13.525**	**5.830**
42. What is a shooting star? **A. Dying star** **B. Meteor** **C. Comet** **D. Asteroid** **E. Supernova**	**B**	**26**	**34.6**	**12.308**	**4.548**
43. What percent of people live north of the equator? **A. 70%** **B. 75%** **C. 80%** **D. 85%** **E. 90%**	**E**	**28**	**3.6**	**13.472**	**5.060**
44. Which one of these is not an insect? **A. Flea** **B. Tick** **C. Mosquito** **D. Beetle** **E. Butterfly**	**B**	**28**	**21.4**	**19.246**	**9.324**
45. What will you get if you shake whipping cream in a glass can for 10 minutes? **A. Whipped Cream** **B. Cheese** **C. Milk** **D. Butter** **E. Yogurt**	**D**	**28**	**21.4**	**20.748**	**11.353**

46. What animal first reached Earth's orbit alive? A. Mice B. Ape C. Cockroach D. Cat E. Dog	E	27	63.0	9.343	3.292
47. What is the official World Health Organization abbreviation for the current global pandemic? A. COVID-19 B. Coronavirus C. SARS-CoV-2 D. 2019-nCoV E. SARS-CoV	A	29	65.5	14.810	7.558
48. How many days are in a year? A. 151 B. 243 C. 365 D. 411 E. 502	C	29	100.0	4.835	1.769
49. Which is the longest river on earth? **A. Nile** **B. Amazon** **C. Congo** **D. Lena** **E. Mekong**	**A**	**26**	**50.0**	**7.829**	**4.337**
50. Which is the tallest mountain in the world? A. Fuji B. Mount Kilimanjaro C. Table Mountain D. Mont Blanc E. Mount Everest	E	29	86.2	8.520	4.056
51. Where is Kenya located? A. Asia B. Africa C. Europe D. South America E. Australia	B	29	100.0	5.175	2.144

52. Which planet is known as the Red Planet? A. Neptune B. Jupiter C. Mars D. Mercury E. Earth	C	26	96.2	5.278	2.551
53. Which animal is known as the 'Ship of the Desert?' A. Fox B. Camel C. Lizard D. Whale E. Scorpion	B	29	100.0	10.090	4.935
54. Who was the first man to walk on the moon? A. Cristiano Ronaldo B. Captain Cook C. Neil Armstrong D. Ibn Battuta E. Laika	C	28	100.0	6.316	2.318
55. Which is the most spoken language in the world? **A. English** **B. Spanish** **C. Arabic** **D. French** **E. Chinese**	**E**	**26**	**34.6**	**6.497**	**2.640**
56. What is the capital of the United States? A. Washington, DC B. New York City C. Los Angeles D. Chicago E. Miami	A	29	100.0	5.923	4.357
57. Who is the founder of Microsoft? A. Donald Trump B. Bill Gates C. Mark Zuckerberg D. Steve Jobs E. Elon Musk	B	27	96.3	4.939	2.280

58. Global warming is caused by the excess of which type of gas? A. Oxygen B. Nitrogen C. Argon D. Carbon dioxide E. Carbon monoxide	D	28	78.6	12.225	6.414
59. Which country is home to the kangaroo? A. Australia B. Congo C. Angola D. Zambia E. New Zealand	A	29	100.0	5.682	2.652
60. Who invented the telephone? A. Thomas Edison B. Wright Brothers C. Guglielmo Marconi D. Leonardo da Vinci E. Alexander Graham Bell	E	29	75.9	12.117	11.289

Appendix B-3: Order of the stress stimuli within the blocks

Specific mathematical/general knowledge questions for each of the three blocks are provided below; accuracy percentages of each question from the baseline study and the average percentage accuracy for each block are included. The precalculated deadlines of two methods are shown, namely the mean minus one standard deviation (i.e., *M*-*SD*) and 70% of the mean (i.e., 70% of *M*).

Block	**Question Type**	**Control Group**	**Stress Group**			
		Selected question	**Selected question**	**Baseline accuracy (%)**	**Deadline (secs)**	
					M-SD	**70% of *M***
A	Numerical Reasoning	Q16. A car travels along a road at a rate of 40 mph for 4.5 hours. How far does the car travel in miles? (*please type the number only*).	Q11. Ruth goes walking to work at 8.10 a.m. She stops to buy a tea and read her newspaper for 15 minutes and arrives at work at 9.55 a.m. The distance between her house and work is 6 miles. What is her average walking speed in miles per hour? (*please type the number only*).	50.0	42	56
	Tricky Questions	Q7. A boy had five bananas and ate all but three. How many bananas were left? (*please type the number only*)	Q1. A pen and pad cost one dollar and ten cents. The pen costs one dollar more than the pad. How much does the pad cost (in	33.3	6	29

		cents)? (*please type the number only*).			
Basic Arithmetic	Q25. What is the answer? $75 - 8 =$	Q21. What is the answer? $8 \div 4 + 2 \times 5 =$	65.5	9	11
	Q27. What is the answer? $5 \times 2^2 - 3 =$	Q29. What is the answer? $3 - 3 \times 3 \div 3^2 =$	44.4	5	17
General Knowledge Questions	Q56. What is the capital of the United States? A. Washington, DC B. New York City C. Los Angeles D. Chicago E. Miami	Q34. If $x/3 = x^2$, then value of x can be which of the following? I. -1/3; II. 0; III. 1/3 A. I only B. II only C. III only D. II and III only E. I, II, and III	25.0	15	32
	Q60. Who invented the telephone? A. Thomas Edison B. Wright Brothers C. Guglielmo Marconi D. Leonardo da Vinci E. Alexander Graham Bell	Q36. Which one of the following categories is NOT awarded a Nobel Prize? F. Physics G. Chemistry H. Biology I. Medicine J. Literature	39.3	6	14
	Q51. Where is Kenya located?	Q37. What percentage cocoa	20.7	6	11

		A. Asia B. Africa C. Europe D. South America E. Australia	solids must chocolate contain to be legally called chocolate? F. 99 G. 50 H. 15 I. 25 J. 35			
		Q59. Which country is home to the kangaroo? A. Australia B. Congo C. Angola D. Zambia E. New Zealand	Q39. Samite is a type of? A. Cake B. Stone C. Dog D. Fabric E. Horse	17.2	3	9
			Average Score for Stress Group	**36.9**	**-**	**-**
B	Numerical Reasoning	Q18. A rooster and a hen together weigh 27 pounds. If the rooster weighs twice as much as the hen does, how much does the hen weighs in pounds? (*please type the number only*).	Q12. The Arsenal football club “games won” to “games lost” record last season was 2:3. How many games did they play last season if all the games were either won or lost and Arsenal won 6 games? (*please type the number only*).	46.2	30	41
	Tricky Questions	Q4. In a juice factory, it takes 5 machines 5 seconds to produce 5 boxes of juice.	Q6. Divide 30 by ½ and add 10. What is the result?	28.6	7	10

	How many seconds would it take 100 machines to make 100 boxes of juice? (*please type the number only*).				
Basic Arithmetic	Q22. What is the answer? $93 - 16 =$	Q20. What is the answer? $2 - 6 \div 3 =$	88.5	5	7
	Q30. What is the answer? $12 \div 6 \times 3 \div 2 =$	Q28. What is the answer? $4 \times 3^2 \div 1^4 + 4 =$	61.5	14	19
General Knowledge Questions	Q31. In the following sentence, there are two missing words; which two words from the options (A–E) below fit the missing words best: The higher court's reversal of its previous ruling on the issue of suspected terrorists ______________ its reputation for_____________. F. sustained, infallibility G. compromised, consistency	Q43. What percent of people live north of the equator? F. 70% G. 75% H. 80% I. 85% J. 90%	3.6	8	9

	H. bolstered, doggedness I. aggravated, inflexibility J. dispelled, vacillation				
	Q57. Who is the founder of Microsoft? A. Donald Trump B. Bill Gates C. Mark Zuckerberg D. Steve Jobs E. Elon Musk	Q40. Who wrote “don’t count your chickens before they are hatched”? F. Shakespeare G. Ben Franklin H. Chaucer I. Aesop J. Dickens	22.2	5	7
	Q53. Which animal is known as the “Ship of the Desert?” A. Fox B. Camel C. Lizard D. Whale E. Scorpion	Q44. Which one of these is not an insect? F. Flea G. Tick H. Mosquito I. Beetle J. Butterfly	21.4	10	13
	Q54. Who was the first man to walk on the moon? A. Cristiano Ronaldo B. Captain Cook C. Neil Armstrong D. Ibn Battuta E. Laika	Q45. What will you get if you shake whipping cream in a glass can for 10 minutes? F. Whipped cream G. Cheese H. Milk I. Butter J. Yogurt	21.4	9	15

			Average Score for Stress Group	**36.7**	**-**	**-**
C	Numerical Reasoning	Q13. The total entrance price for a theatre show for 2 adults and 2 children is \$24. The ticket price for an adult is twice the price for a child's ticket. How much does an adult's ticket cost? (*please type the number only*).	Q15. A plant grows by 5% each year. Its height was 90 cm when it was planted. In which year will the plant exceed 1 m in height. (*please type the number only*).	39.3	22	32
	Tricky Questions	Q9. If there are 12 one-cent stamps in a dozen, then how many two-cent stamps are in a dozen? (*please type the number only*).	Q5. A person was born on May 6, 30 B.C. He died on May 6, 30 A.D. How old was he when he died? (*please type the number only*).	0.0	9	23
	Basic Arithmetic	Q19. What is the answer? $3 \times 4 - 9 =$	Q26. What is the answer? $20 - 5 + 2 \times 3 =$	57.1	10	14
		Q23. What is the answer? $84 - 19 - 7 =$	Q24. What is the answer? $3 \times 3 - 2 \times 2 =$	58.6	4	10
	General Knowledge Questions	Q48. How many days are in a year? A. 151 B. 243 C. 365 D. 411 E. 502	Q35. If the population of a town doubles every 10 years, then population in the year (X + 100) years will be how many times	24.1	27	46

		the population in the year (X)? F. 512 G. 100 H. 1,024 I. 10 J. 1,000			
	Q58. Global warming is caused by the excess of which type of gas? A. Oxygen B. Nitrogen C. Argon D. Carbon dioxide E. Carbon monoxide	Q38. What is the major vitamin found in brown rice? A. A B. B C. C D. D E. E	34.5	6	12
	Q52. Which planet is known as the Red Planet? A. Neptune B. Jupiter C. Mars D. Mercury E. Earth	Q42. What is a shooting star? F. Dying star G. Meteor H. Comet I. Asteroid J. Supernova	34.6	8	9
	Q50. Which is the tallest mountain in the world? A. Fuji B. Mount Kilimanjaro C. Table Mountain D. Mont Blanc E. Mount Everest	Q41. What is the capital of Cambodia? A. Luang Prabang B. Vientiane C. Ho Chi Minh D. Phnom Penh E. Hebei	39.3	8	9
		Average Score for Stress Group	**35.9**	**-**	**-**

Appendix B-4: Self-evaluation questionnaire (PART A)

A number of statements that people have used to describe themselves are given below. Read each statement and then choose the response to indicate how you feel *right now*, that is, *at the moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer that seems to describe your feelings best. Note that item 13 is an additional item was added to the questionnaire as an attention check screener.

		Not at all	Somewhat	Moderately so	Very much so
1	I feel calm	1	2	3	4
2	I feel secure	1	2	3	4
3	I am tense	1	2	3	4
4	I feel strained	1	2	3	4
5	I feel at ease	1	2	3	4
6	I feel upset	1	2	3	4
7	I am presently worrying over possible misfortunes	1	2	3	4
8	I feel satisfied	1	2	3	4
9	I feel frightened	1	2	3	4
10	I feel comfortable	1	2	3	4
11	I feel self-confident	1	2	3	4
12	I feel nervous	1	2	3	4
13	Please tick "somewhat"	1	2	3	4
14	I am jittery				
15	I feel indecisive	1	2	3	4
16	I am relaxed	1	2	3	4
17	I feel content	1	2	3	4
18	I am worried	1	2	3	4
19	I feel confused	1	2	3	4
20	I feel steady	1	2	3	4
21	I feel pleasant	1	2	3	4

Appendix B-5: Self-evaluation questionnaire (PART B)

A number of statements which people have used to describe themselves are given below. Read each statement and then choose the response to indicate how you *generally feel*. Note that PART B of the questionnaire is different from PART A that you completed earlier. In part A, it was about how you feel right now, but here, in part B, it is how you generally feel as a person, NOT how you specifically feel at this moment.

		Almost Never	Sometimes	Often	Almost Always
1	I feel pleasant	1	2	3	4
2	I feel nervous and restless	1	2	3	4
3	I feel satisfied with myself	1	2	3	4
4	I wish I could be as happy as others seem to be	1	2	3	4
5	I feel like a failure	1	2	3	4
6	I feel rested	1	2	3	4
7	I feel "calm, cool, and collected"	1	2	3	4
8	I feel that difficulties are piling up so that I cannot overcome them	1	2	3	4
9	I worry too much over something that really doesn't matter	1	2	3	4
10	I am happy	1	2	3	4
11	I have disturbing thoughts	1	2	3	4
12	I lack self-confidence	1	2	3	4
13	Please tick "often"	1	2	3	4
14	I feel secure				
15	I make decisions easily	1	2	3	4
16	I feel inadequate	1	2	3	4
17	I am content	1	2	3	4
18	Some unimportant thought runs through my mind and bothers me	1	2	3	4

19	I take disappointments so keenly that I can't put them out of my mind	1	2	3	4
20	I am a steady person	1	2	3	4
21	I get in a state of tension or turmoil as I think over my recent concerns and interests	1	2	3	4

Appendix C: Supplementary Materials for Chapter 6

Appendix C-1: Assessment by nine fingerprint examiners of 23 pairs of fingermarks/prints for matching difficulty.

Fingerprint Pairs			**Fingerprint Experts**									
Pair	Ground Truth	ID	A	B	C	D	E	F	G	H	I	Avg
		YoE	13	10	6.5	11	8	3	34	28	11	13.8
1	match	Difficulty rating (1 = easy, 5 = difficult)	2	1	1	1	1	2	1	2	1	1.3
2	match		3	2	3	2	2	2	1	2	2	2.1
3	non-match		5	2	4	5	3	5	4	5	5	4.2
4	match		4	3	3	4	4	4	1	3	3	3.2
5	match		5	5	3	5	3	5	3	4	4	4.1
6	non-match		1	1	1	1	1	1	1	1	1	1.0
7	match		3	1	2	3	2	4	1	1	2	2.1
8	non-match		4	2	2	2	2	4	1	3	3	2.6
9	non-match		2	2	2	2	3	3	1	1	1	1.9
10	non-match		5	4	3	4	4	5	3	4	4	4.0

11	non-match	3	1	2	3	2	3	1	2	1	2.0
12	match	5	5	4	4	4	5	4	5	5	4.6
13	match	2	1	1	2	1	1	1	1	1	1.2
14	non-match	4	1	2	3	2	1	1	2	2	2.0
15	non-match	4	1	2	3	2	4	1	1	2	2.2
16	match	4	1	2	3	2	2	1	2	2	2.1
17	non-match	3	1	3	3	2	2	1	2	3	2.2
18	match	5	1	3	3	4	3	1	3	3	2.9
19	match	3	2	2	3	1	3	1	2	1	2.0
20	non-match	4	4	4	5	5	5	1	3	3	3.8
21	match	2	1	2	2	2	1	1	1	1	1.4
22	match	4	1	2	3	2	3	1	2	3	2.3
23	non-match	4	1	3	3	1	2	1	2	2	2.1

Key: YoE = Years of experience; ID = Identification letter of participant.

Appendix C-2: Example of fingerprint mark/print assessed for matching difficulty by fingerprint examiners.

Example
مثال

Easy (1) سهل	2	3	4	Difficult(5) صعب
	✓			

Appendix C-3: Fingerprint decision-making of novices and experts.

C-3-A: Novices

Pair	ID-C	ID-S	Ex-C	Ex-S	Inc-C	Inc-S	Total-C	Total-S
				Same-Source Evidence				
A0	8	7	5	8	16	10	29	25
B0	19	22	7	3	5	2	31	27
C0	0	1	13	7	20	19	33	27
D0	10	12	17	11	4	4	31	27
E0	15	13	16	13	0	2	31	28
F0	20	17	6	5	5	3	31	25
	38.7%	45.3%	34.4%	29.6%	26.9%	25.2%	*N* = 186	*N* = 159
				Different-Source Evidence				
A1	3	3	18	19	11	7	32	29
B1	21	16	3	5	6	6	30	27
C1	4	1	16	8	8	18	28	27
D1	6	5	18	15	6	7	30	27
E1	1	1	25	20	4	5	30	26
F1	16	10	11	11	3	8	30	29
	28.3%	21.8%	50.6%	47.3%	21.1%	30.9%	*N* = 180	*N* = 165

C-3-B: Experts

Pair	ID-C	ID-S	Ex-C	Ex-S	Inc-C	Inc-S	Total-C	Total-S
				Same Source Evidence				
A0	1	0	2	0	6	8	9	8
B0	1	1	5	5	2	1	8	7
C0	1	1	1	0	5	8	7	9
D0	5	9	0	0	1	0	6	9
E0	9	10	2	1	0	0	11	11
F0	6	7	0	0	2	0	8	7
	46.9%	54.9%	20.4%	11.8%	32.7%	33.3%	*N* = 49	*N* = 51
				Different Source Evidence				
A1	0	0	6	9	1	1	7	10
B1	1	0	6	8	1	3	8	11
C1	0	0	0	0	9	9	9	9
D1	0	0	10	9	0	0	10	9
E1	0	0	5	7	0	0	5	7
F1	0	0	8	11	0	0	8	11
	2.1%	0.0%	74.5%	77.2%	23.4%	22.8%	*N* = 47	*N* = 57

Key: ID = Identification decision; Ex = Exclusion decision; Inc = Inconclusive decision; C = Control condition; S = Stress condition; Total = all fingerprint pairs assessed in each condition; A0 = same-source pair of latent print A; A1 = different-source pair of latent print A.