Tom Pakkanen

Behavioural Crime Linking in Serial Homicide

Towards Practical Application

Åbo Akademi University

Behavioural Crime Linking in Serial Homicide

Towards Practical Application

Tom Pakkanen

Psychology
Faculty of Arts, Psychology, and Theology
Åbo Akademi University
Åbo, Finland, 2021

Supervised by

Professor Pekka Santtila, PhD
NYU-ECNU Institute for Social Development
New York University Shanghai
China

Professor Patrick Jern, PhD
Department of Psychology
Åbo Akademi University
Finland

Reviewed by

Professor Samantha Lundrigan, PhD
Department of Humanities and Social Sciences
Anglia Ruskin University
United Kingdom

Associate Professor Alasdair Goodwill, PhD
Department of Psychology
Ryerson University
Canada

Opponent

Professor Samantha Lundrigan, PhD
Department of Humanities and Social Sciences
Anglia Ruskin University
United Kingdom

Cover art by Henri Salonen

ISBN 978-952-12-4108-6 (printed)
ISBN 978-952-12-4109-3 (digital)
Painosalama, Åbo, Finland 2021

To my mother and my father
who taught me that curiosity is a virtue

Acknowledgements

The work encapsulated by these covers is the product of many wonderful collaborations, inspiring people, and loving support over the years. I am grateful for the academic home and financial support that Åbo Akademi University has provided for me. I also wish to thank the participants of Study II for their valuable contribution to this work, and the Italian police for providing us with the data that made these studies possible.

I am enormously grateful to my supervisor, the brilliant Professor Pekka Santtila, who has been so generous with his wisdom, insight, patience, and support over the better part of two decades. This thesis would simply not have come about without him. Professor Santtila's guidance, and belief in me when my own faltered, has made all the difference. Also, beyond this work, Professor Santtila's passion for scientific rigour has been one of my primary professional inspirations. I am also grateful to my other supervisor Professor Patrick Jern, for committing to my work, sharing his knowledge, and supporting me through the wrap-up of the thesis.

I wish to thank my co-authors, especially the Italian connection: Angelo Zappàla, Dario Bosco, and Andrea Berti. Their data, insightful comments, and support made up the foundation for this work. I am also indebted to Jukka Sirén, who with his exceptional talent for statistics, made it all that much more understandable to me through his timely and generous explanations and comments. Thank you also Sara Landström and Isa Skeppar for their help with the Swedish summary.

The external reviewers, Professor Samantha Lundrigan and Associate Professor Alasdair Goodwill, improved this work through their insightful comments and encouraging observations. I am especially grateful to Professor Lundrigan for agreeing to appear in the role of opponent at the public defence of this thesis.

A special thank you goes out to Matthew Tonkin, Jessica Woodhams, Amy Burrell, and the rest of the C-LINK crew. I feel privileged for having been included in the group, for the opportunity to learn from all the talented researchers and practitioners, and for the friendships that have formed during the collaboration.

My colleagues and dear friends Taina, Liisa, and Julia. I lack the words to describe what our endless discussions and your support has meant to me. Thank you for challenging me in a way that has allowed me to grow, both professionally and personally.

Finally, I wish to express my gratitude and love for my family and friends. Many of you I have known all my life. Some since April. Thank you for standing by me and showing me that you care. Even when you have been asking how the thesis is progressing. You give meaning to my life outside of my study. Thank you to my brother and his family, for so generously hosting me this summer, allowing me to concentrate on finishing the long journey that has been this thesis. And a special thank you to my nieces Ninni and Noomi, for giving me a lovely reason to leave the screen and rejuvenate in the best of company.

Helsinki, October 2021

Tom Pakkanen

List of Original Publications

I. Santtila, P., Pakkanen, T., Zappàla, A., Bosco, D., Valkama, M., & Mokros, A. (2008). Behavioural crime linking in serial homicide. *Psychology, Crime & Law, 14*, 245–265.

II. Pakkanen, T., Zappàla, A., Grönroos, C., & Santtila, P. (2012). The effects of coding bias on estimates of behavioural similarity in crime linking research of homicides. *Journal of Investigative Psychology and Offender Profiling, 9*, 223–234.

III. Pakkanen, T., Zappàla, A., Bosco, D., Berti, A., & Santtila, P. (2015). Can hard-to-solve one-off homicides be distinguished from serial homicides? Differences in offence behaviours and victim characteristics. *Journal of Criminal Psychology, 5*, 216–232.

IV. Pakkanen, T., Sirén, J., Zappàla, A., Jern, P., Bosco, D., Berti, A., & Santtila, P. (2020). Linking serial homicide – towards an ecologically valid application. *Journal of Criminological Research, Policy and Practice, 7*, 18–33.

Sammanfattning

Brottslänkning syftar till att länka samman två eller flera, för brottsutredaren observerbara brott, på basis av gärningspersonens brottsplatsbeteende. Metoden bygger på antagandet att en gärningspersons beteende är konsekvent från ett brott till ett annat och unikt från en gärningsperson till en annan. Systematisk brottslänkning på basis av brottsplatsbeteende introducerades under 1970-talet och sedan 1990-talet har området utvecklats, såväl metodologiskt som vetenskapligt. Under det senaste decenniet har det skett en kraftig ökning av vetenskapliga studier i brottslänkning. Forskningsfältet har utvecklats avsevärt, från empirisk validering av de underliggande antagandena, till kartläggning av praxis och mer ekologiskt valida tester av brottslänkningens precision. Utvecklingen är positiv eftersom en reliabel och valid brottslänkning, i bästa fall, kan hjälpa polisen prioritera utredningslinjer och lösa svåra och resurskrävande seriemord.

Forskningen kring brottslänkning vid seriemord har i huvudsak utförts i USA och fokus har legat på hur konsekvent och unikt seriemördare beter sig. En brist i denna forskning är att den saknar empiriska tester av så kallade brottslänkningsmodeller och hur väl dessa modeller identifierar mordserier, genom korrekt sammanlänkning av en gärningspersons mordfall. En annan kunskapslucka i forskningsfältet är att forskare till största delen använt länkning utgående från seriebrott, medan verkliga brottsdatabaser innehåller såväl seriebrott som enskilda brott. Det finns ett fåtal vetenskapliga studier där man har testat vilken inverkan enskilda inbrott och våldtäkter har på precisionen i sammanlänkningen av inbrotts-, och våldtäktsserier. Vilken inverkan enskilda mord har på precisionen i sammanlänkningen av seriemord, är dock ännu oklart. Vidare, finns det en del studier där man har jämfört seriemord med enskilda mord, men det saknas forskning kring möjligheten att förutsäga om ett mord hör till en serie eller om det är ett enskilt fall. Tidigare forskning har visat att kognitiva bias, så som konfirmeringsbias eller förväntanseffekt kan ha betydande inverkan på brottsutredningar. Trots detta, saknas det forskning kring hur dessa bias påverkar brottslänkning.

Syftet med föreliggande avhandling var att utvidga forskningen kring brottslänkning av seriemord. För att maximera resultatens relevans för brottsutredning, genomfördes en series brottslänkningsstudier med hög ekologisk validitet, med avseende på de brott som inkluderades och de frågeställningar som analyserades. Det huvudsakliga underlaget för avhandlingen bestod av 116 italienska seriemord utförda i 23 separata serier

samt på 45 svårlösta mord som inte ingår i någon serie. **Studie I** identifierade sju dimensioner av beteende bland seriemorden (t.ex., mord med sexuella motiv och aspekter av kontrollbeteende), som överensstämde med tidigare forskning. Anmärkningsvärt var att även andra motiv än sexuella, fanns bland seriemorden. De flesta brotten (63%) kunde sammanlänkas till rätt serie i den prediktiva delen av studien. **Studie II** var ett experiment som undersökte huruvida förhandskunskap om vilken serie enskilda brott tillhörde, ökade deltagarnas uppfattning om likhet i brottsplatsbeteende. Studien visade ingen förväntanseffekt, men detta kan ha berott på att uppgiften var för enkel, vilket innebär att studien behöver replikeras, innan meningsfulla slutsatser kan dras. **Studie III** identifierade flera signifikanta skillnader mellan serie- och enskilda mord. Vidare kunde dessa skillnader användas till att, med god precision, förutspå huruvida brottet var del av en serie eller ej. **Studie IV** kombinerade alla befintliga metodologiska framsteg dittills och visade att brottslänkning fortfarande bibehöll god precision även när en stor andel (10:1) enskilda fall inkluderades i urvalet. Modellens specificitet försämrades (flera falska positiva), som en funktion av inkluderade enskilda fall. Då varje enskilt fall jämfördes mot en rankad lista, från beteendemässigt likartade fall till mindre likartade, kunde en liknande försämring iakttas i andelen brott, som låg nära toppen. Modellens förmåga att sammanlänka fall förblev dock god, vilket vidare gav stöd för utförandet av brottsläkning med data av hög ekologisk validitet.

Studierna i den föreliggande avhandlingen bidrar till en förbättring av forskningsmetodologin gällande brottslänkning av seriemord på basis av brottsplatsbeteende. För att maximera den praktiska användbarheten av forskningsmetodologin och de beprövande modellerna krävs dock ytterligare replikering på databaser i olika länder och olika jurisdiktioner. För att hålla forskningen relevant för den kliniska verksamheten och för att utveckla evidensbaserad praxis, krävs ett nära samarbete mellan forskare och kliniker. I och med att vi får en tydligare bild av precisionen och felfrekvensen för modeller för brottslänkning, ökar deras användbarhet i såväl förundersökning, som vid huvudförhandling i domstol, där brottslänkning presenteras av sakkunniga.

Summary

Behavioural crime linking refers to the practice of trying to tie two or more offences to the same offender using behaviour observable at the crime scene. It rests on the assumptions that offenders behave consistently enough from one offence to another, and distinctively enough from other offenders allowing offences to be successfully linked together. Conceptualised in the 70s, and developed methodologically with increased scientific rigour from the 90s, the last decade has seen a sharp rise in published studies on behavioural crime linking. From empirical validation of the underlying assumptions to mapping out practice and more ecologically valid tests of linkage accuracy, the field has developed considerably. Considering that investigating homicide is resource intensive, not to mention serial homicide, reliable and valid behavioural crime linking has the potential to aid and prioritise investigative avenues and help solve serial homicide.

Most studies on serial homicide have been carried out on North American samples. While some research has studied the consistency and distinctiveness of serial homicide offenders, few have empirically tested models of behavioural crime linking and linkage accuracy with serial homicide. Another shortcoming in behavioural crime linking research is the use of mostly serial cases to model crime linking, even though real crime databases include both serial and one-off offences. Some studies have tested the effect of added one-offs on the linkage accuracy of burglary and rape, but none so far the effect this would have on homicide. Additionally, while some studies have compared serial homicide offences to one-off homicides, none have tested whether it would be possible to predict whether a homicide belongs to a series or is a singular offence. Cognitive bias, especially confirmation bias or the expectancy effect, has been shown to have a considerable effect on crime investigation. No studies to date have explored the effect of such bias in behavioural crime linking.

The general aim of the thesis was to increase ecological validity of behavioural crime linking research, especially with regard to sampling choices and analyses that strive to answer questions relevant for homicide investigation. The main sample consisted of 116 Italian serial homicides, committed in 23 separate series of homicide. Additionally, information about 45 cases of hard-to-solve one-off homicide was gathered, coded, and added to the sample. **Study I** found seven behavioural dimensions of offending (e.g., sexually motivated homicides and aspects of control-behaviour) in line with previous research. Notably, also other motives than sexual were found in the

killings. A majority of offences (63%) were correctly classified to their actual series in the predictive part of the study. **Study II** was an experiment that investigated whether knowledge of series membership increased perceived (coded) behavioural similarity in homicides committed by the same offender. While no support was found for a strong expectancy effect, the experimental task may have lacked in sufficient complexity, and replication is thus needed. **Study III** found several key differences between serial and singular homicides and was able to successfully use these differences to predict with good accuracy whether an offence was part of a series. **Study IV** combined all the advances in the methodology thus far and showed that behavioural crime linking was still viable even with a large proportion (10:1) of one-off homicides added into the sample. As a function of added one-off homicides, the specificity of the model worsened (more false positives), as did the proportion of offences belonging to a series found near the top of a ranked listing from more behaviourally similar to less behaviourally similar. Overall model accuracy remained good, though, further validating the practice of behavioural crime linking with more ecologically valid data.

The studies of the present thesis contribute to the methodology of behavioural crime linking research. Replication on local crime databases is needed to maximise the practical usefulness of the models in different jurisdictions. Going forward, a close-knit collaboration between researchers and practitioners is called for, to keep the research relevant for practice and to develop evidence-based practice. As we gain a clearer picture of the accuracy and error rate of behavioural crime linking models, their usefulness increase in both the criminal investigative phase and in the trial phase with behavioural crime linking being presented as expert evidence.

Abbreviations

AUC	area under the curve
BCL	behavioural crime linking
SHO	serial homicide offender
ViCLAS	Violent Crime Linkage Analysis System

Table of Contents

1. Introduction

1.1. Criminal Profiling

The basic concept of making inferences about an offender from crime scene evidence is an old one, traceable back to the beginning of the Common Era and the investigation of Blood libel murders (Turvey, 2012). The coining of the term *offender profiling* and the systematic development of the practice is, however, most commonly credited to the members of the Behavioral Science Unit of the Federal Bureau of Investigation in the 1970s (Davis, Rainbow, Fritzon, West, & Brooks, 2018). The method and model of offender profiling initially developed by the FBI have been criticised for lack of empirical validation of the claims made (Rainbow, Gregory, & Alison, 2014). In an attempt to remedy this shortcoming, UK psychologist David Canter advocated for a more robust scientific approach that ushered in a new era of scientific exploration of profiling.

As of yet, there are no internationally agreed upon definitions or standards of profiling or profilers (Davis, 2021). Profilers approach the task from a variety of disciplines, typically criminology (the study of crime and criminal behaviour), psychology and psychiatry (the study of behaviour and mental health), and forensic sciences (the examination of physical evidence) (Turvey, 2012). Some argue that experience of criminal investigation is essential for competent profiling (Douglas, Ressler, Burgess, & Hartman, 1986; Hazelwood, Ressler, Depue, & Douglas, 1995), while others stress the need for anchoring practice in empirical research (Alison & Canter, 1999).

A couple of decades of profiling research has not solved the classical problem of how to apply group level findings to individual cases. Summing up this research in a critical review, Snook and colleagues (Snook, Cullen, Bennell, Taylor, & Gendreau, 2008) aptly asked what was behind the smoke and mirrors of offender profiling. The gist of their criticism was that criminal profiling had no basis in scientific theory and lacked empirical support. Dern and colleagues (Dern, Dern, Horn, & Horn, 2009) replied that Snook and colleagues were throwing out the baby with the bathwater. They pointed out that Snook and colleagues failed to consider the practices of profiling: that there are practitioners mindful of the science and the evidence base, and especially of the limitations that the current evidence holds.

Contemporary profiling is not limited to predicting offender characteristics from crime scene behaviour. Alison and colleagues (Alison, Goodwill, Almond,

van den Heuvel, & Winter, 2010) indicated that a more pragmatic, interdisciplinary practitioner-academic model has emerged in the UK, termed behavioural investigative advice. The advice given includes recommendations for investigative interviewing, behavioural crime linking (BCL) advice, and risk assessment. Subsequently Rainbow and colleagues (2014) described behavioural investigative advice as scientifically based, pragmatic activities that support police investigations, and are based on replicable, transparent, and valid knowledge and research. They stress that behavioural investigative advice should always make the supportive rationale explicit, and that the advice should be falsifiable and given in the form of probabilistic expressions. Rainbow and Gregory (2011) go on to explicate how the scientific method itself can be applied through behavioural investigative advice to enhance the investigative process:

> "through appropriate provision of hypotheses, evidence based prioritisation of the 'most likely' and associated decision support strategies, grounded firmly in psychological principles and available empirical research findings" (Rainbow & Gregory, 2011, p. 21)

Albeit challenging to measure, researchers (Snook, Taylor, Gendreau, & Bennell, 2009; Kocsis, 2013; Chifflet, 2015; Fox & Farrington, 2018; Davies & Woodhams, 2019) and practitioners (Dern et al., 2009) alike agree that there is desperate need for research on the validity of the advice given by profilers.

1.2. Serial Homicide

Violence in general and homicide specifically has been in a steady decline for decades (UN, 2021; WHO, 2021). Globally, homicides have decreased from 6.8/100,000 in 1990 to 5.8/100,000. In Italy, (on which the current research focuses) homicides have decreased from a peak 3.4/100,000 in 1991 to a record low of 0.6/100,000 in 2017 and 2018 (UN, 2021). International estimates on the prevalence of serial homicide are similar. The FBI estimates that less than one percent of homicides committed annually in the US are the work of serial homicide offenders (SHOs; Morton & Hilts, 2008). Estimates in Australia (Mouzos & West, 2007; McKinley & Petherick, 2021) and in the UK (Wilson, 2007) put the proportion of serial homicides at approximately one percent of all homicides, and in Sweden at 1.6% (Sturup, 2018). Upon closer inspection of data from the US, Yaksic and colleagues (2019) found that there had been a decline also in serial homicide since the 1980s.

According to the early definitions of serial homicide, the murderer had to have been involved in three or more separate homicide events with an emotional cooling-off period between the offences (Douglas, Ressler, Burgess, & Hartman, 1986). Some authors have called for an inclusion of motive into the definition suggesting serial homicide offenders (SHOs) are primarily driven by personal gratification, or lust (Hazelwood & Douglas, 1980; Skrapec, 2001). Others have included victimology in their definition, equating serial homicide with the killing of strangers (Egger, 1984). Nearly all the aspects of these definitional notions have been contested. The FBI lowered their requirement of three victims to two nearly a decade ago (Morton & Hilts, 2008), although some argue that offenders who kill only two victims differ from those that kill more and should therefore not be included in the category (Fridel & Fox, 2018). Others argue that restricting the definition based on victim count unnecessarily segments a population that share similar pathologies (Yaksic, 2018). While most victims of SHOs are strangers to their killers, there are also exceptions to this (Morton & Hilts, 2008). Also, while sexual motives seem to be over-represented in serial homicides, as compared to apparent one-off homicides, there are also many other motives present in serial homicides (Harbort & Mokros, 2001; Kraemer, Lord, & Heilbrun, 2004; Morton & Hilts, 2008). There is also considerable overlap between serial homicide and rape: many SHOs also commit rape that does not turn into homicide (Davies, Woodhams, & Rainbow, 2018; Salfati & Sorochinski, 2019). Even the cooling-off period has been debated, as some SHOs occasionally go on killing sprees killing several victims in one event. This has led to the inclusion of the separate category of spree-killing, where an offender kills several persons in two or more temporally related instances without a cooling-off period, into the definition of serial homicide (Yaksic, Simkin, & Roychowdhury, 2021). The contemporary, rather inclusive, and mostly agreed upon definition of serial homicide is the unlawful killing of two or more persons by one or more offenders in separate events (Morton & Hilts, 2008; Adjorlolo & Chan, 2014).

SHOs generally target highly vulnerable victims associated with circumstances of low social control (Egger, 1984; Horning, Salfati, & Labuschagne, 2015). Women and strangers are targeted more often, while one-off homicide offenders most commonly target family and friends (Sturup, 2018). Especially prostitutes run a significantly higher risk of being targeted by SHOs (Sorochinski & Salfati, 2019) than one-off homicide offenders (Sturup, 2018).

Homicide offenders and perpetrators of violence in general display a greater degree of alienation, social maladjustment, and psychiatric pathology compared to the normal population (e.g., Farrington, Gaffney, & Ttofi, 2017). SHOs even more so, when compared to offenders who have killed only once (Fox & Levin, 1998; Harbort & Mokros, 2001; Kramer et al., 2004; Sturup, 2018). In a recent Swedish study, Sturup (2018) compared SHOs to apparent one-off homicide offenders ("apparent", because it is impossible to have full certainty that none of the offenders had committed additional offences they have not been linked to). He found, in line with earlier studies, that the SHOs more often had a personality disorder or an autism spectrum disorder. Most incidents of violence, including homicides, are expressive or impulsive in nature (Feshbach, 1964). Salfati and Bateman (2005) found, though, that a large portion (33%) of SHOs could be classified as instrumental in their offences. Similarly, Sturup (2018) found that SHO were more likely to premeditate their offences than one-off offenders: they displayed a greater degree of planning in their offences, and had a higher forensic awareness (e.g., had taken steps to rid themselves of evidence). In contrast, one-off offenders behaved more impulsively, using opportunistic weaponry such as knives in their killings.

Chan and Heide (2016) found a similar difference between sexual offenders who killed their victims and those who did not: the former were more frequently diagnosed with maladaptive personality traits, sexual sadism and paraphilias compared to the latter. Sturup (2018), again in line with earlier research, found that SHOs commonly had sexual motives, more often so than offenders who only killed once (e.g., Kraemer et al., 2004). It is important to note, however, that many other motives, both instrumental (e.g., a homicide carried out in association with the acquisition of money in burglary or robbery, or professional killers working within organised crime) and expressive (e.g., motivated by rage or the lust to kill), have been found among SHOs. Furthermore, there are considerable overlaps in that one homicide can have several motives and that a series of homicides can have varying motives for each individual killing (e.g., Kraemer et al., 2004; Morton & Hilts, 2008; Petreca, Burgess, Stone, & Brucato, 2020; McKinley & Petherick, 2021).

In summary, serial homicide offenders are a heterogeneous group. Although they display some similarities at a group level that differ in terms of characteristics and crime scene behaviour from offenders that kill only once, their offences reflect a variety of different motives and *modi operandi* (a particular way or method of doing something).

1.3. Behavioural Crime Linking

Behavioural crime linking (BCL, sometimes called crime linking, linkage analysis, case linking or linkage, to name a few variations of the term) refers to an analysis done on crime scene behaviour in an attempt to connect two or more offences to the same offender(s) (Woodhams & Bennell, 2014). BCL falls under the broader task of criminal profiling, together with geographic profiling and predicting offender characteristics based on crime scene behaviour. With successful crime linking, individual offences can be investigated and prosecuted as a whole, pooling both investigative resources and evidence (Grubin, Kelly, & Brunsdon, 2001; Davies & Woodhams, 2019). Another advantage of BCL is that forensic evidence (such as fingerprints, fibres, or DNA) is not always available (Labuschagne, 2006). Once automated applications for BCL are in place, the analysis itself is also fast and cheap by comparison to the analysis of other forensic evidence (Davies & Woodhams, 2019). Especially in cases of serial homicide, the costs of not linking offences are very high, as Yaksic and colleagues (2021a) point out: failing to identify offence series and catching the offenders cost additional lives.

1.3.1. Theoretical Assumptions and Empirical Validation

Behavioural crime linking rests on two theoretical assumptions: behavioural consistency (Canter, 1995) and distinctiveness (Alison, Bennell, Mokros, & Ormerod, 2002; Bennell & Canter, 2002). In practice this means that offenders have to behave consistently enough from one offence to another, and distinctively enough from other offenders, for BCL to reliably identify actually linked offences.

The last two decades has seen a sharp increase in BCL research. Earlier reviews, focusing exclusively on BCL (Woodhams, Hollin, & Bull, 2007; Bennell, Mugford, Ellingwood, & Woodhams, 2014), have concluded that both the assumptions of consistency and distinctiveness have been supported by a good amount of empirical evidence. BCL has been shown to be viable across a variety of crime types, in samples from all over the world: the United Kingdom, the United States, Canada, Finland, South Africa, Japan, Italy and Australia. These studies include rape (Santtila, Junkkila, & Sandnabba, 2005; Yokota, Fujita, Watanabe, Yoshimoto, & Wachi, 2007; Woodhams & Labuschagne, 2012; Winter, Lemeire, Meganck, Geboers, Rossi, & Mokros, 2013; Slater, Woodhams, & Hamilton-Giachritsis, 2015; Oziel, Goodwill, Beauregard, 2015; Sorochinski & Salfati, 2018; Woodhams et al., 2019; Davidson & Petherick, 2020), robbery

(Woodhams & Toye, 2007; Burrell, Bull, & Bond, 2012), and also volume- and property crime: arson (Santtila, Fritzon, & Tamelander, 2004), burglary (Bennell & Canter, 2002; Benell & Jones, 2005; Tonkin, Santtila, & Bull, 2012; Bouhana, Johnson, & Porter, 2016), and car theft (Tonkin, Grant, & Bond, 2008; Davies, Tonkin, Bull, & Bond, 2012).

In a systematic review and meta-analysis of four decades (1976–016) and 426 publications on offender profiling, Fox and Farrington (2018) found that there was significant improvement in the scientific rigour of the research. This advancement was especially pronounced in the BCL research. The authors conducted a meta-analysis on BCL studies published between 2012 and 2016, that reported effect sizes and measurements of variance for their results on linkage accuracy. The 18 studies that were included showed a moderate to strong accuracy rate (overall *AUC* = .83, 95% *CI* = .76–.89) for BCL as measured by the area under the curve (*AUC*). The authors concluded that of all the offender profiling publications reviewed, the BCL studies had the highest overall scientific quality.

A couple of decades of BCL research on serial homicide has also deepened our understanding of SHOs consistency and distinctiveness across their homicide series. Most of the research has been conducted on data from the United States (Salfati & Bateman, 2005; Bateman & Salfati, 2007; Sorochinski & Salfati, 2010; Melnyk, Bennell, Gauthier, & Gauthier, 2011; Salfati & Sorochinski, 2019), but some BCL research has also been conducted on data from South Africa (Salfati, Horning, Sorochinski, & Labuschagne, 2015) and Italy (Salo, Sirén, Corander, Zappalà, Bosco, Mokros, & Santtila, 2013; studies in present thesis). Themes of behaviour have, generally, been found to have higher consistency than individual behaviours. In one of the first empirical studies of behavioural consistency in serial homicide, Salfati and Bateman (2005) showed that SHOs demonstrated some consistency for instrumental and expressive themes over their series. Planning behaviour (both before and after the homicide) seems to be among the most consistent behaviours (Bateman & Salfati, 2007; Sorochinski & Salfati, 2010; Salfati et al., 2015). Interestingly, victim selection does not appear to be as consistent (Salfati et al., 2015; Salfati & Sorochinski, 2019). For example, SHOs that target sex workers commonly also target other types of victims during their series. In a cross-national comparison, Salfati (2014) pointed out the need to study behavioural consistency over homicide series separately in different cultures as, for example, the victims of South African SHOs did not include an overrepresentation of sex workers like other studies on serial homicide have

found. Vulnerable victims seemed to be a consistent theme, though. For example, people looking for and being offered day labour were prone to get into a car with strangers and end up victims of SHOs in South Africa (Salfati et al., 2015).

Melnyk and colleagues (2011; the only study on serial homicide that was included in the Fox and Farrington meta-analysis) analysed 237 North American serial homicides, committed by 79 individual offenders. The authors found that linked crime pairs showed significantly higher behavioural similarity than unlinked pairs, leading them to conclude that their findings supported both the assumption of consistency and distinctiveness in serial homicide. To determine the discrimination accuracy between linked and unlinked crime pairs (i.e. the accuracy of the crime linking decisions), Melnyk and colleagues used receiver operating characteristics and calculated the *AUC* to determine the sensitivity and specificity of their predictions. The discrimination accuracy was found to be high (*AUC* = .96, 95% *CI* = .94–.98), significantly higher than that for burglaries, using the same methodology (*AUC* = .62, 95% *CI* = .50–.74). The accuracy of the crime linking for the homicides was also significantly higher than for any other crime type analysed by Fox and Farrington (2018): sexual offences (*N* = 8 studies; *AUC* = .86, 95% *CI* = .74–.98), burglary (*N* = 11; *AUC* = .73, 95% *CI* = .57–.89), robbery (*N* = 5; *AUC* = .84, 95% *CI* = .60–1.00), and car theft (*N* = 7; *AUC* = .72, 95% *CI* = .54–.90). Melnyk and colleagues (2011) concluded, however, that replication was needed to validate the finding.

BCL accuracy (as measured by the *AUC*) seems to be somewhat higher for person crimes (rape and homicide) compared with property crimes (Melnyk et al., 2011; Fox & Farrington, 2018). It would seem that behaviours that are less situation-dependent and more under the offenders' control show greater consistency (Woodhams & Toye, 2007; Davies, Woodhams, & Rainbow, 2018; Tonkin, Lemeire, Santtila, & Winter, 2019). Some practitioners have called this more consistent behaviour of the offender their signature (e.g., Douglas & Munn, 1992; Keppel, 2000; Hazelwood & Warren, 2004) or psychological fingerprint (Keppel, 2000). In contrast to the *modus operandi* which is thought to be more susceptible to change because of learning and varying situational variables the signature (or ritualistic) behaviours are theorised to be the personal expression of an offender, unnecessary for committing the crime, and indicative of the offender's inner psychology and fantasies (*ibid*.). Empirical testing of the signature construct has shown, however, that serial murderers do not engage in the same rituals and leave unique signatures at every scene,

but that their crime scene behaviour is complex and varied (Schlesinger, Kassen, Mesa, & Pinizzotto, 2010). Hence, while there is a good body of evidence supporting the underpinning assumptions for BCL, Allison and colleagues' (2002) critique remain as poignant as ever: most models are still overly simplistic and fail to fully capture situational variables as well as offender-victim interactions and the way these may influence offender behaviour. The methodology is ever developing, though, and researchers have begun to zero in on these shortcomings.

1.3.2. Behavioural Crime Linking in Practice

Although BCL is practiced widely all over the world and academic research on BCL has taken off well during the last decades there is an urgent lack of publications describing the varying methods that are being employed in practice (Davies & Woodhams, 2019). The earliest descriptions came from the profilers of the FBI's Behavioral Science Unit in the United States (Douglas & Munn, 1992; Keppel, 2000; Hazelwood & Warren, 2004). More recently, descriptions from all over the world have started to emerge, including from South Africa (Labuschagne, 2006), Germany (Dern et al., 2009), the United Kingdom (Rainbow, 2014), Japan (Yokota, Kuraishi, Wachi, Otsuka, Hirama, & Watanabe, 2017), China (Chi, Lin, Jin, Xu, & Qi, 2017), and New Zealand (Tonkin & Weeks, 2021). BCL can be seen as an umbrella term for a family of related techniques (Davies & Woodhams, 2019). In his practitioner's perspective, Rainbow (2014) differentiates between the task of comparative case analysis and crime linkage analysis. He argues that comparative case analysis, the less common task of the two, is what most BCL research is actually looking at: trying to identify linked crimes in a larger database of offences. In the UK, this task is typically carried out by crime analysts, while behavioural investigative advisers are tasked with crime linkage analysis: evaluating two or more offences that a crime investigator suspects may be perpetrated by the same offender. This is a problem, Rainbow argues, as research methodology is focused mainly on the first task, while practitioners for the most part are focused on the latter. Woodhams, Bull, and Hollin (2007) additionally point out, that comparative case analysis can be divided into reactive and proactive linking tasks (crime linkage analysis by definition is reactive). In the reactive comparative case analysis task, an investigator might ask an analyst to find them behaviourally similar offences for an index offence, from a database. In proactive comparative case analysis, analysts might trawl through their databases in an attempt to find previously unidentified crime series.

Rainbow (2014) further points out, that (the British) behavioural investigative advisors are mainly involved with BCL from an intelligence rather than evidentiary perspective. In other words, the advice on crime linking is mainly used in the investigative phase, not in court. BLC advice is sometimes also divided into clinical and statistical (*ibid.*). The former emphasises the need for investigative experience, while the latter relies on scientific and academic research on the matter. Rainbow says that they often are seen as incompatible or mutually exclusive, while they in fact are complementary. A close collaboration between academics and practitioners is needed (Labuschagne & Salfati, 2015). An attempt at bridging this gap and advancing the field of BCL was made in 2013 with the founding of Crime Linkage International NetworK (C-LINK), an international collaborative network of researchers and practitioners specialised in BCL (C-LINK, 2021).

In their recent review of the practice of BCL, Davies and Woodhams (2019) concluded that there is great variability in how BCL is carried out and utilised internationally. They agree with Rainbow (2014), that there is a gap between the empirical validation research, and practice. Davies and Woodhams recommend that more focus be placed on the systematic mapping of BCL practice, so that the empirical parts of the research would better correspond to practitioner needs. Also, while some studies have attempted to pinpoint what makes for good (accurate) profiling and BCL advice, the ecological validity of the studies is generally poor, mainly lacking in task complexity (Davies & Woodhams, 2019). Davies and Woodhams (2019) go on to talk about the human factor of BCL, and how that is another major gap in the research literature. There is a growing body of research on cognitive bias, crime investigation, and forensic expertise (e.g., Ask & Granhag, 2005; Kassin, Dror, & Kukucka, 2013; Dror & Murrie, 2018). In their recent review, Cooper and Meterko (2019) concluded that the finding of cognitive bias influencing expertise in this area is robust, and suggest measures to counteract it, including the reduction of superfluous information, blinding procedures, and peer replication of tasks without the biasing knowledge of the others' conclusion.

Mapping out the practice of BCL in New Zealand, Tonkin and Weeks (2021) concluded that the procedural variability within a single country was also great and called for a standardisation of both data recording and linking methodologies. Reliable and valid BCL requires reliable and valid data: efforts to enhance the quality of crime data recorded by the police will enhance BCL. Once data quality is sufficient and a reliable database is constructed, Tonkin and Weeks suggest the use of computerised tools as an aid for BCL advice. Such

large data repositories already exist in many countries, the internationally most widespread system being the Violent Crime Linkage Analysis System (ViCLAS) developed by crime analysts in Canada (Davies, Imre, & Woodhams, 2021). The question of whether such complex and resource intensive systems are worth maintaining and updating hinges on their continued assessment (regarding reliability and validity) and development. While these systems are far from perfect, it is worth noting that behavioural crime linking is only as reliable and valid as the data it is based on (Pakkanen, Santtila, & Bosco, 2014).

1.3.3. Development of the Methodology

One of the initial attempts at a systematic classification of serial homicide was done by special agents at the Behavioural Science Unit of the FBI. They talked about the lust murderer, being sexually motivated to kill, and being either organised or disorganised in both their actions and to their personality (Hazelwood & Douglas, 1980). Ressler and colleagues (1986) interviewed 36 convicted SHOs and analysed data on 118 of their victims and crime scenes, and as a result categorised 24 of the offenders as organised and 12 as disorganised. The authors postulated that organised crime scene behaviours involve a greater degree of planning and control than disorganised behaviours. Further, organised offenders were hypothesised to commit organised offences, while disorganised offenders would behave in a disorganised fashion at the crime scene. The dichotomous typology and the methodology by which it was developed has been criticised widely (e.g., Canter, Alison, Alison, & Wentink, 2004). For instance, Canter and colleagues (2004) questioned the validity of the sampling: a small convenience sample of 36 convicted killers who agreed to talk to the special agents in exchange for benefits at their prison. Another central point of criticism has been that more emphasis should be put on observable features of the interaction between the offender and the victim, rather than inferences about the motives and intentions of the offender (Canter et al., 2004; Morton & Hilts, 2008). Profiling advice should be given based on data available when the police only have a crime scene, that is, prior to the identification of the offender (Knabe-Nicol, Alison, & Rainbow, 2011). In their own empirical analysis of 100 North American serial homicides, Canter and colleagues (2004) found that rather than the typology being able to distinguish between serial killings, all the offences had organised features to them. They argued that this is what lies at the core of the definition of serial homicide: offenders that are able to plan and control the circumstances of their offending

to a point where they are able to commit several homicides without being caught.

The organised/disorganised typology is perhaps the most cited serial homicide typology and is still echoed in current research on SHOs even though attempts at empirical validation have failed to give it unequivocal support. While these initial studies on serial homicide classification contributed to a deeper understanding of the dynamics of serial homicide, the methodological problems limit its practical usefulness for BCL.

1.3.3.1. Sampling

Authors have frequently brought up the need to improve the ecological validity of BCL research. One of the main problems is research being carried out on unrepresentative samples (Alison et al., 2010). More specifically, a concern has been raised for BCL research using only solved and serial cases (Bennell & Canter, 2002; Woodhams, Bull, & Hollin, 2007). Using only solved cases might lead to an overestimation of behavioural similarity and inflated estimates of linking accuracy as more behaviourally similar cases may be easier to link and solve in the first place (Melnyk et al., 2011). Researchers have started to address this methodological shortcoming by identifying series in police databases that are linked by DNA, as opposed to similarity in MO, or linked by DNA, but still unsolved.

Woodhams and Labuschagne (2012) were the first to study behavioural similarity in crimes that were either unsolved or first linked by DNA. Their sample consisted of 119 South African rapes committed in 22 series. Comparing behavioural similarity of the rapes linked by *modus operandi* vs. DNA, the difference (as measured by Jaccard's coefficient) was significant, but small (M = .51 vs. M = .47; t = 1.98, p = .05). For their total sample, linkage accuracy remained high (AUC = .88, 95% CI = .86–.90). Similarly, Tonkin and colleagues (Tonkin, Woodhams, Bull, & Bond, 2012) gathered a sample of 132 offenders that had committed 264 offences of different crime types in the United Kingdom, of which three quarters were unsolved but linked by DNA. Comparing their findings to their earlier study conducted on only solved cases (Tonkin, Woodhams, Bull, Bond, & Palmer, 2011), the authors concluded that the results were for the most part replicable and BCL viable, although linkage accuracy was negatively affected (mean reduction in AUC = .06, range = .00–.21). Lastly, Woodhams and colleagues (2019) analysed a large international sample of 3,364 sexual crimes, including 668 series, and 92 unsolved cases that were linked by DNA. Overall, they found a good level of discrimination accuracy

between linked and unlinked cases (*AUC* = .85, 95% *CI* = .84–.86). The accuracy was decreased when conducted solely on the unsolved cases (*AUC* = .79, 95% *CI* = .77–.81), but remained moderate and in line with the results on previous studies on BCL in sexual offences. In summary, even though decrements in linking accuracy were observed with the inclusion of unsolved cases, differentiating linked cases from unlinked cases remained viable.

Another major issue with sampling is that early research on BCL was carried out using only serial cases. Tonkin and colleagues (Tonkin, Santtila, & Bull, 2011) were the first to address this particular concern in their study of 508 Finnish burglaries, where a fourth of their sample included apparent one-off offences. The authors found that including one-off offences in their analyses had no significant effect on linking accuracy. Later studies have added one-off offences to their samples of sexual assaults (Winter et al., 2013; Slater et al., 2015; Woodhams et al., 2019; Davidson & Petherick, 2020). While there are some negative effects on BCL accuracy (discussed in detail in Study IV), the assumptions of consistency and distinctiveness and the ability to distinguish between linked and unlinked crimes in mixed samples of serial and one-off offences still stand. The studies have used relatively small proportions of one-off to serial offences, which may not be a big problem in BCL research on sexual offences (or volume crime such as burglary) as many stranger assaults are committed by serial offenders. For homicides, however, the problem is more pressing as most homicides are one-off offences. In his comparison of serial and one-off homicides Sturup (2018) showed that the two could be distinguished from each other with moderate accuracy (*AUC* = .76, 95% *CI* = .70–.83), which could prove useful for making investigative decisions in homicide investigations.

Not all homicide investigations require the assistance of behavioural investigative advisers. Most homicides are committed by people close to the victim (e.g., Lehti, 2020), and are both faster and easier to solve than serial homicide. Some authors have therefore focused on samples of stranger homicides (e.g., Salfati & Canter, 1999; Greenall & Wright, 2020) as a ground for profiling advice. From the point of view of what is known at the onset of the investigation, this can be problematic. Often the relationship between the offender and the victim may not be known at the point in time, when behavioural investigative advice is sought (Salfati & Canter, 1999). Therefore, we proposed a more pragmatic definition of hard-to-solve homicide by looking at the time between when the offence came to the knowledge of the police and when the police first interviewed the offender as a suspect. In a Finnish sample

of 1017 homicides committed between 1990–2001, 71 offences met the criteria of "at least 24 hours" for that time interval (Pakkanen, 2006).

A related issue is that studies have restricted sample selection to sexual homicides (e.g., Sturup, Rodre, Karlberg, von Vogelsang, Rying, & Caman, 2019; Greenall & Wright, 2020). While these studies do contribute to the understanding of the dynamics of sexual homicide, their applicability to behavioural investigative advice on homicide and serial homicide in general is limited, as serial homicide encompasses many other motives as well (e.g., Morton & Hilts, 2008). Motive in general might make for a poor inclusion criterion, as it is oftentimes not apparent at the onset of the investigation. On the other hand, sexual motives are over-represented in serial homicide, and there is considerable overlap between the crime types of homicide and rape (Davies, Woodhams, & Rainbow, 2018; Beauregard, DeLisi, & Hewitt, 2018). This raises another important question of conducting BCL over these crime types, which has not yet been done (although the behavioural dynamics of the two have been compared in, e.g., Salfati & Porter, 2006). Some tentative studies have successfully attempted BCL over crime types with different types of property crime (Tonkin et al., 2011; Tonkin et al., 2012b; Tonkin & Woodhams, 2017), looking at a combination of intercrime distance, temporal proximity, and behavioural similarity, suggesting it could be feasible with homicide and rape also.

1.3.3.2. Methods of Analysis

There are several ways of conducting BCL. Some methods are more established than others (Bennell et al., 2014). Offences can, for example, either be compared pairwise for behavioural similarity (e.g., Woodhams et al., 2007b), or by trying to predict series membership directly (e.g., Salo et al., 2013; Porter, 2016). The different methods could be seen as reflecting different kinds of BCL tasks: the pairwise comparison being more in line with Rainbow's (2014) crime linkage analysis, and the prediction of series membership corresponding to comparative case analysis and a proactive BCL approach (Woodhams et al., 2007b).

A method developed by Craig Bennell (2002) (Bennell & Canter, 2002; Bennell & Jones, 2005) has become standard in BCL research for examining behavioural similarity and its ability to distinguish between linked and unlinked crimes (Tonkin et al., 2017). The method comprises of calculating the similarity between two offences using Jaccard's similarity coefficient *J*: $\boldsymbol{J = a / (a + b + c)}$, where **a** is the number of behaviours present in both offences, **b** the

number of behaviours present in offence one but not in offence two, and **c** the number of behaviours present in offence two but not in offence one. A coefficient of 0 would indicate no similarity whatsoever between the two offences, while a coefficient of 1 would indicate perfect similarity, where all the same behaviours are present in both homicides. Jaccard's coefficient is usually the preferred measure of similarity because it emphasises the joint occurrence of a variable in a crime pair while disregarding joint non-occurrence. This is advantageous with behavioural offence data, as there is a larger degree of uncertainty associated with the absence of observation, which can arise because the perpetrator has not engaged in a behaviour, or because data are missing for said behaviour. For example, investigators at the crime scene might have missed that something of non-monetary value was stolen from the victim, or a researcher coding the data might have missed that particular detail in an extensive data file.

To determine what level of similarity should qualify a decision of two crimes being linked, Bennell, Jones and Melnyk (2009) argue for the use of receiver operating characteristic curves. The curve illustrates the diagnostic ability of a binary classifier system (e.g., "linked" or "not linked") as its discrimination threshold (e.g., behavioural similarity, as measured by Jaccard's similarity coefficient) is varied. The curve plots the true positive rate (a "hit"; a correctly identified link) against the false positive rate (a "miss"; an incorrectly identified link) at various thresholds, thus enabling an examination of the BCL-model's sensitivity and specificity. The area under the curve (*AUC*) can be used as a standardised measure of the model's accuracy. According to Swets (1988), an *AUC* of under .50 is non-informative (no better than chance), an *AUC* of .50–.69 represents low accuracy, .70–.89 moderate, and over .90 high accuracy. BCL research has generally been found to yield moderate to high levels of accuracy, as measured by the *AUC* (Bennell et al., 2014; Fox & Farrington, 2018), see Table 1.

Table 1

The Distribution of Effect Sizes of Behavioural Crime Linking Accuracy Compared by Two Meta-Analyses

Study	Non-informative *AUC* = 0–.49	Low accuracy *AUC* = .50–.69	Moderate accuracy *AUC* = .70–.89	High accuracy *AUC* = .90–1.0
Bennell et al. (2014): 19 studies, 146 *AUC* values	2%	29%	54%	15%
Fox & Farrington (2018): 18 studies, 40 *AUC* values	0%	25%	58%	18%

Note. AUC = area under the curve.

Adjusting the thresholds allows for application of the model to different scenarios that may have different requirements. For example, in the investigative phase an analyst would want to make sure not to miss any possible links (a greater tolerance for false positives), while a court, considering BCL-evidence, would want to make sure the defendant indeed is guilty (minimise risk for false positives) (Turvey & Freeman, 2016).

Even though BCL studies use a variety of different methods, few studies compare methods directly (e.g., Winter et al., 2013; Tonkin et al., 2017; Tonkin et al., 2019). In a recent comparison, Tonkin and colleagues (2019) compared 7 different statistical methods on three separate sets of data: residential burglary, car theft, and commercial robbery. The authors produced both *AUC*s and ranked listings of the cases from most similar to least similar. The *AUC*s varied from .44 to .90, and the authors concluded that there is no robust, generalisable evidence to support the superiority of any statistical method over others, and that the choice of method depends (at least) on crime type, the base rate frequency of linked vs. unlinked offences, and what type of offender

behaviour is available. Tonkin and colleagues also noted that when comparing linked and unlinked crime pairs, relying on *AUC* as a sole measure of accuracy may be misleading. This is due to the class imbalance problem: when all offences in a dataset are paired with each other, the unlinked pairs vastly outnumber the linked ones, resulting in a large number of correct negative predictions and thus inflating the *AUC*. For example, their best model for series of car thefts displayed a very high accuracy as measured by the *AUC* (*AUC* = .82, 95% *CI* = .82–.83), yet the ranked listing revealed that less than 10% of the linked car thefts could be found among the 1000 most similar cases.

Another topic of debate has been whether to use behavioural dimensions (clusters of behaviours that relate to the same psychological meaning, such as planning behaviour, escape behaviour, and controlling the victim) or separate, individual behaviours (Salfati & Bateman, 2007; Winter et al., 2013) in analyses of behavioural consistency. The choice has in part been guided by method: some statistical methods (e.g., discriminant function analysis) are not able to handle hundreds of behavioural variables, while others (e.g., Bayesian modelling) are. Direct comparisons are few, but Tonkin and colleagues (2019) found the differences to be small in general, although there was a trend towards dimensions being more efficient. While still needing replication on other samples (especially person-on-person offences) dimensions may capture behavioural consistency better than individual variables (*ibid*.) as situational factors vary from one offence to the next, and the offender may not have the opportunity to behave in the exact same manner. For example, controlling the victim verbally may exclude the need for the use of bindings, but both behaviours are still means of controlling the victim.

In order to address the variability of the situation, some complementary methodology and research is beginning to emerge. While the interviews with SHOs conducted by Ressler and colleagues (1986) was a good start, similar but more systematic and scientifically rigorous paradigms are being utilised. Woodhams and colleagues (2008) examined "if (situation)-then (behaviour)" contingencies in sexual assaults, pointing out that most BCL research has exclusively been looking at the latter ("then") part, disregarding situational variability. Failing to find behavioural consistency in the contingencies, the authors concluded that while being a preliminary study, the mounting evidence for behavioural consistency and distinctiveness with the traditional approaches may actually be enough. Winter and Rossi (2020) tried another approach with sequence analysis. Traditional paradigms, they argue, only represents the outcome (all crime scene behaviour examined simultaneously),

while disregarding the deeper dynamics of how sequences of behaviour, both of the victim and offender, develop. The authors argued that this model does not reflect daily practice, and crime linkage analysis (as defined by e.g., Rainbow, 2014). The resulting discriminatory accuracy was moderate (*AUC* = .74, 95% *CI* = .73–.75) compared to previous crime linkage studies. Woodhams and Komarzynska (2014) underline the importance of combining these complementary paradigms in order to get a clearer picture of offence behaviour, its consistency and variability, and how it affects behavioural crime linking.

1.3.4. Behavioural Crime Linking as Expert Evidence in Court

What Rainbow (2014) termed intelligence and evidentiary applications of BCL, Turvey and Freeman (2016) divide into investigative and forensic. Nomenclature aside, BCL has been presented as evidence in court all over the world for a long time under different names and rules of admissibility (Labuschagne, 2014; Pakkanen et al., 2014). While there is no generally agreed upon gold standard for how to perform BCL and present it as evidence, Labuschagne (2014) points out that expert evidence on BCL can never solely rely on statistical calculations. Canter (2004), on the other hand, points out that it cannot be based solely on systematised experience that has not been subjected to some form of psychometric testing either. Although rules of admissibility vary in different countries and jurisdictions, some have argued that the Daubert standard, set forth by the United States Supreme Court, is the closest thing we have to a judicial definition of the scientific principle (e.g., Pakkanen et al., 2014; Davies, Woodhams, & Tonkin, 2019).

The Daubert standard is meant to guide the evaluation of admissibility for scientific evidence. The evaluated aspects are (from Pakkanen et al., 2014, p. 227):

1. *Reliability*. The Daubert test requires that the methods used and the theories referred to must be tested and reliable and that statements made by the expert are falsifiable.
2. *Peer review and scientific publication*. To verify reliability of theories offered by an expert witness, they must have been previously published or submitted to the scientific community for review.
3. *Error rate*. The expert must be able to offer to the court the potential error rate of the findings and methods.

4. *General acceptance*. The general acceptance from the scientific community, while important, should not be considered a precondition for the admissibility of scientific evidence.
5. *Standard*. It is important to identify the existence and respect of standards known and recognised for the procedure and the methods adopted in the expert opinion.
6. *Applicability to the concrete case (fit)*. Finally, for the method or theory to be scientifically reliable, it must fit the concrete case at hand and must be logically linked to it.

Another key concern is the generalisability of the BCL research to a particular case at hand; the so-called group to individual (g2i) -inference (Faigman, Monahan, & Slobogin, 2014). In their review of the intricacies of this process, Faigman and colleagues (2014), point out that the Daubert standard is applied differently at both (g and i) levels. In the case of BCL, the expert evidence could ideally be given at the g-level by a researcher of BCL, while a particular case could be commented on at the i-level by, for example, a behavioural investigative adviser with clinical expertise. The evaluation of the experts' evidence at the i-level should, according to Faigman and colleagues (2014), start with the examination of "Has the expert followed the current guidelines of the field?" Such an evidence-based standard for BCL, that the researcher and practitioner community would agree on, is yet to emerge in the field of BCL.

In the most updated review of the state of the field of BCL research, Davies and colleagues (2019) concluded that while the field has made significant advances in the last years, there are still issues to be addressed before BCL meets the criteria set forth in the Daubert standard. While more research is needed on how BCL is carried out in practice, Tonkin and Weeks (2021) suggest researchers also explore the use of BCL in court more in detail. As the ecological validity of the research is increased, methods are refined, coding standardised, and databases cultivated, our grasp of the accuracy and error rates of BCL become clearer and the scientific foundation on which the practice of BCL rests, sturdier. This advancement also makes BCL easier to consider from an evidentiary perspective.

2. General Aims of the Thesis

The aim of this thesis was a pragmatic one: to develop BCL research in a direction useful for homicide investigators. This general aim could be divided into two more specific aims. First, to develop the methodology of behavioural crime linkage research by increasing ecological validity. Specifically, this means developing research datasets to better correspond with the form and content of real crime databases and to base predictive analyses on information readily available at the beginning of the homicide investigation. This includes sample selection: not limiting case selection by inferred motive and including hard-to-solve one-off homicides, and selection of variables used for the predictive analyses. Second, to pose research questions that are relevant for practitioners. On a practical level, this entails conducting analyses that attempt to answer actual investigative and analytical questions encountered in homicide investigation. Can series of homicides be successfully linked based on crime scene behaviour? (Study I) Is this also possible when hard-to-solve one-off homicides are added? (Study IV) Do we need to apply blinding procedures for coders of the research data (or analysts preforming comparative case analysis)? (Study II). Can we distinguish a serial homicide from a one-off homicide based on crime scene behaviour? (Study III).

A lot has happened in BCL research during the decade and a half that the studies in this thesis encompass. The first study set out to do BCL on homicides using an approach (Mokken scaling and discriminant function analysis) previously applied on series of rape (Santtila, Junkkila, & Sandnabba, 2005) and arson (Santtila, Fritzon, & Tamelander, 2004). The second study and the additional analysis were designed to critically examine the result of the first study by looking for possible confounds in the results. The second study was an experiment examining possible expectancy effect in the form of coder bias for the data, while the additional analysis looks at the effect of the data source on behavioural similarity. The third and fourth study examined the differences between serial and hard-to-solve one-off homicides, with an aim to distinguish between the two. The fourth and final study, taking advantage of methodological advances made by Salo and colleagues (2013), examined how BCL accuracy was affected when the research sample is more in resemblance with real crime databases with a mix of both serial and one-off offences.

The general discussion synthesised the results of the studies placing them in the context of the last decade of BCL research with a specific focus on

implications for (1) research on BCL, (2) the applicability and use of BCL in homicide investigation, and (3) the use of BCL as expert evidence in court.

3. Summaries of the Studies

3.1. Study I: Behavioural Crime Linking in Serial Homicide

Earlier studies on serial homicide have sought to understand and model crime scene behaviour through simple dichotomous categorisations (e.g., organised/disorganised, and expressive/instrumental). While researchers have found these dimensions to be present in homicidal behaviour, these overly simplistic models have failed to categorise samples of homicides in a way that would benefit BCL in practice (e.g., Canter, Alison, Alison, & Wentink, 2004; Salfati & Bateman, 2005). The intricacies of behavioural patterns that make up homicide may not lend themselves to simple categorisation. Most offences include aspects of both aforementioned categories. Consequently, these models have fallen short in offering aid to behavioural crime linking of homicides. The first aim of the study was, therefore, to identify dimensions of variation in cases of serial homicide beyond a dichotomous categorisation. The second aim was to use the identified dimensions to behaviourally link offences committed by the same offender. Successful linking would also provide additional support for the consistency and distinctiveness hypotheses for serial homicide.

3.1.1. Method

A sample of 116 Italian serial homicides committed during the years 1970–2001 were gathered, through public records; mainly courts and mass media. A total of 23 homicide series were identified and the crime features were coded from court files (n = 15) and criminological literature (n = 8). The inclusion criteria were two or more victims and a cooling-off period of at least 24 hours between the homicides. The number of victims varied in the series from 2 to 17, with a median of 6 victims (M = 6.7, SD = 4.5). The length of the series varied from all offences carried out within a single year to 16 years, with a median duration of 3 years (M = 4.8, SD = 5.1). As some offences had more than one offender, and some more than one victim, all offender-victim interactions were coded separately for a total of 155 unique pairings. All offenders were convicted of their crimes.

The crime features were coded building on coding schemes developed by earlier research by Samantha Lundrigan (Hodges, 1998) and Gabrielle Salfati (1998). Most variables were dichotomous, coded as either present (1) or absent (0). The variables included situational variables (pertaining to the when

and where of the homicide), offence related variables observable at the crime scene, associated with the killing itself (e.g., the use of weapons, binds, and gags, as well as injuries of the victim), post-mortem activity (e.g., moving, hiding, or destroying the body), and victim characteristics (e.g., age, gender, relationship status, employment status, known health issues). The inter-rater reliability of the coding scheme was considered good, with a mean Cohen's K of 0.72 (*SD* = 0.13, *Range* = 0.59–0.89).

The dimensions of variation were identified using Mokken scaling, a non-parametric equivalent to factor analysis that produces nonexclusive scales. The linkage analysis itself was carried out using discriminant function analysis, with the identified dimensions as independent variables and series membership as the dependent variable. In other words, the discriminant analysis provided a predictive model of series membership based on the observed crime features (scores of the Mokken scales).

3.1.2. Results

Seven dimensions were identified. Five of them described motivational aspects of the offences and two the level of planning. Three of the five motivational scales were interpreted, based on the crime features that made up the scales, to have to do with instrumental motives: the acquisition of money, status, and revenge. The two remaining motivational scales were interpreted to reflect a sexual motive: either more normative rape or rape with paraphilic interests at play. The last two scales were interpreted to reflect on the level of planning of the offence: either controlled or impulsive.

The crime linkage analysis resulted in 63% of the homicides correctly attributed to their right series, compared to a chance expectation of 6%. Three of the scales (Sexual: Paraphilic, Sexual: Rape, and Level of planning: Controlled) contributed the most to the model, correctly linking 56% of the cases by themselves.

3.1.3. Conclusion

The identified dimensions of offending were well in line with previous models of homicide and interpersonal violence (e.g., Buss, 1998; Wilson & Daly, 1998) where motives of money, status, revenge, and level of planning have repeatedly been found as underlying psychological dimensions of homicidal aggression. The emergence of these motives also lent support to the hypothesis that not all serial homicides are sexually motivated. The dimensions of sexual motivation and level of planning correlated with each other and corresponded to Keppel

and Walter's (1999) earlier model on sexual violence. The large amount of correctly linked cases provided additional support to the hypotheses of consistency and distinctiveness in serial homicidal behaviour. The accuracy of the BCL was significantly higher than in earlier studies using the same methodology that managed to classify 26% of rapes (chance expectation 3%; Santtila, Junkkila, & Sandnabba, 2005) and 23% of arsons into their correct series (chance expectation 8%; Santtila, Fritzon, & Tamelander, 2004).

3.2. Study II: The Effects of Coding Bias on Estimates of Behavioural Similarity in Crime Linking Research of Homicides

Seeing how the effectiveness of the BCL in Study I was significantly higher than in earlier studies, the aim for Study II was to look for possible confounding errors in Study I. One possible candidate was the expectancy effect. According to the effect, our observations are affected by what we expect them to be. The expectancy effect (or confirmation bias) has been shown to affect decision making in a number of studies and contexts, also in criminal investigations (e.g., Rosenthal & Jacobson, 1968; Risinger, Saks, Thompson, & Rosenthal, 2002). No prior studies on the expectancy effect on BCL had been carried out. The aim of Study II was, therefore, to investigate whether prior knowledge of series membership effected perceived behavioural similarity in cases researchers or analysts code. An experiment with three conditions was devised to test this. The three experimental groups were given either correct, incorrect or no information on case linkage prior to being asked to code crime scene behaviour from case vignettes. The first hypothesis was that the cases the participants were told were linked would be perceived as more similar than the cases that they were told were not. The second hypothesis was that the incorrectly informed group would code less similarity in homicides that were actually linked, than the not informed group. An additional hypothesis was tested, predicting that the group without any prior linkage information would more often correctly link together homicides that they perceived (coded) having higher behavioural similarity.

3.2.1. Method

The experiment utilised the help of 60 Italian university students, who were randomly assigned to three groups of 20. Ten case vignettes where 5 offenders had committed 2 offences each were created by heavily editing and summarising court transcripts, resulting in approximately 15 lines of text per vignette. The independent variable was the information provided on series membership prior to the coding (correct/incorrect/no information). The dependent variable was behavioural similarity of the coded crime scene behaviour. Each group was given information on linkage according to their assigned experiment group and asked to code all 10 cases. After the coding task the “no information” -group was told the cases included 5 linked offences and were asked to identify these links.

Behavioural similarity in linked pairs was analysed using the φ coefficient and compared between the groups to check if group allocation would predict differences in coded behavioural similarity. Finally, the linkage decisions of the third group (not informed) were checked to see whether correct decisions were predicted by increments in coded behavioural similarity.

3.2.2. Results

The analyses found no significant differences in coded behavioural similarity between the three experiment groups. The highest behavioural similarity for the linked pairs was coded by the not informed group, but the difference did not reach formal levels of significance (p = .07). The overall inter-rater reliability (N = 60; Kuder-Richardson Formula 20) for the coding was very high for all groups (r_{KR20} = .93). The not informed group made 61% correct linking decisions (n = 17). The pairs of homicide that the participants had coded as having higher behavioural similarity were significantly more often correctly linked together.

3.2.3. Conclusion

The first and the second hypothesis were not confirmed. The experiment thus failed to provide evidence for confirmation bias in the coding of data for BCL. The additional hypothesis was confirmed: participants intuitively linked cases more often correctly if they had coded the cases as being more behaviourally similar. The main weakness of the study was the poor ecological validity of the case vignettes. Being heavily edited and summarised, the vignettes likely failed to reproduce the ambiguity oftentimes found in case materials in regard to the coding of particular elements of crime scene behaviour. Homicide investigation protocols are usually also very extensive making it laborious to find the pertinent information for the coding, another possible source of coding error. The very high overall inter-rater reliability would seem to support this interpretation.

While the study did not confirm the existence of a clear confirmation bias, it was not able to exclude a weak one. Replication of the experiment is needed with a design that improves on the ecological validity of the stimulus material for the coding-task. The risk of the expectancy effect influencing an analysis of similarity is particularly salient in crime linkage analysis, where a behavioural investigative advisor gives an opinion on the linkage status of two or more offences. If the investigator asks the behavioural investigative adviser “Are these particular offences committed by the same offender?”, it is obvious what

the expected suspicion of the investigator is. To mitigate the risk of expectancy effect in research in general, and crime investigation in particular, authors recommend utilising blinding procedures (e.g., Risinger et al., 2002; Sheldrake, 1998; Wilkinson, 1999).

3.3. Additional Analysis: The Effect of the Source of the Data on Estimates of Behavioural Similarity

Another possible candidate for a confounding error in Study I was the data source. Of the 23 series of homicide, roughly one third (n = 8) were coded from true crime literature, while the rest (n = 15) used court transcripts as the primary data source. An over-reporting of behavioural similarity in the source material would distort the findings of the BCL resulting in more efficient linking. It was speculated that true crime literature might not aim for objectivity with the same rigour as court proceedings, as it is produced in large part for entertainment. Also, true crime literature is usually produced on cases that are already tried in court and the guilt of the offender has been established to some degree of certainty (e.g., beyond a reasonable doubt). On the other hand, the legal system is designed to examine evidence for and against the guilt of the defendant, and might thus more systematically look for and report, differences in behaviour, alongside similarities. The hypothesis to be tested was, therefore, that the cases coded from true crime literature would demonstrate higher behavioural similarity than the cases coded from the court transcripts. Consequently, the null hypothesis was that there would not be a significant difference in behavioural similarity between the cases coded from true crime literature and court transcripts. This, in turn, would increase the validity of the results of Study I, as another potential source of error could be eliminated.

3.3.1. Method

For the analyses, all of the 116 offences of the 23 series were included. The same set of dichotomous variables (N = 89) analysed in studies three and four were used, and missing values were substituted with 0. Behavioural similarity was calculated for all linked crime pairs using Jaccard's coefficient, in line with methodological considerations made by Bennell (2002). The mean similarity of series coded from court transcripts (n = 15), and true crime literature (n = 8) were compared with each other using a Welch two sample t-test.

3.3.2. Results

The mean behavioural similarity, as measured by the Jaccard's coefficient, for the series coded from court transcripts was slightly higher (J = .55) than in the series coded from true crime literature (J = .43). The difference was not significant (t = 1.67, df = 16.78, p = 0.11).

3.3.3. Conclusion

Based on the result of the t-test, the hypothesis that true crime literature of serial homicide would demonstrate greater behavioural similarity than court transcripts was not confirmed. Since no significant difference in behavioural similarity between the two data sources was found, and thus no evidence of this possible confounding error, the result of the BCL of the first study was validated further. The result, however, does not exclude the possibility that both sources could overestimate behavioural similarity compared to, for example, pre-trial investigation protocols. Cases that are more behaviourally similar could be more likely to get prosecuted and lead to convictions. An overestimation of behavioural similarity could also have happened even earlier during the investigation, in that behaviourally dissimilar cases could have been falsely dismissed or not detected as part of the series.

3.4. Study III: Can Hard-to-Solve One-off Homicides be Distinguished from Serial Homicides? Differences in Offence Behaviours and Victim Characteristics

Behavioural crime linking research has traditionally been carried out on serial cases exclusively. This oversight in ecological validity might overestimate linking efficiency, as police databases are a mixture of serial and non-serial offences. While recent studies on rape have included one-off offences, serial homicide research has not. In approaching the question of "How does including one-offs in the data affect BCL?", a first step is to ask, "Do serial offences differ from one-off offences and can they be distinguished from each other?" Also, being able to distinguish between serial and one-off homicides based on offence features of a new crime scene could be practically useful for investigators managing finite resources. While there are a few studies comparing serial to one-off homicides (e.g., Fox & Levin, 1998; Harbort & Mokros, 2001; Kraemer, Lord, & Heilbrun, 2004), no study to date (2015) had attempted to distinguish serial homicides from one-off homicides. The first aim of Study III was to add to the existing literature of exploring the differences

between serial and one-off homicides. The second aim was to determine if it is possible to distinguish between the two types of homicides based on offence behaviour and victim characteristics.

3.4.1. Method

The sample gathered for Study I (23 Italian homicide series, encompassing 116 cases), was utilised for the analyses. Generally speaking, most homicides are expressive and solved comparatively quickly. Cases where the offender is identified from the start of the investigation make a poor target for profiling research, however, since investigators of these cases are unlikely to need or ask for behavioural investigative advice. Hence, for the one-off homicides, hard-to-solve cases were selected. Hard-to-solve was operationalised by looking at the time from when the offence came to the knowledge of the police to when the offender was caught or interrogated as a suspect of the offence. Hard-to-solve was defined as this time delay being at least three days (72 hours). A total of 45 cases of hard-to-solve one-off homicides were identified by Reparto Investigazioni Scientifiche di Roma (the department of scientific investigation within the Arma dei Carabinieri) committed during the years 2001–2014. As with the serial homicides, some of the one-off cases (n = 3) had multiple victims, and some (n = 5) had multiple offenders. For the analyses, however, a single offender-victim pairing per case was used (N = 161), to reduce possible error of inflation of certain crime scene behaviours. All offenders had a guilty verdict. Crime features of the one-off offences were coded from the case information (pre-trial investigation protocols) of the local police districts. The coding-scheme developed for Study I was used also for the one-off offences. All missing data, for both the serial and one-off homicides, were imputed because logistic regression is unable to handle missing values.

The crime features of the serial offences were compared to the crime features of the one-off offences using cross-tabulations (χ^2 and Fisher's exact test). A Benjamini-Hochberg correction was used to compensate for the high number of calculations. The first and last offence of a series were compared separately with the one-off homicides, to check whether there was a difference in the homicides from the beginning or whether some sort of progressive change was happening through a series. Logistic regression with a leave-one-out cross-validation was done to predict which group an offence belonged to. The accuracy of the predictions were analysed with receiver operating characteristics and further validated with a generalised estimating equation to

ensure that dependency in the data (several homicides done by one offender in the serial cases) did not alter the predictions significantly.

3.4.2. Results

Several statistically significant differences were found in crime behaviour (n = 10) and victim characteristics (n = 3) between the serial homicides and the hard-to-solve one-off homicides. Serial homicide offenders more often targeted strangers, prostitutes, and younger victims while one-off killers more often targeted females. The serial homicide offenders displayed a higher level of forensic awareness both before (brought a weapon to the crime scene) and after (evidence destroyed by the offender) the killing and had more often an apparent sexual element to their offence. Overall, the one-off homicides had evidence of being more impulsive and expressive (more violence, more violence without the use of weapons). Comparing the first and last offences of a series with the one-off homicides gave further support to the hypothesis that there was a qualitative difference between the two groups from the start of a series, providing evidence against the null hypothesis of serial homicide offenders starting out the same as one-off offenders and then learning or progressing their *modus operandi*. The regression analyses further confirmed that the difference in crime features was measurable and that it could be used to distinguish between serial and one-off homicides at a level that significantly exceeds chance (AUC = .88, 95% CI = .82–.94).

3.4.3. Conclusion

A number of differences in crime features, both in victimology and crime scene behaviour, between serial homicide offences and hard-to-solve one-off offences were identified. For the most part, the differences identified in Study III were in line with earlier studies comparing serial and none-serial homicides. The results further supported the hypothesis that a homicide could be identified as either being part of series or a one-off killing, based on the crime scene behaviour. This in turn could prove useful for homicide investigations, in alerting the investigator to consult for further behavioural investigative advice where needed. It is likely that local and cultural variations and trends in offence features and crime scene behaviour vary to the degree where local replication is needed to establish practically useful models for different countries and cultures.

3.5. Study IV: Linking Serial Homicide – Towards an Ecologically Valid Application

In the fourth and final study, we wanted to move towards higher ecological validity with our BCL of serial homicides by including hard-to-solve one-off offences into our sample. This was a step closer to what actual crime databases look like, where most cases are one-off homicides. Tonkin and his colleagues (2011) were the first ones to add one-off offences into their crime linking analysis of burglaries. Some studies have followed suit with BCL of rapes (Winter et al, 2013; Slater et al., 2015; Woodhams et al., 2019), showing a slight and negligible negative effect on BCL accuracy. The proportion of one-off cases in these samples of rape has been small, though; a problem that accentuates in homicides, as the overwhelming majority of homicides are one-off offences. BCL of homicides, utilising mixed samples, had not been carried out previously. Also, most previous research conducted on mixed samples have used a fixed proportion of serial to one-off offences, which limits the evaluation of the effect of added one-off offences on BCL accuracy. The purpose of the fourth study was, therefore, to investigate how BCL accuracy of serial homicide is affected as a function of added hard-to-solve one-off offences. The hypothesis was that increasing one-off offences to the sample would add noise to the data, and thus decrease linkage accuracy.

3.5.1. Method

The fourth study utilised the same serial homicides (n = 116) as the first study, and the same hard-to-solve one-off homicides (n = 45) as the third study. Single offender-victim pairings per case were used for the analyses and simulation. The familiar coding scheme from all three previous studies was used, and 89 dichotomous variables were analysed. Missing values were retained in the main analyses but substituted with zero in the similarity comparison calculated with Jaccard's coefficient. With a global estimate of 1–2% prevalence of serial homicides, and an earlier estimate of roughly 10% of homicides being hard-to-solve, our estimate put the ratio of hard-to-solve one-off to serial homicides at roughly 10:1. As the exact proportion is unknown, we sought a study design that would allow for a varied proportion.

From the original 45 hard-to-solve one-off homicides, a total of 1–1044 were simulated (N = 117–1160). Consequently, the ratio of serial to one-off homicide was 116:1 to 1:10. A simulation approach was chosen, as it was deemed prohibitively laborious to identify and code more than 1000 cases. The validity

of the simulation was estimated by making a number of simulated datasets and comparing the results of the BCL for the simulated one-off homicides to the original ones.

The primary BCL analysis was performed using a Bayesian approach based on the work of Salo and colleagues (2013). Every series was modelled separately and series membership was predicted for each case with a leave-one-out principle. The hard-to-solve one-off offences were treated as their own series, as study three showed that the one-off offences are more like each other than like the serial cases. For ease of comparison to the aforementioned earlier studies, receiver operating characteristics analyses were carried out. The predicted probability of most likely series was plotted against the accuracy of that prediction as a function of 1–1044 added one-off offences. Linking accuracy was evaluated further by examining the sensitivity and specificity of the model using Youden's index to find the statistically optimal threshold between the two.

A secondary analysis was carried out for a case-by-case comparison of similarity using Jaccard's coefficient. Utilising a leave-one-out principle, each case was compared to a ranked list of all other homicides from most to least behaviourally similar as a function of added one-off homicides. In other words, corresponding to a situation where the investigator asks the analyst to give the most similar cases from the database to an index offence (comparative case analysis). Here the efficiency of BCL was measured by how high up on the ranked list the correctly linked cases were found.

3.5.2. Results

Overall BCL accuracy increased slightly as measured by the area under the curve (only serial *AUC* = .88, 95% *CI* = .81–.93; mixed sample *AUC* = .90, 95% *CI* = .81–.94), when adding a large number of one-offs (1 serial to 10 one-offs). Sensitivity increased while specificity decreased as a function of added one-offs. In other words, with the mixed sample it became easier to identify series correctly (increase in correct positives), while a number of series were identified erroneously (increase in false positives).

Rank ordering cases according to similarity, linkage accuracy decreased as a function of added one-off offences, confirming the main hypothesis. The proportion of cases found among the top 5 most similar cases dropped from 56% with a serial only sample to 51% with a mixed sample and among the top 20 most similar cases from 77% to 63%.

3.5.3. Conclusion

A more natural data set, which is more similar to a real crime database, introduces error to BCL in the form of increased false positives and as a smaller proportion of linked cases found in a ranked list. BCL still remains feasible, and even though rejecting false links become more difficult, correct classifications increased with the more ecologically valid mixed sample. In the investigation phase, high sensitivity is to be preferred, as no possible links will want to be missed for further scrutiny. In the trial phase, however, increased false positive linkage decisions pose a bigger problem, as thresholds in considering evidence are higher. The next natural step for BCL research is to apply the developed models on actual crime databases, and to compare BCL efficiency with and without the use of the model as an aid.

Table 2

Summary of the Research Questions, Samples, Methods, and Results of the Four Studies, and the Additional Analysis

Study	Research Question	Sample	Method	Results
Study I	-Identify dimensions in offence behaviour in serial homicide -Behaviourally link serial homicides	116 Italian serial homicides in 23 series	-Mokken scaling -Discriminant function analysis	-7 dimensions: 5 reflecting motives (money, status, revenge, rape, paraphilic rape), 2 reflecting planning (controlled and impulsive) -63% of the cases attributed to their correct series
Study II	Does knowledge of series membership affect perceived behavioural similarity in linked homicides?	3 groups of 20 participants linked 5 series of 2 homicides	-Independent variable: information on series membership (correct / incorrect / none) -Dependent variable: coded similarity	-No significant difference between the groups in coded behavioural similarity -Inter-rater reliability very high for all groups

Additional Analysis	Are homicide series coded from true crime literature more behaviourally similar than series coded from court transcripts?	23 series of homicides: 8 from true crime literature, 15 from court transcripts	-Jaccard's coefficient of similarity -Welch two-sample *t*-test	The difference in behavioural similarity between the two sources was not significant
Study III	-Explore differences between serial and one-off homicides -Can serial / non-serial be predicted?	116 serial homicides and 45 Italian hard-to-solve one-off homicides	-Cross tabulations (χ^2 and Fisher's exact test) -Logistic regression analysis	-Statistically significant differences in victim selection, forensic awareness, sexual motive, and impulsiveness -Serial / non-serial could be predicted with good accuracy
Study IV	-Is behavioural crime linking accuracy decreased as a function of added one-off homicides?	116 serial homicides and 1-1044 simulated hard-to-solve one-off homicides	-Bayesian approach for predicting series membership -Jaccard's coefficient for ranking according to similarity	-Overall crime linking accuracy increased slightly with 1 to 10 serial to one-off homicides -Sensitivity increased and specificity decreased -Series were somewhat harder to find among ranked cases when adding one-offs

4. General Discussion

The aim of the thesis was to refine the methodology of behavioural crime linking by increasing the ecological validity of the sample and formulate research questions that are relevant for giving behavioural investigative advice and questions faced in homicide investigation. We set out to add to the scarce BCL literature on homicide where most of the data up to that point had come from North America. There is surprisingly little research on crime linking accuracy with serial homicide samples, considering how much has been published on other crime types, especially serial sexual offending. Although Salfati, Labuschagne, and Sorochinski in particular have published several empirical examinations of the consistency of different aspects of SHOs behaviour, very few studies (e.g., Melnyk et al., 2011) had looked at BCL accuracy directly.

The second study was motivated by the fact that the results of the first one (and subsequently the study done on the same data by Salo et al., 2013) were so much better than previous linkage studies using the same methodology. A search for confounds that could have caused an inflation of behavioural similarity ensued. The three most likely sources of bias and error were hypothesised to be the expectancy effect of the data coders; the source of the data; and the fact that we were looking at a data pool of only serial offences. Neither the expectancy effect nor the source of the data were found to introduce significant levels of error or bias, although they could not wholly be ruled out. The last two studies were aimed to rectify the issue of previous BCL studies on homicide only being carried out on serial cases. Study III sought to add to the limited literature of comparing serial homicides to apparent one-off homicides, and to examine how well serial homicides could be distinguished from non-serial homicides. In Study IV we set out to conduct an updated BCL study, combining all the developments so far. A more ecologically valid sample of both serial and one-off homicides, using the best and most established methods available (Salo et al.'s, 2013 Bayesian reasoning, Bennell & Canter's, 2002 pairwise comparison utilising Jaccard's coefficient), and producing *AUC*s and similarity rankings for ease of comparison and better applicability of the results.

The first main finding was that BCL accuracy with Italian serial homicides was very good. Of all 116 cases, 63% could be classified to their correct series (Study I). Measured by the area under the curve, the result was slightly worse than Melnyk and colleagues (2011) study on North American serial homicides

(Study IV: *AUC* = .88, 95% *CI* = .81–93; Melnyk et al.: *AUC* = .96, 95% *CI* = .94–.98), and on par with linkage accuracy in studies on sexual offending (*N* = 8) (*AUC* = .86, 95% *CI* = .74–.98) reviewed by Fox and Farrington (2018). It must be noted that the main result of Study I (63% correctly classified cases) and Study IV (*AUC* = .88) are not directly comparable, as the *AUC* is not a measure of proportion of correct linkage decisions. Rather, the *AUC* indicates the percentage of times that a randomly selected correct case-series link will have a higher predicted probability value than a randomly selected case-series prediction where the case does not belong to the series in question. For example, an *AUC* of .88 indicates that 88% of the time randomly selected correct case-series predictions have a higher probability value than randomly selected incorrect case-series predictions. The "multivariate behavioural linking" (simultaneous input of all crime scene behaviours; Winter et al., 2013) used in Study IV improved on the performance of the "dimensional behavioural linking" (dimensions, consisting of thematically related crime scene behaviours, as linkage predictors; *ibid*.) calculated with the Mokken scaling and discriminant function analysis in Study I. In addition, the multivariate (Bayesian) method further improved accuracy, classifying 84% of the series correctly (Salo et al., 2013). This was in line with Winter and colleagues' (2013) comparison between the same methods (Mokken scaling and discriminant function analysis vs. a Bayesian method). In their study, series of rapes could also be classified more accurately with a Bayesian model (*AUC* = .84, 95% *CI* = .82–.85) than with discriminant function analysis (*AUC* = .74, 95% *CI* = .72–.76). It is worth noting, however, that Winter and colleagues' comparison is not a fair comparison of the dimensional and multivariate approach (in contrast to Tonkin et al., 2019), because the dimensions and single behaviours were being compared with different statistical methods (discriminant function analysis vs. Bayesian). Study II and the additional analysis further validate the main finding by failing to find support for a major expectancy effect or a difference as a function of data source, as measured by coded behavioural similarity. It is worth noting, however, that the chance expectation of classifying a case to its correct series was rather high (6%; Study I). This was because there were only 23 series of homicide, and the sample only consisted of serial cases.

The second main finding was that BCL with serial homicide remained viable, even when the data better reflected a real crime database with most homicides being one-off offences. When hard-to-solve one-off homicides were added at a ratio of 10 to 1 compared to serial homicides, the classification accuracy remained good. In fact, measured by the *AUC*, it increased to an excellent level

(*AUC* = .90; 95% *CI* = .81–.94). As Tonkin and colleagues (2019) point out, however, looking at the *AUC* only does not tell the whole story. Inspecting the sensitivity and specificity of the model separately revealed that while sensitivity (correct positives) increased when adding one-off offences to the mix, the specificity decreased (causing an increase in false positive predictions). The impact of the added one-off homicides on BCL accuracy, and on its practical implications become clearer when looking at the similarity rankings (Study IV). When the homicides were ordered from most to least behaviourally similar utilising a leave-one-out cross validation principle, the overall hit rate was high. In more than 80% of the cases, the most behaviourally similar case in the sample was a correct positive. The hit rate dropped only nominally, when all (*n* = 1044) one-off offences were added into the mix (85% to 83%). This would indicate an overall high behavioural consistency and distinctiveness among the SHOs, which is consistent with Melnyk and colleagues' (2011) findings. To account for some homicide series being longer, and to get a better picture of where the majority of the series was located on the ranking, the median rank was calculated for each series. Three quarters of the cases (77%) could be found in the top 20 with serial cases only (63% in the mixed sample). In practice this means that if an investigator asks for a comparative case analysis of the local database, some linked offences are likely to be missed (false negatives) even if they exist within the database. The role of this potential source of error is likely to become more pronounced when the database includes one-off offences. The combined result of the prediction of series membership (with Bayesian reasoning) and pairwise comparison of behavioural similarity (with Jaccard's similarity index) also goes to show how the two approaches complement each other and shed light on different aspects of specific BCL tasks.

Melnyk and colleagues suggest that SHOs may be more consistent (and distinctive) in their crime scene behaviour than offenders of other type of crime. The findings in the present thesis would seem to verify this hypothesis. The findings are also in line with Fox and Farrington's (2018) conclusion that person-to-person crimes (homicide and sexual offences) seem to display greater consistency and distinctiveness (as measured by the *AUC*). On the "Why?" –part of it, some speculate that SHOs more often follow inner scripts and fantasies, that are rehearsed and repeated even long before they commit their first offence (e.g., Keppel, 2000; Hazelwood & Warren, 2004). Offenders of property crime may have to adapt their offence behaviour to situational variables more (Bouhana et al., 2016), depending on what they encounter

during a burglary (e.g., a pet or an unexpected lock). In addition, the *modus operandi* may also develop and change during the course of a series as a result of improved skills, for example, for breaking into houses (*ibid*.). In summary, the initial worry of "Are these results in Study I too good to be true?" seems to subside with the comparison to Melnyk and colleagues' (2011) results and Fox and Farrington's (2018) review, which were corroborated by the results of Study IV.

A methodological shortcoming that was not tested by any study in the present thesis is that all the included cases (both serial and one-off) were solved. In other words, the present methodology does not deal with the concern of inflation of behavioural similarity, because behavioural similarity might be higher in detected and solved serial cases in the first place (Bennell & Jones, 2005; Melnyk et al., 2011). The sampling of "unsolved, but linked by DNA" (Woodhams & Labuschagne, 2012; Tonkin et al., 2012b; Woodhams et al., 2019) is yet to be tested with serial homicide.

The studies that constitute the present thesis, especially Study I and Study III, also added to the descriptive literature on serial homicide. Although apparent sexual motive was more than twice as common in the serial homicides as in the hard-to-solve one-off homicides (30% vs. 13%; Study III), more serial homicides were associated with burglary or robbery (19%) than rape (16%) (Study I). The dimensional analysis confirmed the idea that these were different kinds of serial homicide offences all together (Study I). This obviously has to do with sampling, that is, not restricting inclusion to only sexual homicide. However, it should be borne in mind that both previous research and the present studies suggest that several different motives are at play in serial homicide (e.g., Morton & Hilts, 2008). This also has implications for future research endeavours. The most obvious crossover in crime types, which has been suggested several times before (e.g., Slafati & Taylor, 2006; Davies et al., 2018), is between (serial) homicide and rape. This finding suggests it would also be pertinent to consider crossover offending between homicide and property crime. In their small qualitative analysis, Woodhams and Komarzynska (2014) found that their interviewees reported on consistent *modus operandi* behaviours over crime types from sex offences to other types of crime. Larger, quantitative analyses (such as Tonkin et al., 2011) linking different crime types from property offences to person-to-person crimes, are thus warranted.

Many features of our Italian sample were in line with earlier and later descriptive studies on serial homicide, and comparisons between serial and

non-serial homicides (e.g., Harbort & Mokros, 2001; Sturup, 2018). For example, the targeting of vulnerable victims (e.g., Salfati et al., 2015; Sturup, 2018): among the victims of the SHOs, there was an overrepresentation of prostitutes (31%) and immigrants and refugees (17%) (Study I). Compared to the hard-to-solve one-off homicides, the SHOs displayed a higher level of forensic awareness (59% vs. 20%) in that they had attempted to rid themselves of evidence (Study III). SHOs also planned their attack more often, which was apparent in that they more often brought a weapon to the crime scene (76% vs. 47%). One-off homicides, on the other hand, showed more traits of opportunistic and impulsive violence (e.g., the victim was more often hit or kicked: 38% vs. 12%; a kitchen knife or axe was used in the killing: 27% vs. 7%). In general, the non-serial homicides, even though being selected for being hard-to-solve, displayed more opportunistic and expressive violence, while the serial homicides were more planned and controlled. This would suggest the hard-to-solve one-off homicides were more like homicide (and violence) in general: expressive and impulsive, and that serial homicide truly differs qualitatively. It would also seem to confirm what Canter and colleagues (2004) stated about serial homicides and SHOs being more organised, by definition.

When examining the serial homicides on a dimensional level (rather than individual behaviours), the found dimensions (Study I) also coincided well with dynamics (e.g., victim control, expressive vs. instrumental behaviour) described to be central in earlier research (e.g., Keppel & Walter, 1999; Salfati & Bateman, 2005). The dimensions centred on themes that have been suggested to be more in the offenders' control and thus more consistent behaviour across a series of offences (*ibid*.), such as planning before and after the killing and controlling the victim. In conclusion, several similarities between the Italian serial homicides and SHOs were found to earlier (and later) samples from North America, South Africa, and Germany. Likewise, the differences found between serial homicides and one-off homicides (Study III), were well in line with earlier studies. Although serial homicides clearly differ from non-serial homicides, there is still plenty of heterogeneity and distinctiveness among SHOs and their motives and offending behaviour.

Another first for serial homicide research was using the observed differences between serial and non-serial homicides to distinguish between the two (Study III). The regression analysis found seven of the differences to be efficient for the prediction: wounds to the hands, body found outside, victim's gender, victim was a prostitute, forensic awareness, and weapon brought by the offender (all being crime features that can be known at the initial stages of

the investigation). In the context of Italian homicide investigation, these would be the crime scene behaviours to look at when making an assessment on whether the case at hand could belong to a larger series of homicides. The accuracy of the prediction was good (*AUC* = .88, 95% *CI* = .82–.94). Interestingly all seven variables that best predict whether the offence was part of a series, or a stand-alone homicide can be considered behaviours that the offender is in control of. This would further seem to corroborate the hypothesis that what makes SHOs' behaviour distinct from that of one-off offenders' is that they are characterised by greater instrumentality and consistency.

Sturup (2018) replicated the two main findings of Study III on a sample of Swedish serial and non-serial homicide offenders. The replication analysis was carried out using four of the seven factors found to contribute the most in Study III, and the discrimination accuracy was moderate (*AUC* = .76, 95% *CI* = .70–.83; *AUC* = .69, 95% *CI* = .62–.76 with all seven variables). A significant reason for the lower accuracy may have been Sturup's choice of variables. Instead of predicting category (serial or one-off homicide) using the variables that differed the most between the Swedish subsamples, Sturup did the predictive analysis using the variables that differentiated the Italian serial homicide offenders from the one-off offenders. In other words, the seven variables (Study III) are not as such globally valid when distinguishing between serial and one-off homicides. This ties in to the debate on dimensional vs. multivariate analysis of behavioural consistency. It can be detrimental for BCL accuracy to look at specific variables instead of dimensions and validating models on local data. Replication needs to be done separately on local data, to take into account cultural variations in offending behaviour and behavioural consistency (Labuschagne & Woodhams, 2012; Salfati, 2014). These variations could also have implications for data recording. This is a good example of how practice needs to inform research and research practice: the salient variables for BCL in one jurisdiction and culture may not be as salient in another. Overall, the results of the thesis show that the use of a coding scheme can be recommended to the police: coding behavioural data is relatively cheap and can be useful for homicide investigations. Data coded directly by the police could also increase the reliability of BCL research (Bennell & Jones, 2005), as it likely would decrease misinterpretations of missing information as non-occurrence of behaviour. On the other hand, coding data comprehensively and reliably is resource intensive (Davies, Imre, & Woodhams, 2021). In terms of process efficiency, future research would do well in replicating (again, locally) Salo and colleagues' (2013) finding that BCL could be done as accurately using only 15

of the full 92 variables. Interestingly (although not perhaps surprisingly), many of those 15 variables could also be interpreted as being crime features the offenders had control over (Davies, Woodhams, & Rainbow, 2018).

As Cooper and Meterko pointed out in their 2019 review, the evidence for the effect of cognitive bias influencing expertise in forensic science is robust. Davies and Woodhams (2019) point to a gap in the BCL research on this part. The effect is stronger in ambiguous tasks that rely heavily on evaluation by the expert (Dror & Murrie, 2018). Blinding procedures, omitting superfluous information that potentially are biasing, hypothesis testing, and peer evaluation are increasingly being recommended and applied procedures to counteract these biases (e.g., Riesinger et al., 2002; Kassin et al., 2013; Cooper & Meterko, 2019). It is on this point of task ambiguity that Study II failed. In their review of inter-rater reliability, Davies, Imre and Woohams (2021) found that the percentage occurrence agreement of the ViCLAS coding scheme ranged from as low as 25% in earlier studies to 52–65% in their latest, most ecologically valid coding task. For reference, a percentage occurrence agreement of 80% is usually considered acceptable. The overall inter-rater reliability for the subjects of Study III was .93. While the results are not directly comparable (e.g., the ViCLAS coding scheme has a lot more variables to code than the subjects were asked to code in Study II), the difference does provide an explanation for the result of the coding task in Study II. Thus, the conclusion that the expectancy effect does not influence data coding (or behavioural investigative advice give on crime linkage analysis in general) is premature based on the results and limitations of Study II.

While our study was designed mainly with the coding of data in mind to study the validity of the findings of Study I, the expectancy effect would certainly be hypothesised to be present in the crime linkage analysis task presented to behavioural investigative advisers. “Is the same offender responsible for these crimes?” implies an expectation on the part of the one posing the question. It is precisely these kinds of questions that are vulnerable to the effect and to confirmation bias (e.g., Dror & Murrie, 2018). This question, within the context of behavioural investigative advice on crime linkage analysis, needs further study. The finding that the uninformed group correctly linked the cases that they coded for higher behavioural similarity more often (Study II) could be interpreted so that the coders (and perhaps humans in general) intuitively know to look for behavioural similarity and that we are, at least to some extent, successful at detecting it. This puts the pressure on the BCL model to contribute something additional to a trained analyst or

behavioural investigative adviser. Future research should look at this: how well do analysts or behavioural investigative advisers fare with (and without) the BCL model as a decision support tool to aid them?

Concerning the effect of the data source and the additional analysis conducted, there does not seem to be a significant difference in behavioural similarity between court transcripts and true-crime literature as sources. Perhaps the true-crime literature strives well enough toward realistic and objective description of the cases, so as not to inflate behavioural similarity. However, the comparison and finding needs to be replicated. Bennell and Jones (2005) argue that, from a perspective of ecological validity, pre-trial investigative protocols might be the best option, as they more directly represent the information that the police operate with, in the investigative phase. Emphasis should be put on research focusing on the improvement of inter-rater reliability for the actual coding schemes that are used in practice (such as Davies et al., 2021).

Perfect linking results are not to be expected from any model. The complex dynamics that make up the resulting behaviour of any offender at the crime scene goes beyond what the offender controls. The behaviour of the victim, inconsistencies in the mental state of the offender (see e.g., Woodhams & Komarzynska, 2014), and situational variables beyond the control of either the victim or the offender (e.g., a witness unexpectedly showing up) are but a few examples. There is, however, also a deep-rooted source of uncertainty in a researcher, crime analyst or behavioural investigative adviser analysing the behaviour of an offender. The philosopher Peter Winch postulated that there is a crucial qualitative difference between the natural and the social sciences. Whereas the natural sciences are concerned with causes of correlated events, the social sciences contribute to the understanding of the world through the seeking of shared meaning (Winch, 1990). To gain an understanding of SHOs' behaviour, we need, in Winch's view, to understand the actors themselves; what they think they are doing and how they themselves conceptualize their actions. Winch is arguing that their motives are an integral part of the behaviour they display when committing their homicide. Even though we are coding carefully predefined categorizations of behaviour, and analysing the behaviour using increasingly valid statistical methods, we still need to interpret the results in order for them to make any real practical sense. It is in this interpretive step, that we cannot escape surmising the intent of the actors.

What does shared meaning mean in the case of trying to understand the actions of a serial murderer? Consider a crime scene that the offender has set

ablaze. We would code this as presence ("1") for the variable "murder scene burned". One offender might have done it for pragmatic reasons, displaying "forensic awareness", trying to destroy evidence. For another it might have a symbolic, metaphysical meaning. For a third offender it might have been an action without intent: the result of discharging a firearm near flammables. The same behaviour – burning the scene of the murder – might have varying intentions for the same offender from one offence to another. Current models still consider all instances to be the same behaviour, constituting an oversimplification of the motives at the different crime scenes. In other words, there is a similarity in their psychological fingerprint, that we subsequently use to try to determine if the two crimes are linked, that is lacking. Although there is a superficial similarity in what the offenders are doing at the different crime scenes, Winch would argue that their motives vary to a degree where we cannot claim that they are doing the same thing.

Consequently, statistical analyses of crime scene behaviour as a basis for BCL, cannot transcend the need to examine the intent of the offender. Here we come back to Rainbow's (2014) assertion, about it not being about either statistics or clinical expertise, as could easily, although mistakenly be inferred from the debate between Snook and colleagues and Dern and colleagues (2008–2009), or Canter and colleagues' (2004) criticism of the organised/ disorganised -model of serial homicide. Perhaps the most fruitful view, as advocated by Rainbow, is that statistical and clinical/expertise-based linking techniques complement rather than compete with each other. Whichever the application may be – giving behavioural investigative advice or expert evidence – one must be mindful, transparent, and forthcoming about the strengths (probability) and weaknesses (error rates) of their behavioural crime linking.

The example of the burned murder scene also makes for an additional argument for why looking at behavioural themes or domains has its advantages over the examination of individual behaviours. In other words, confirming consistency of a theme by finding several different indicators for it. Tonkin and Weeks (2021) noted that behavioural investigative advisers in New Zealand described something similar in their practice of BCL. They were not trying to link additional offences to an index offence (as Rainbow, 2014 described in the comparative case analysis scenario), but rather to a *modus operandi* they were trying to establish for a perpetrator. With valid and accurate enough methods, proactive comparative case analysis could lead to series of offences "unsolved but linked by BCL".

While Rainbow's distinction between the crime linkage analysis and comparative case analysis tasks seems valid, the questions of "Are these crimes perpetrated by the same offender [based on the psychological fingerprint of their crime scene behaviour]?" and "Can you find any similar cases to this one in your database [based on the psychological fingerprint of the crime scene behaviour]?" do share commonality. Arguably, enough so that the criticism of BCL research mainly looking at the comparative case analysis scenario and not at the crime linkage analysis -task is not wholly convincing. Granted, it may mirror the task of comparative case analysis more than crime linkage analysis, but the science and numbers supporting the argument are essentially the same. Comparing offences in the crime linkage analysis scenario requires backing up claims of consistency and distinctiveness with data on both. BCL research based on local, sufficiently reliable, and valid data can begin to answer the questions of how distinct or unique an offender's behaviour at a particular crime scene is. The data gathered and analysed in this thesis provides such a starting point for Italian homicide investigators and behavioural investigative advisers, as no centralised police database on crime scene behaviour in serial homicide exist.

4.1. Strengths

Most research on serial homicide has been conducted on North American samples. Studying an Italian sample, and producing comparable results, has contributed to the diversity of the field. The results align well with both previous and subsequent research. One of the advantages of taking a decade and a half to produce one's thesis is that it allows for taking advantage of research published both before and after one's own results. BCL research has seen a big rise in publications during the last few decades (Fox & Farrington, 2018; Davies & Woodhams, 2019). The studies in the present thesis managed to tap into this development and draw use of the progress that had been made in the field. Here and there, some incremental contributions were made. Some of our findings replicated earlier findings and some of our analyses were completely novel. Some of our results got validated through replication studies conducted by others (e.g., Sturup, 2018).

From the perspective of ecological validity and BCL in practice, the sampling choices (e.g., broad definition of serial homicide, inclusion of hard-to-solve one-off homicides) of the studies in the thesis can be seen as a strength. Choices were made to create data sets that would closely resemble the information and situation homicide investigators are faced with. This increases the applicability

of the findings, especially in the Italian context, when making assessments on the uniqueness of an offender's behaviour or when trying to determine if a homicide is part of a series. The chosen methods of analysis follow the same ideal: choices were made based on the most recent developments of the field and with maximising applicability in mind. Especially the fourth study combined what we had learned over the years: a Bayesian method for series membership prediction, calculating *AUC*s, and pairwise comparison of behavioural similarity using Jaccard's similarity coefficient. Combining the complementary methods of series membership prediction and pairwise comparison of behavioural similarity allowed us to address both the comparative case analysis- and crime linkage analysis scenarios.

4.2. Limitations

While the present studies addressed some issues concerning ecological validity, other issues could not be solved. The results of the current studies are best applied on Italian hard-to-solve and serial homicide. While attempting to come as close to a real crime database as possible, the data in the present studies were not. Law enforcement databases have their own set of challenges ranging from data coding reliability (e.g., Davies, Imre, & Woodhams, 2021) to the fact that the linkage status of crimes in a real database are unknown for a large portion of the crimes. These issues were not addressed by the studies in the current thesis. The results are also not generalizable to all homicide, or necessarily to other cultures. Local replications are needed, in order to account for possible cultural variation (Salfati, 2014).

One major limitation is that the current sample only included solved cases. This may be a confound that inflates behavioural similarity (e.g., Bennell & Jones, 2005; Melnyk et al., 2011). Although the inclusion of unsolved cases (but linked by, e.g., DNA) have been shown to have negligible effect on BCL accuracy of other crime types (e.g., Woodhams & Labuschagne, 2012; Woodhams et al., 2019), studies are missing for homicide. As results on BCL research on homicide seem to indicate that consistency and distinctiveness may be higher than for other crime types (at least as measured by BCL accuracy), the impact of unsolved cases may also be different.

The serial homicides in the current sample were from 1970–2001, while the hard-to-solve one-off homicides were from 2001–2014. Some research on homicide has indicated that there has been a general development in the quality of homicides, during the decades, in, for example, an increase in forensic awareness on the part of offenders (e.g., Lehti, 2020). Ideally the sample of non-

serial homicide would have been larger and from the same timeframe as the serial homicides, for an even more valid comparison. The inter-rate reliability of the coding scheme used could have been tested more rigorously, especially since studies have shown this to be a particular weakness (Davies, Imre, & Woodhams, 2021). Study II had some serious flaws, mainly in the stimulus material. While teaching the first author a lot about conducting experiments, and about the expectancy effect in general, the conclusions that can be drawn from the study are limited. Replication, with vignettes that better resemble the complexity of the real task, is needed.

Finally, standing on its own, the applicability of the findings of the studies in the thesis (especially Study IV) have limited applicability on the task of BCL, even in Italy. As discussed, it is unlikely a single model will be developed that does the whole task (i.e., gives a definite answer on a particular case of crime linkage analysis). The dynamics behind observed offender crime scene behaviour are too complex. Other research paradigms (e.g., Woodhams, Hollin, & Bull, 2008; Winter & Rossi, 2020) and qualitative research are needed to gain a more complete picture. Rather the results of the present studies should be put in this context and seen as a complementary piece of the puzzle that make up BCL.

4.3. Future Directions for BCL Research

Practice and research have begun converging, through efforts such as C-LINK. It is critical that this development continues (e.g., Labuschagne & Salfati, 2015). Behavioural investigative advice on BCL is as good as the science it rests upon, and the science is as valid and reliable as the data used to produce it. It is important that validation research on all the different aspects of BCL continue and expand. While the practice of BCL benefits from standardised and validated data recording procedures (such as ViCLAS), research needs to be conducted on relevant and valid data. Datasets compiled for research can only get us so far in terms of realism and ecological validity. The next step is to conduct these analyses on real databases. This requires trust and established cooperation between the academic and law enforcement communities. Research on real databases is also the next step towards automation of computerised decision support tools for BCL. The validation of these models (e.g., the ones reviewed by Fox & Farrington, 2018) needs to expand to practice, real tasks, and cases: How well do behavioural investigative advisers or crime analysts fare in their BCL with and without the aid of said decision support tools? Using these tools on real databases for comparative case analyses could detect previously

undetected series of offences. Studying how these undetected series have remained undetected, could help improve BCL and investigative methods in general.

Replications of the findings of the current study on other datasets of serial homicide (and non-serial homicide) are needed, as they are still scarce. Replication on local databases is also needed from the viewpoint of applicability (e.g., Salfati, 2014), as for example, discussed in regard to the factors that help distinguish Italian and Swedish serial and non-serial offences from each other.

Linkage studies over different crime types needs to be expanded. The comparison and overlap of sexual offending with homicides have been the subject of some studies, but not BCL so far. The present sample of Italian serial homicides shows that there are overlaps to other types of crime as well – at least robbery and burglary. Future studies could gain a deeper understanding of consistency in looking at how it persists and changes from person-to-person crime and property offences.

Davies and Woodhams' (2019) call for more research on the practice of BCL has yielded fruitful responses from practitioners and academics. In one of the latest descriptions of the practice of BCL Tonkin and Weeks (2021) point out that especially systematic research on how BCL has been done and applied in court, is needed. Also, research is needed to map out BCL practices in Italy, for a deeper and more detailed understanding of how practice and research can inform each other.

In their landmark paper on the application and evaluation of scientific knowledge and expert evidence in criminal court cases, Faigman and colleagues (2014) suggests that the starting point for a court in evaluating expert testimony on a specific case, should start with an examination of whether current evidence-based standards have been complied with. Such standards for BCL do not yet exist. The research, and thus the field of BCL, is still evolving and far from ready. The field would benefit from a large-scale "state of the art" –white paper, or an evidence-based standard, that chalks out the agreed minimum requirements for conducting BCL and presenting expert evidence on it in court.

4.4. Conclusions

Italian serial homicides can be assigned correctly as committed by the same perpetrator with good accuracy based on the crime scene behaviour of the offender. Even with a more ecologically valid sample of added hard-to-solve one-off homicides, BCL remains viable with a good classification accuracy, though classification accuracy is decreased slightly. The only study on the expectancy effect on data coding to date failed to find a strong effect. However, more ecologically valid research is needed on both the effect on data coding and on behavioural investigative advice in the form of crime linkage analysis. No significant difference in terms of coded behavioural similarity was found comparing data from court transcripts and true-crime literature. Italian serial and non-serial offences could be distinguished from each other with a good accuracy. Replication is needed on local data from jurisdictions where the models are to be applied.

Whatever the road ahead looks like for BCL, it would do best as a close-knit research-practitioner collaboration where the research informs practice and practitioners help researchers study relevant questions. Be it comparative case analysis or crime linkage analysis, proactive or reactive BCL, the validity of the advice given is only as good as the datasets it is based upon. Local databases need to be developed and maintained with continued efforts to improve their validity and reliability. As BCL methodology is refined and crime linkage analysis is carried out on increasingly ecologically valid data we gain a more precise picture of its efficiency, the accuracy of the linking decisions, and their error rates. This also has implications for the use of BCL as expert testimony because it makes it possible to get a more precise picture of the validity and reliability of the presented expert evidence. The question of admissibility is ultimately up to the courts, but the responsibility of the expert is to be transparent and forthcoming about validity and reliability and the (ever growing) evidence base of behavioural crime linking.

References

Adjorlolo, S., & Chan, H. (2014). The controversy of defining serial murder: Revisited. *Aggression and Violent Behavior*, *19*, 486–491.

Alison, L., Bennell, C., Mokros, A., & Ormerod, D. (2002). The personality paradox in offender profiling: A theoretical review of the processes involved in deriving background characteristics from crime scene actions. *Psychology, Public Policy, and Law*, *8*, 115–135.

Alison, L., & Canter, D. (1999). Professional, legal and ethical issues in offender profiling. In *Profiling in Policy and Practice*, 21–54. Routledge.

Alison, L., Goodwill, A., Almond, L., van den Heuvel, C., & Winter, J. (2010). Pragmatic solutions to offender profiling and behavioural investigative advice. *Legal and Criminological Psychology, 15*, 115–132.

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*, 43–63.

Bateman, A., & Salfati, G. (2007). An examination of behavioral consistency using individual behaviors or groups of behaviors in serial homicide. *Behavioral Sciences & the Law*, *25*, 527–544.

Beauregard, E., DeLisi, M., & Hewitt, A. (2018). Sexual murderers: Sex offender, murderer, or both? *Sexual Abuse*, *30*, 932–950.

Bennell, C. (2002). *Behavioural consistency and discrimination in serial burglary*. Unpublished doctoral dissertation, University of Liverpool.

Bennell, C., & Canter, D. (2002). Linking commercial burglaries by modus operandi: Tests using regression and ROC analysis. *Science & Justice*, *42*, 153–164.

Bennell, C., & Jones, N. (2005). Between a ROC and a hard place: A method for linking serial burglaries by modus operandi. *Journal of Investigative Psychology and Offender Profiling*, *2*, 23–41.

Bennell, C., Jones, N. J., & Melnyk, T. (2009). Addressing problems with traditional crime linking methods using receiver operating characteristic analysis. *Legal and Criminological Psychology*, *14*, 293–310.

Bennell, C., Mugford, R., Ellingwood, H., & Woodhams, J. (2014). Linking crimes using behavioural clues: Current levels of linking accuracy and strategies for moving forward. *Journal of Investigative Psychology and Offender Profiling*, *11*, 29–56.

Bouhana, N., Johnson, S. D., & Porter, M. (2016). Consistency and specificity in burglars who commit prolific residential burglary: Testing the core assumptions underpinning behavioural crime linkage. *Legal and Criminological Psychology*, *21*, 77–94.

Burrell, A., Bull, R., & Bond, J. (2012). Linking personal robbery offences using offender behaviour. *Journal of Investigative psychology and offender profiling*, *9*, 201–222.

Buss, D.M. (1998). Evolutionary psychology: the new science of the mind. Allyn & Bacon.

C-LINK (2021). Crime Linkage International NetworK (C-LINK). Available at: crimelinkage.org (accessed 30 May 2021).

Canter, D. V., Alison, L. J., Alison, E., & Wentink, N. (2004). The Organized/Disorganized Typology of Serial Murder: Myth or Model? *Psychology, Public Policy, and Law, 10*, 293–320.

Chi, H., Lin, Z., Jin, H., Xu, B., & Qi, M. (2017). A decision support system for detecting serial crimes. *Knowledge-Based Systems*, *123*, 88–101.

Chifflet, P. (2015). Questioning the validity of criminal profiling: An evidence-based approach. *Australian & New Zealand Journal of Criminology*, *48*, 238–255.

Cooper, G., & Meterko, V. (2019). Cognitive bias research in forensic science: a systematic review. *Forensic science international*, *297*, 35–46.

Davis, M., Rainbow, L., Fritzon, K., West, A., & Brooks, N. (2018). Behavioural investigative advice: A contemporary commentary on offender profiling activity. In *The Psychology of Criminal Investigation*, 203–225. Routledge.

Davis, T. (2021). The Art and Science of Offender Profiling. In *Forensic Psychology – Fact and Fiction*, 157–182. Red Globe Press.

Davies, K., Imre, H., & Woodhams, J. (2021). The utility of the Violent Crime Linkage Analysis System for conducting comparative case analysis. *Journal of Criminological Research, Policy and Practice, 7*, 77–90.

Davies, K., Tonkin, M., Bull, R., & Bond, J. (2012). The course of case linkage never did run smooth: A new investigation to tackle the behavioural changes in serial car theft. *Journal of Investigative Psychology and Offender Profiling*, *9*, 274–295.

Davies, K. & Woodhams, J. (2019). The practice of crime linkage: A review of the literature. *Journal of Investigative Psychology and Offender Profiling, 16*, 169–200.

Davies, K., Woodhams, J., & Rainbow, L. (2018). Crime linkage and sexual homicide. In *Routledge international handbook of sexual homicide studies*, 527–543. Routledge.

Davies, K., Woodhams, J., & Tonkin, M. (2019). Reviewing the use of crime linkage evidence within a legal context. In *The Routledge International Handbook of Legal and Investigative Psychology*, 197–211. Routledge.

Dern, H., Dern, C., Horn, A., & Horn, U. (2009). The fire behind the smoke: A reply to Snook and colleagues. *Criminal justice and behavior*, *36*, 1085–1090.

Douglas, J. & Munn, C. (1992). Violent crime scene analysis: Modus operandi, signature, and staging. *FBI Law Enforcement Bulletin*, *61*, 1–10.

Douglas, J., Ressler, R., Burgess, A., & Hartman, C. (1986). Criminal profiling from crime scene analysis. *Behavioral Sciences & the Law*, *4*, 401–421.

Dror, I., & Murrie, D. (2018). A hierarchy of expert performance applied to forensic psychological assessments. *Psychology, Public Policy, and Law*, *24*, 11–23.

Egger, S. (1984). Working Definition of Serial Murder and the Reduction of Linkage Blindness. *Journal of Police Science and Administration*, *12*, 348–357.

Faigman, D., Monahan, J., & Slobogin, C. (2014). Group to individual (G2i) inference in scientific expert testimony. *The University of Chicago Law Review*, *110*, 417–480.

Farrington, D., Gaffney, H., & Ttofi, M. (2017). Systematic reviews of explanatory risk factors for violence, offending, and delinquency. *Aggression and violent behavior*, *33*, 24–36.

Fesbach, S. (1964). The function of aggression and the regulation of aggressive drive. *Psychological Review*, *71*, 257–272.

Fox, B., & Farrington, D. (2018). What have we learned from offender profiling? A systematic review and meta-analysis of 40 years of research. *Psychological Bulletin*, *144*, 1247–1274.

Fox, J. & Levin, J. (1998). Multiple homicide: patterns of serial and mass murder. *Crime and Justice, 23*, 407–455.

Fridel, E., & Fox, J. (2018). Too few victims: Finding the optimal minimum victim threshold for defining serial murder. *Psychology of violence*, *8*, 505–514.

Greenall, P., & Wright, M. (2020). Stranger sexual homicide: an exploratory behavioural analysis of offender crime scene actions. *Journal of sexual aggression*, *26*, 163–177.

Grubin, D. Kelly, P., & Brunsdon, C. (2001). *Linking serious sexual assaults through behaviour: HORS 215*. Home Office Research, Development and Statistics Directorate. London.

Harbort, S. & Mokros, A. (2001). Serial murderers in Germany from 1945 to 1995. *Homicide Studies, 5*, 311–334.

Hazelwood, R., & Douglas, J. (1980). The lust murderer. *FBI Law Enforcement Bulletin*, *49*, 18–22.

Hazelwood, R, Ressler, R, Depue, R., & Douglas, J. (1995). Criminal Investigative Analysis: An Overview. In *Practical Aspects of Rape Investigation: A*

Multidisciplinary Approach, 137–149. CRC Press.

Hazelwood, R., & Warren, J. (2004). Linkage analysis: Modus operandi, ritual, and signature in serial sexual crime. *Aggression and Violent Behavior*, *9*, 307–318.

Hodge, S. (1998). *The spatial nature of serial murder: A conceptual model of disposal site location choice.* Unpublished dissertation thesis, University of Liverpool.

Horning, A., Salfati, G., & Labuschagne, G. (2015). South African serial homicide: A victim-focused behavioural typology. *Journal of Investigative Psychology and Offender Profiling*, *12*, 44–68.

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*, 42–52.

Keppel, R. (2000). Investigation of the serial offender: Linking cases through *modus operandi* and signature. In *Serial offenders; current thoughts, recent findings*, 121–133. CRC Press.

Keppel, R.D., & Walter, R. (1999). Profiling killers: a revised classification model for understanding sexual murder. *International Journal of Offender Therapy and Comparative Criminology, 43*, 417–437.

Knabe-Nicol, S., Alison, L., & Rainbow, L. (2011). The cognitive expertise of behavioural investigative advisers in the UK and Germany. *Professionalizing Offender Profiling: Forensic and Investigative Psychology in Practice,* 72–125. Routledge.

Kocsis, R. (2013). The criminal profiling reality: What is actually behind the smoke and mirrors? *Journal of Forensic Psychology Practice*, *13*, 79–91.

Kraemer, G., Lord, W., & Heilbrun, K. (2004). Comparing single and serial homicide offenses. *Behavioral Sciences and the Law, 22*, 325–343.

Labuschagne, G. (2006). The use of a linkage analysis as evidence in the conviction of the Newcastle serial murderer, South Africa. *Journal of Investigative Psychology and Offender Profiling*, *3*, 183–191.

Labuschagne, G. (2014). The Use of Linkage Analysis Evidence in Serial Offense Trials. In *Crime Linkage: Theory, Research, and Practice*, 197–224. CRC Press.

Labuschagne, G., & Gabrielle Salfati, C. (2015). An examination of serial homicide in South Africa: The practice to research link. *Journal of Investigative Psychology and Offender Profiling*, *12*, 4–17.

Lehti, M. (2020). *Henkirikoskatsaus 2020*. [*Homicide review 2020.*] Helsingin yliopisto, Kriminologian ja oikeuspolitiikan instituutti.

Melnyk, T., Bennell, C., Gauthier, D. J., & Gauthier, D. (2011). Another look at across-crime similarity coefficients for use in behavioural linkage analysis: An attempt to replicate Woodhams, Grant, and Price (2007). *Psychology, Crime & Law*, *17*, 359–380.

McKinley, A., & Petherick, W. (2021). Serial homicide in Australia, 1820–2020. *Salus Journal*, *9*, 39–50.

Morton, R. & Hilts, M. (Eds) (2008). *Serial Murder: Multi-Disciplinary Perspectives for Investigators*. Federal Bureau of Investigation, Quantico, VA.

Mouzos, J. & West, D. (2007). *An Examination of Serial Murder in Australia*. Australian Institute of Criminology. Canberra.

Oziel, S., Goodwill, A., & Beauregard, E. (2015). Variability in behavioural consistency across temporal phases in stranger sexual offences. *Journal of Police and Criminal Psychology*, *30*, 176–190.

Pakkanen, T. (2006). *Profiling hard-to-solve homicides – identifying dimensions of offending and associating them with situational variables and offender characteristics*. Unpublished master's thesis: Åbo Akademi University. Available at: https://www.researchgate.net/publication/262563269

Pakkanen, T., Santtila, P., & Bosco, D. (2014). Crime Linkage as Expert Evidence: Making a Case for the Daubert Standard.

In *Crime Linkage: Theory, Research, and Practice*, 225–249. CRC Press.

Petreca, V., Burgess, A., Stone, M., & Brucato, G. (2020). Dismemberment and Mutilation: A Data-Driven Exploration of Patterns, Motives, and Styles. *Journal of forensic sciences*, *65*, 888–896.

Rainbow L. (2014). A practitioner's Perspective: Theory, Practice, and Research. In *Crime Linkage: Theory, Research, and Practice*, 173–196. CRC Press.

Rainbow L., Gregory A., & Alison L. (2014). Behavioral Investigative Advice. In *Encyclopedia of Criminology and Criminal Justice*, 125–134. Springer.

Rainbow, L., & Gregory, A. (2011). What behavioural investigative advisers actually do. *Professionalizing offender profiling: Forensic and investigative psychology in practice*, 318–342. Taylor & Francis.

Ressler, R., Burgess, A., Douglas, J., Hartman, C., & D'Agostino, R. (1986). Sexual killers and their victims: Identifying patterns through crime scene analysis. *Journal of Interpersonal Violence*, *1*, 288–308.

Risinger, M., Saks, M., Thompson, W., & 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.

Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review, 3*, 16–20.

Salfati, G. (1998). Homicide: a behavioural analysis of crime scene actions and associated offender characteristics. Unpublished doctoral dissertation. University of Liverpool, UK.

Salfati, G. (2014). A Cross-National Comparison of Behavioral Consistency in serial homicide. In *Crime Linkage: Theory, Research, and Practice*, 141–171. CRC Press.

Salfati, G., & Bateman, A. (2005). Serial homicide: an investigation of behavioural consistency. *Journal of Investigative Psychology and Offender Profiling, 2*, 121–144.

Salfati, G., & Canter, D. (1999). Differentiating stranger murders: Profiling offender characteristics from behavioral styles. *Behavioral Sciences and the Law*, *17*, 391–406.

Salfati, G., Horning, A., Sorochinski, M., & Labuschagne, G. (2015). South African serial homicide: Consistency in victim types and crime scene actions across series. *Journal of Investigative Psychology and Offender Profiling*, *12*, 83–106.

Salfati, G., & Sorochinski, M. (2019). Match: a new approach for differentiating & linking series of sex worker homicides and sexual assaults. *International journal of offender therapy and comparative criminology*, *63*, 1794–1824.

Salfati, G., & Taylor, P. (2006). Differentiating sexual violence: A comparison of sexual homicide and rape. *Psychology, Crime & Law*, *12*, 107–125.

Salo, B., Sirén, J., Corander, J., Zappalà, A., Bosco, D., Mokros, A., & Santtila, P. (2013). Using Bayes' theorem in behavioural crime linking of serial homicide. *Legal and Criminological Psychology*, *18*, 356–370.

Santtila, P., Fritzon, K., & Tamelander, A. (2004). Linking arson incidents on the basis of crime scene behavior. *Journal of Police and Criminal Psychology*, *19*, 1–16.

Santtila, P., Junkkila, J., & Sandnabba, K. (2005). Behavioural linking of stranger rapes. *Journal of Investigative Psychology and Offender Profiling, 2*, 87–103.

Schlesinger, L., Kassen, M., Mesa, B., & Pinizzotto, A. (2010). Ritual and signature in serial sexual homicide. *Journal of the American Academy of Psychiatry and the Law Online*, *38*, 239–246.

Sheldrake, R. (1998). Experimenter effects in scientific research: How widely are they neglected? *Journal of Scientific Exploration, 12*, 73–78.

Skrapec, C. (2001). Defining serial murder: A call for a return to the original Lustmörd. *Journal of Police and Criminal Psychology*, *16*, 10–24.

Slater, C., Woodhams, J., & Hamilton-Giachritsis, C. (2015). Testing the assumptions of crime linkage with stranger sex offenses: a more

ecologically-valid study. *Journal of Police and Criminal Psychology, 30*, 261–273.

Snook, B., Cullen, R., Bennell, C., Taylor, P., & Gendreau, P. (2008). The criminal profiling illusion: What's behind the smoke and mirrors? *Criminal Justice and Behavior*, *35*, 1257–1276.

Snook, B., Taylor, P., Gendreau, P., & Bennell, C. (2009). On the need for scientific experimentation in the criminal profiling field: A reply to Dern and colleagues. *Criminal Justice and Behavior*, *36*, 1091–1094.

Sorochinski, M., & Salfati, G. (2010). The consistency of inconsistency in serial homicide: Patterns of behavioural change across series. *Journal of Investigative Psychology and Offender Profiling*, *7*, 109–136.

Sorochinski, M., & Salfati, G. (2018). A multidimensional approach to ascertaining individual differentiation and consistency in serial sexual assault: Is it time to redefine and refine? *Journal of police and criminal psychology*, *33*, 63–83.

Sorochinski, M., & Salfati, G. (2019). Sex worker homicide series: profiling the crime scene. *International journal of offender therapy and comparative criminology*, *63*, 1776–1793.

Sturup, J. (2018). Comparing serial homicides to single homicides: A study of prevalence, offender, and offence characteristics in Sweden. *Journal of Investigative Psychology and Offender Profiling*, *15*, 75–89.

Sturup, J., Rodre, S., Karlberg, D., von Vogelsang, E., Rying, M., & Caman, S. (2019). Male-on-female sexual homicides in Sweden, 1990 to 2013: a population-based controlled study of incidents, victims, and offenders. *International journal of offender therapy and comparative criminology*, *63*, 1557–1574.

Tonkin, M., Grant, T., & Bond, J. W. (2008). To link or not to link: A test of the case linkage principles using serial car theft data. *Journal of Investigative Psychology and Offender Profiling*, *5*, 59–77.

Tonkin, M., Pakkanen, T., Sirén, J., Bennell, C., Woodhams, J., Burrell, A., ... & Santtila, P. (2017). Using offender crime scene behavior to link stranger sexual assaults: A comparison of three statistical approaches. *Journal of criminal justice*, *50*, 19–28.

Tonkin, M., Santtila, P. and Bull, R. (2012a). The linking of burglary crimes using offender behaviour: testing research cross-nationally and exploring methodology. *Legal and Criminological Psychology, 17*, 276–293.

Tonkin, M., & Weeks, M. J. (2021). Crime linkage practice in New Zealand. *Journal of Criminological Research, Policy and Practice*, *7*, 63–76.

Tonkin, M., Woodhams, J., Bull, R., & Bond, J. (2012b). Behavioural case linkage with solved and unsolved crimes. *Forensic Science International*, *222*, 146–153.

Tonkin, M., Woodhams, J., Bull, R., Bond, J., & Palmer, E. (2011). Linking different types of crime using geographical and temporal proximity. *Criminal Justice and Behavior*, *38*, 1069–1088.

Turvey, B. (2012). A history of criminal profiling. *Criminal Profiling: An Introduction to Behavioral Evidence Analysis*, 1–42. Elsevier.

Turvey, B., & Freeman, J. (2016). Applied Case Linkage Analysis. In *Behavioral Evidence Analysis*, 299–335. Academic Press.

UN (2021). Victims of intentional homicide, 1990–2018. United Nations Office on Drugs and Crime, available at: https://dataunodc.un.org/content/data/homicide/homicide-rate (accessed 21 May 2021).

WHO (2021). Homicide Estimates by country, 2000–2019. World Health Organization, Global Health Observatory data repository, available at: https://apps.who.int/gho/data/view.main.VIOLENCEHOMICIDEv (accessed 21 May 2021).

Wilkinson, L. (Task Force on Statistical Interference, Board of Scientific affairs, APA) (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594–604.

Wilson, D. (2007). *Serial Killers: Hunting Britons and Their Victims, 1960 to 2006*. Waterside Press.

Wilson, M., & Daly, M. (1998). Sexual rivalry and sexual conflict: recurring themes in fatal conflicts. *Theoretical Criminology, 2*, 291–310.

Winch, P. (1990). The social studies as science. In *The idea of a social science and its relation to philosophy*, 66–94. Routledge.

Winter, J., Lemeire, J., Meganck, S., Geboers, J., Rossi, G. and Mokros, A. (2013). Comparing the predictive accuracy of case linkage methods in serious sexual assaults. *Journal of Investigative Psychology and Offender Profiling, 10*, 28–56.

Winter, J., & Rossi, G. (2020). Closer to reality? The application of sequence analysis in crime linkage. *Journal of Criminological Research, Policy and Practice*, *7*, 34–50.

Woodhams, J., & Bennell, C. (2014). Introduction: Time to consolidate and reflect. In *Crime Linkage: Theory, Research, and Practice*, 1–9. CRC Press.

Woodhams, J., Bull, R., Hollin, C. (2007a). Case linkage: identifying crimes committed by the same offender. In Criminal Profiling: International Theory, Research, and Practice, 117–133. Humana Press Inc.

Woodhams, J., Hollin, C., & Bull, R. (2007b). The psychology of linking crimes: A review of the evidence. *Legal and Criminological Psychology*, *12*, 233–249.

Woodhams, J., Hollin, C., & Bull, R. (2008). Incorporating context in linking crimes: An exploratory study of situational similarity and if-then contingencies. *Journal of Investigative Psychology and Offender Profiling*, *5*, 1–23.

Woodhams, J., & Komarzynska, K. (2014). The effect of mental disorder on crime scene behavior, its consistency, and variability. In *Crime Linkage: Theory, Research, and Practice*, 55–82. CRC Press.

Woodhams, J., & Labuschagne, G. (2012). A test of case linkage principles with solved and unsolved serial rapes. *Journal of Police and Criminal Psychology*, *27*, 85–98.

Woodhams, J., Tonkin, M., Burrell, A., Imre, H., Winter, J. M., Lam, E. K., ... & Santtila, P. (2019). Linking serial sexual offences: Moving towards an ecologically valid test of the principles of crime linkage. *Legal and criminological psychology*, *24*, 123–140.

Woodhams, J., & Toye, K. (2007). An empirical test of the assumptions of case linkage and offender profiling with serial commercial robberies. *Psychology, Public Policy, and Law*, *13*, 59.

Yaksic, E. (2018). The folly of counting bodies: Using regression to transgress the state of serial murder classification systems. *Aggression and violent behavior*, *43*, 26–32.

Yaksic, E., Allely, C., De Silva, R., Smith-Inglis, M., Konikoff, D., Ryan, K., … & Keatley, D. (2019). Detecting a decline in serial homicide: Have we banished the devil from the details? *Cogent Social Sciences*, *5*.

Yaksic, E., Bulut Allred, T., Drakulic, C., Mooney, R., De Silva, R., Geyer, P., ... & Ranger, R. (2021a). How much damage do serial homicide offenders wrought while the innocent rot in prison? A tabulation of preventable deaths as outcomes of sentinel events. *Psychology, Crime & Law*, *27*, 76–88.

Yaksic, E., Simkin, M., & Roychowdhury, V. (2021b). A statistical study of the time intervals between serial homicides. *Journal of Criminal Justice*, *73*.

Yokota, K., Fujita, G., Watanabe, K., Yoshimoto, K., & Wachi, T. (2007). Application of the behavioral investigative support system for profiling perpetrators of serial sexual assaults. *Behavioral Sciences and the Law*, *25*, 841–856.

Yokota, K., Kuraishi, H., Wachi, T., Otsuka, Y., Hirama, K., & Watanabe, K. (2017). Practice of offender profiling in Japan. *International Journal of Police Science & Management*, *19*, 187–194.

ISBN 978-952-12-4109-3