



# University of HUDDERSFIELD

## University of Huddersfield Repository

Rogerson, Michelle

The Utility of Applying Textual Analysis to Descriptions of Offender Modus Operandi for the Prevention of High Volume Crime

### Original Citation

Rogerson, Michelle (2016) The Utility of Applying Textual Analysis to Descriptions of Offender Modus Operandi for the Prevention of High Volume Crime. Doctoral thesis, University of Huddersfield.

This version is available at <http://eprints.hud.ac.uk/id/eprint/28709/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: [E.mailbox@hud.ac.uk](mailto:E.mailbox@hud.ac.uk).

<http://eprints.hud.ac.uk/>

# The Utility of Applying Textual Analysis to Descriptions of Offender Modus Operandi for the Prevention of High Volume Crime

Michelle Rogerson

A thesis submitted to the University of Huddersfield in partial fulfilment of the  
requirements for the degree of Doctor of Philosophy

The University of Huddersfield  
May 2016



## Copyright statement

- i. The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the "Copyright") and s/he has given The University of Huddersfield the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.
- ii. Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the University Library. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made. Postgraduate Research Degrees: Guidelines for the submission of work for examination.
- iii. The ownership of any patents, designs, trade marks and any and all other intellectual property rights except for the Copyright (the "Intellectual Property Rights") and any reproductions of copyright works, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.

## Acknowledgements

I would like to thank my supervisors, Professor Alex Hirschfield and Professor Paul Ekblom who have provided intellectual support, expertise and guidance throughout the PhD but also in my wider research career. I would like to thank them both for being generous with their knowledge, for their attention to detail but also for their good humour, patience and understanding.

I would like to show my appreciation to colleagues, past and present, including *everyone* at the Applied Criminology Centre and many from the wider School of Human and Health Sciences. I would like to extend particular appreciation to fellow PhD students with whom I have shared the ups and downs of the experience.

I would like to thank Monish Bhatia and Kris Christmann for their friendship and support. Jeannette Garwood has been a source of continued support over the years, and I'm grateful for the time and effort that she has given to help me work through ideas as the research progressed. Alina Haines has always been there to offer encouragement and extra motivation when mine had all but run out.

Special thanks go to my family, particularly Mum, Dad and Grandad but especially Grandma whose support has been the most 'tenacious'.

Finally, I would like to thank Osk for all of his support but mostly for his inspirational advice on getting the thesis finished.

# Abstract

Police crime information systems contain modus operandi (MO) fields which provide brief text descriptions of the circumstances surrounding crime events and the actions taken by offenders to commit them. This Thesis aims to assess the feasibility of undertaking systematic analysis of these descriptions for high volume crimes. In particular, it seeks to ask the following three questions:

- 1) Are police recorded MO data a potential source of actionable intelligence to inform crime prevention?
- 2) Can techniques drawn from computer-aided text analysis be used to identify meaningful patterns in MO data for high volume crimes?
- 3) Do conceptual frameworks add value to the analysis and interpretation of patterns in MOs?

The study focuses on a sample of theft from the person and robbery of personal property offences ( $n \sim 30,000$ ). Although existing studies have utilised similar data, they have tended to focus on crime detection and have been beset with problems of data quality.

To explore these aims, it was first necessary to conduct a thorough review of MO fields to identify the challenges they present for analysis. Problems identified include various types of error but a more prominent challenge is the inherent flexibility found within natural language, i.e. human language as opposed to languages that are artificially constructed. Based on the data review, it was possible to select, and develop, appropriate techniques of computer-aided content analysis to process the data ready for further statistical investigation. In particular, a cluster analysis successfully identified and classified groups of offences based on similarities in their MO fields.

The findings from the analysis were interpreted using two conceptual frameworks, the conjunction of criminal opportunity and crime scripts, both of which are informed by situational crime theories. The thesis identified that the benefits of these frameworks were twofold. As methods of analysis the frameworks ensure that the interpretation of results is systematic. As theoretical frameworks they provide an explicit link between patterns in the data, findings from previous literature, theories of crime causation and methods of prevention. Importantly, using the two frameworks together helps to build an improved understanding of offender's ability both to cope with and to exploit crime situations.

The thesis successfully demonstrates that MO fields contain a potential source of intelligence relevant to both practical crime prevention and research, and that it is possible to extract this information using innovative computer-aided textual analysis techniques. The research undertaken served as a pathfinding exercise developing what amounts to a replicable technique applicable to datasets from other localities and other crime types. However, the analysis process is neither fully objective nor automated. The thesis concluded that criminological frameworks are a pre-requisite to the interpretation of this intelligence although the research questioned the strict categories and hierarchies imposed by the frameworks which do not entirely reflect the flexibilities of real-life crime commission.

# Contents

Chapter 1	Introduction .....	16
1.1	Structure of the Thesis.....	18
Chapter 2	Theoretical and Conceptual Frameworks .....	20
2.1	Explaining and Preventing Crime: The Role of the Crime Situation.....	20
2.1.1	Routine activity theory.....	21
2.1.2	Crime pattern theory .....	22
2.1.3	The rational choice perspective.....	23
2.1.4	Critiques of situational crime theories.....	25
2.1.5	Situational action theory.....	31
2.1.6	Summary .....	33
2.2	Crime Causation: The Conjunction of Criminal Opportunity .....	34
2.2.1	Introduction .....	34
2.2.2	Causality and contingency .....	34
2.2.3	The Conjunction of Criminal Opportunity.....	36
2.2.4	Summary .....	46
2.3	Criminal Methods and Crime Scripts .....	47
2.3.1	Introduction .....	47
2.3.2	Cognitive scripts.....	47
2.3.3	The utility of the script framework in understanding behaviour.....	52
2.3.4	Summary .....	57
2.4	Previous Studies Analysing Offender MO .....	58
2.4.1	Linking offenders to offences.....	58
2.4.2	Comparing MOs .....	59
2.5	Conclusion.....	61
Chapter 3	Research Aims and Objectives.....	63
3.1	Data Review .....	64

3.2	Data Analysis.....	66
3.3	Interpretation of Findings.....	67
3.4	Summary.....	71
Chapter 4	Methodology.....	72
4.1	An Introduction to the Dataset.....	72
4.2	Text as Natural Language.....	75
4.3	Methods of Analysing Natural Language.....	78
4.3.1	Content analysis of natural language.....	79
4.3.2	Data and Text Mining.....	84
4.3.3	Narrative Analysis.....	86
4.4	Statistical Analysis.....	88
4.4.1	Cluster analysis.....	89
4.5	Summary.....	96
Chapter 5	Integrating and Applying Frameworks: Research Literature.....	97
5.1	Introduction.....	97
5.2	Exploring and Explaining Theft from the Person Scripts.....	101
5.2.1	Precondition: Readiness.....	101
5.2.2	Act: Preparation.....	103
5.2.3	Act: Target selection.....	105
5.2.4	Act: Transfer.....	115
5.2.5	Act: Post conditions.....	122
5.3	Summary.....	123
Chapter 6	Understanding the Limitations of Police MO Descriptions.....	126
6.1	Background: The Process of Police Crime Recording.....	126
6.1.1	Recording MO.....	127
6.2	The Quality of MO Descriptions.....	129
6.2.1	MO fields: Issues that reflect the nature of crime.....	129
6.2.2	MO fields: Issues that reflect the nature of crime recording.....	132

6.2.3	MO fields: Issues that reflect the nature of natural language.....	138
6.2.4	MO fields: Issues that reflect the nature of narrative accounts.....	139
6.3	Summary.....	140
Chapter 7	Identifying Crime Scripts Through Computer-aided Content Analysis.....	143
7.1	Introduction.....	143
7.2	Data Preparation.....	144
7.2.1	Dictionary development.....	144
7.2.2	Coding iterations.....	150
7.2.3	Summary.....	152
Chapter 8	Identifying Crime Scripts and Tracks.....	154
8.1	Distinguishing Between Protoscripts.....	154
8.2	Identifying Patterns in Methods of Offending: Distinguishing Between Scripts.....	157
8.2.1	The clusters described.....	158
Chapter 9	Reviewing and Validating the Clusters.....	220
9.1	Comparing Scripts and Tracks.....	220
9.1.1	Pre-and-post script conditions.....	220
9.1.2	Target selection.....	222
9.1.3	Transfer.....	223
9.1.4	Summary.....	227
9.2	Evaluating and Validating the Clusters.....	228
9.2.1	Comparing the cluster results with other variables.....	229
9.2.2	Comparing the stability of clusters over time.....	235
9.2.3	Determining the number of clusters.....	242
9.2.4	Comparing narrative and N-gram analysis.....	243
Chapter 10	Conclusions.....	247
10.1	MO Descriptions: Intelligence for Crime Prevention?.....	247
10.1.1	Relevant for purpose?.....	247
10.1.2	Accurate and complete accounts?.....	248

10.2	MO Descriptions: Extractable Intelligence?.....	250
10.3	Data Analysis: Identifying Similarities and Differences in Offence Commission .....	252
10.4	Interpretation: The Role of Theory in Understanding Offence Commission.....	254
10.4.1	The role of conceptual frameworks.....	254
10.4.2	Applying conceptual frameworks: Some limitations and challenges .....	257
10.4.3	Enriching offender accounts .....	259
10.5	Recommendations .....	260
10.5.1	Recommendations for police recording practices .....	260
10.5.2	Recommendations for techniques and technologies .....	261
10.6	Limitations of the Research .....	263
10.7	Original Contribution .....	264
10.8	Future Research .....	266
	References .....	270
	APPENDICES .....	286
	Appendix 1. Example of an MO field coded according to Franzosi’s semantic triplet.....	287
	Appendix 2: Mapping the script tracks identified in the MO descriptions. ....	289
	WORDCOUNT = 81,119	

## List of Tables

Table 4-1 Home Office crime classifications included in the data .....	74
Table 4-2 Methods for reporting and evaluating clusters .....	96
Table 5-1 Acts, Scenes and Sub-Goals in Theft from the Person Scripts .....	125
Table 7-1 Table showing the alternative words and incorrect spellings of ‘escape’ found in the data .....	146
Table 7-2 Examples of multi-word phrases (including negating phrases) from the data.....	147
Table 8-1 Distinguishing between robbery and theft from the person. distinctive and exclusive tokens.....	155
Table 8-2 Comparing the between-cluster variance to total variance ratio of cluster solutions.....	158
Table 8-3 Clusters identified from theft from the person and robbery of personal property MO descriptions.....	160
Table 8-4 Examples of MO descriptions assigned to cluster 1: pickpocketing at venues. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.....	163
Table 8-5 Examples of MO descriptions assigned to cluster 2: bag snatches. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores. ....	168
Table 8-6 Examples of MO descriptions assigned to cluster 3: jewellery snatches. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.....	173
Table 8-7 Examples of MO descriptions assigned to cluster 4: theft of cycles. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.....	177
Table 8-8 Examples of MO descriptions assigned to cluster 5: theft from bags (shops). Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.....	183
Table 8-9 Examples of MO descriptions assigned to cluster 6: robberies with dwellings. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.....	186
Table 8-10 Examples of MO descriptions assigned to cluster 7: demands with threats. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.....	190
Table 8-11 Examples of MO descriptions assigned to cluster 8: involving taxis. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.....	194
Table 8-12 Examples of MO descriptions assigned to cluster 9: vehicle related thefts. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.....	197

Table 8-13 Examples of MO descriptions assigned to cluster 10: robbery involving young people. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores. ....	201
Table 8-14 Examples of MO descriptions assigned to cluster 11: theft from the person in licensed premises. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores. ....	205
Table 8-15 Examples of MO descriptions assigned to cluster 12: robbery with assault. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.....	208
Table 8-16 Examples of MO descriptions assigned to cluster 13: mobile phone theft. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.....	211
Table 8-17 Examples of MO descriptions assigned to cluster 14: thefts on buses and at bus stops. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores. ....	214
Table 9-1 A cross tabulation of cluster solutions and official crime classifications.....	230
Table 9-2 Clusters in a two cluster (k=2) solution.....	232
Table 9-3 Comparing cluster solutions between the whole sample, Quarter 1 (Jan to March 2007) and Quarter 9 (Jan to March 2009) .....	240

## List of Figures

Figure 4-1 The data preparation process.....	81
Figure 4-2 A hypothetical document term matrix (highlighted cells show the presence of a token in that case).....	89
Figure 4-3 Example dendrogram showing four hierarchical clusters .....	90
Figure 4-4 An illustration of how bisecting K-means operates – Reproduced from Lancia (2012b) ....	92
Figure 4-5 Example wordcloud showing significant tokens in cluster 1. (Font size relative to chi-square score - Bottom panel top 5 words removed).....	95
Figure 5-1 Literature review in progress: a snapshot of the coding structure in NVIVO.....	98
Figure 7-1 A screenshot of a KWIC list showing all instances of <i>escort</i> in the data.....	149
Figure 8-1 Dendrogram showing the hierarchical structure of clusters.....	159
Figure 8-2 Significant tokens in cluster 1 (font size relative to chi-square score). .....	162
Figure 8-3 Significant tokens in cluster 1 – top 5 tokens removed (font size relative to chi-square score).....	162
Figure 8-4 Significant tokens in cluster 2 (font size relative to chi-square score). .....	166
Figure 8-5 Significant tokens in cluster 2. Top 5 tokens removed (font size relative to chi-square score).....	166
Figure 8-6 Significant tokens in cluster 3 (font size is relative to chi-square score).....	172
Figure 8-7 Significant tokens in cluster 3 – top 3 tokens removed (font size is relative to chi-square score).....	172
Figure 8-8 Significant tokens in cluster 4 (font size is relative to chi-square score).....	176
Figure 8-9 Significant tokens in cluster 4 – top 2 tokens removed (font size is relative to chi-square score).....	176
Figure 8-10 Tokens that were used within close proximity to <i>bike</i> in Cluster 4.....	179
Figure 8-11 Significant tokens in cluster 5. (font size is relative to chi-square score).....	181
Figure 8-12 Significant tokens in cluster 5 – top tokens removed (font size is relative to chi-square score).....	181
Figure 8-13 Significant tokens in cluster 6. (font size is relative to chi-square score).....	185

Figure 8-14 Significant tokens in cluster 6 – top 3 tokens removed (font size is relative to chi-square score).....	185
Figure 8-15 Significant tokens in cluster 7. font size is relative to chi-square score. the largest words are the most significantly ‘overused’ tokens in the cluster.....	189
Figure 8-16 Significant tokens in cluster 7. Top 3 tokens removed (font size is relative to chi-square score).....	189
Figure 8-17 Significant tokens in cluster 8. (font size is relative to chi-square score).....	193
Figure 8-18 Significant tokens in cluster 8. top 5 tokens removed (font size is relative to chi-square score).....	193
Figure 8-19 Significant tokens in cluster 9. (font size is relative to chi-square score).....	196
Figure 8-20 Significant tokens in cluster 9. with the top 5 tokens removed.....	196
Figure 8-21 Significant tokens in cluster 10 (font size is relative to chi-square score).....	200
Figure 8-22 Significant tokens in cluster 10. top 5 tokens removed .....	200
Figure 8-23 Significant tokens in cluster 11. (font size is relative to chi-square score).....	204
Figure 8-24 Significant tokens in cluster 11. top 5 tokens removed .....	204
Figure 8-25 Significant tokens in cluster 12. (font size is relative to chi-square score).....	207
Figure 8-26 Significant tokens in cluster 12. top five tokens removed.....	207
Figure 8-27 Significant tokens in cluster 13. font size is relative to chi-square score. the largest words are the most significantly ‘overused’ tokens in the cluster.....	210
Figure 8-28 Significant tokens in cluster 13. top five tokens removed.....	210
Figure 8-29 Significant tokens in cluster 14 (font size is relative to chi-square score).....	213
Figure 8-30 Significant tokens in cluster 14. top 5 tokens removed. ....	213
Figure 8-31 Mapping clusters and sub-clusters to scripts, tracks and permutations.....	217
Figure 9-1 Dendrogram highlighting a two cluster solution.....	232
Figure 9-2 Significant tokens in cluster A of a two cluster solution (font size relative to chi-square score).....	233
Figure 9-3 Significant tokens in cluster B of a two cluster solution (font size relative to chi-square score).....	234

Figure 9-4 Significant tokens in cluster B of a two cluster solution - Top three tokens removed (font size relative to chi-square score) ..... 234

## List of Abbreviations

BKM	Bisecting K-means
CAQDAS	Computer-Aided Qualitative Data Analysis
CCO	Conjunction of Criminal Opportunity
CPTED	Crime Prevention through Environmental Design
CRAVED	Concealable, Removable, Available, Valuable, Enjoyable, Disposable
EPSRC	Engineering and Physical Sciences Research Council
HMIC	Her Majesty's Inspectorate of Constabulary
HOCR	Home Office Counting Rules
HOCR	Home Office Counting Rules
KDD	Knowledge Discovery in Databases
KWIC	Key Words in Context
MO	Modus operandi
NCRS	National Crime Recording Standard
NLTK	Natural Language Took Kit
OFT	Optimal Foraging Theory
SARA	Scanning Analysis Response Assessment
SSE	Sum of squares due to error
TM	Text Mining
VIVA	Value Inertia Visibility Accessibility

### Abbreviations in the Police Data

AP	Aggrieved Party
BMT	Between Material Times
IP	Injured Party
MGE	Made Good Escape
ST	Street
SUS	Suspect
TOMV	Theft of Motor Vehicle
UK	Unknown
X2	Two persons (usually victims or offenders)

## Chapter 1 Introduction

This thesis aims to apply, and to assess the utility of, innovative techniques for the analysis of crime data, for the purposes of crime prevention. It has long been recognised that the analysis of crime problems is central to informing decisions about preventative priorities, approaches and interventions. Crime analysis is at the core of models of action in crime prevention including SARA (Scanning, *Analysis*, Response, Assessment (Clarke & Eck, 2005)) and the more detailed 5is in which analysis of crime problems is central to the generation of *intelligence* (Ekblom, 2011). Crime prevention requires 'crime side' intelligence to provide an understanding of which crimes are committed, where and when and how (Hirschfield, 2005). This can then be combined with 'intervention side' analysis into the interventions that are appropriate and effective to a given crime problem and what conditions are conducive to their success.

Analysis of crime data is far from straightforward, not least because of problems inherent in the data: only a proportion of crimes are reported, locations are often poorly recorded, and timings are often necessarily estimated. Despite these issues, impressive developments have been made in areas such as spatial and temporal analysis of crime (for examples see the academic and practice examples in Chainey & Tompson, 2008 and Hirschfield & Bowers, 2003). Developments in the analysis of 'how' crime has occurred have not been as forthcoming. This type of analysis is important because, despite the fact that crime statistics frequently depict crimes as discrete actions, it is more accurate to think of crimes as processes. Legally defined criminal *actions* are dependent upon a longer chain of events, actions and circumstances. The full crime process includes decisions and actions taken both prior to the defining event, such as pre-planning, and subsequent to it, for example, getting away from the scene without detection, or fencing stolen goods. These prior and subsequent actions can comprise a mixture of legal and illegal activities; activity that is not inherently illegal may become illegal due to its association with the intention to commit an offence (for example, going equipped, or conspiracy). A single crime process may involve the sequential chaining of a succession of illegal activities (Laycock, 2005), for example an illegal firearm can be obtained in order to commit an armed robbery which later requires money to be laundered.

The methods used by offenders to successfully complete crime processes are commonly referred to as Modus Operandi (MO). MO, approximately translated as 'mode of operation,' refers to the characteristic means through which an offender or offenders carry out an offence. MOs can be understood to include all phases of the crime from any pre-planning, use of any tools or resources and the completion of the crime including the disposal of any stolen goods and any covering of tracks. While the term *can* be used to describe all of the component parts of the crime commission

process, its *dominant usage* is to refer more specifically to the techniques employed to commit the defining crime event.

MO is commonly used to refer to the methods of an individual offender, networked group of offenders or criminal organisation and is most often regarded as intelligence to inform the investigation of crimes and the detection of suspects. Of interest to the current research is the use of MOs as intelligence for crime prevention, reducing the probability and harm caused by criminal events by intervening in their causes (Ekblom, 2011). For this task it is necessary to identify the patterns that form typical ways in which similar offences are committed. Thus, in line with Cohen, Vila and Machalek (1995) the object of analysis is the relative frequency of various MOs and their character, not the individuals committing them. There is a useful body of work which explores the process of offending using in-depth qualitative and ethnographic approaches. These studies have produced valuable insights, and show the potential to use knowledge of *how* offenders commit crime to identify crime prevention methods designed to block offender methods. Unfortunately, this type of research is difficult to repeat on a routine basis in order to regularly inform crime analysis and prevention. Although secondary data on MOs exists within police recorded crime databases, their typical free text format presents difficulties for analysis as they are not immediately amenable to statistical analysis. Analysts using these data have faced a long slog of manual coding. The time needed for this type of analysis may be afforded in academic research but is impractical for practice applications where pressure for timely findings is higher and where there may be a need to repeat the analysis on a regular basis. The author's earlier personal experience of this type of manual analysis and the suspicion that these data were not being exploited to their full potential, gave rise to the aims of this thesis.

This thesis aims to assist with the problems experienced with analysing MO data for the purpose of crime prevention using techniques drawn from natural language processing and content analysis.

The realisation of this aim is dependent on answering three broad questions:

1. *Are police recorded MO data a potential source of actionable intelligence to inform crime prevention?*
2. *Can techniques drawn from content analysis, natural language processing and statistical analysis be used to identify meaningful patterns in MO data for high volume crimes?*
3. *Do conceptual frameworks add value to the analysis and interpretation of patterns in MOs?*

These questions were explored with a sample of almost 30,000 cases within police recorded crime data. The cases all related to theft from the person and robbery of personal property. The decision

to select volume crime types was deliberate. The research is practice-orientated and therefore it was important to test the techniques and theories on the crime types that practitioners deal with most frequently and the data sources that are most readily available to them. The selection of two related crime types also meant that it was possible to test whether the methods could distinguish between different crime types. Using two very similar crime categories was felt to be a more appropriate test than using two very distinct types.

## *1.1 Structure of the Thesis*

The next chapter will introduce the key theoretical perspectives which underpin and direct the research. The chapter begins with the description and critiques of a family of perspectives that explain crime through the operation of causes in the crime situation. It will be argued that these theories help to demonstrate how an understanding of *how* crime is committed can reveal situational crime opportunities and the causal factors that create and shape criminal events. The chapter then introduces two conceptual frameworks. The first of these, namely the Conjunction of Criminal Opportunity (CCO) was designed as a method to organise available information on situational factors in order to identify the operation of proximal causes, to explore the interactions between them and hypothesise the potential impact of changes, which deliberately or otherwise, block the operation of one or more of these causes. The second framework introduced, crime scripts, has been proposed as a method of organising knowledge about crime commission in order to reveal the sequential decisions made and actions performed by offenders and to understand the causal factors within the crime situation that shape the process of crime commission. Chapter 3 outlines the research aims and objectives and Chapter 4 describes, and provides, a rationale for the data sources used and methods of analysis employed in this thesis.

Chapters 5 to 9 present empirical results from the research. This begins in Chapter 5 with the application of conceptual frameworks to existing research on theft from the person and robbery of personal property. This application of frameworks aims to examine their use in organising and interpreting knowledge about crime processes but also consolidates existing knowledge on these crime types which is used to interpret findings from the analysis of police data. Chapter 6 reports the results of a review of the nature of police recorded MO data fields and highlights how problems that stem from the nature of crime, the nature of crime recording and the nature of free text fields pose problems for the analysis of these data. Chapter 7 describes the processes through which MO data were cleaned and standardised to allow for further statistical analysis. Chapter 8 presents this analysis of MO fields in police recorded crime data and aims to test the utility of both technical

methods of analysis and conceptual approaches to organising and interpreting these data. Chapter 9 reviews the findings of the analysis and makes an assessment of the validity of the results.

The final chapter of the thesis reflects upon the findings in the light of the research aims and objectives. The chapter also outlines the contributions of the study including methodological, theoretical and practical insights. The chapter also acknowledges the research limitations and identifies avenues for future research.

## Chapter 2 Theoretical and Conceptual Frameworks

This chapter introduces and critiques a family of criminological perspectives that place the crime situation central to explanations of criminal events and efforts to prevent crime. It will highlight the insights about preventing crime that have been gained from considering the crime situation while also paying attention to the consequent pitfalls of this approach, or at least of the manner in which this approach is often applied. The chapter will present two theoretical frameworks, the Conjunction of Criminal Opportunity (CCO) and crime scripts, both of which are closely related to situational explanations of crime and which aim to assist in the organisation and interpretation of data about the crime commission process. The theories and frameworks described in this chapter will provide the foundations for the analysis of the theft from the person literature in Chapter 5 and the analysis of police recorded crime in Chapters 7, 8 and 9. Section 2.4 of this chapter explores previous studies which have used MO data as a source intelligence about crime, this section reviews the methods that have been adopted, highlights some of the challenges of using MO data and identifies gaps in the questions being asked using these data sources.

### *2.1 Explaining and Preventing Crime: The Role of the Crime Situation*

Wortley (2011) defines a situation as a set of circumstances, at a particular point in time usually surrounding a location or place. Situations include, not only, the physical elements of the setting but also, the behaviour that takes place there. A situation is more than a geographical location: it is a system in which the elements of actor, behaviour and setting interconnect to form a complex whole. Criminological theories that focus on the situation have been grouped under a variety of labels including 'Environmental Criminology', 'Opportunity Theories' (Felson & Clarke, 1998) and 'Criminologies of Everyday Life,' (Garland, 2000). For consistency from here onwards this set of theories will be referred to as 'situational crime theories<sup>1</sup>.' These theories stand in contrast to 'predestined actor' (Hopkins Burke, 2009) theories which had dominated explanations of crime for over a century. The latter seek to explain individuals' motivations for offending through the existence of psychological, social and biological factors which predispose individuals to commit crime. The distinguishing feature of situational crime theories is the aim to explain, not the presence of criminal motivation within individuals, but the occurrence of crime events within situations. From this perspective the situation takes a central role in explaining the occurrence of crime events.

---

<sup>1</sup> Environmental Criminology is increasingly being used to refer to the study of crimes that harm the natural environment.

Four central tenets of situational crime theories separate them from traditional dispositional criminology. Firstly, immediate opportunities (as distinct from life opportunities (Cloward & Ohlin, 1960)) have a causal role in the creation of crime events. Secondly, offenders are rational actors (or at least as rational as the general population). Offenders are not viewed as a distinct class of individual, the theories therefore do not focus on dispositional explanations for offending behaviour. Thirdly, crime can and should be prevented by removing opportunities for crime, rather than tackling the motivation to offend (criminality). Finally (as a consequence of the other three tenets), because opportunities are crime specific, crime prevention interventions should be tailored towards specific crime problems. Three perspectives, each of which will be outlined further below, form the core of this branch of criminology: routine activity theory (Cohen & Felson, 1979, Felson & Cohen, 1980), crime pattern theory (Brantingham & Brantingham, 1993) and the rational choice perspective (Cornish and Clarke, 1986). The first two of these theoretical approaches, can be regarded as ecological theories; that is, they explain regularities in crime through the relationships between different entities in a system, although they handle the relationship between offenders and criminal opportunity at a different level of abstraction (Clarke, 2005). Routine activity theory (originally) presented macro level explanations of the supply of criminal opportunities in society at regional and national levels. In its original form, these amounted to generic cross-society explanations but routine activities has increasingly been used to provide specific localised/micro explanations. Crime pattern theory places routine activity theory into a localised context by modelling the concentrations in crime that are created by the regular patterns of everyday routines. The rational choice perspective is a psychological perspective that presents a view of opportunities from the offenders' perspective and describes the decision making processes through which offenders assess and choose whether to act on the risks, effort and reward of crime commission in the immediate situation.

The compelling feature of situational crime theories is that they readily allow for the identification and blocking of causes creating opportunities for crime. In doing so, situational crime prevention approaches have demonstrated the potential to successfully reduce crime in a very short period of time (Clarke, 1997).

### 2.1.1 Routine activity theory

Routine activity theory proposes that for a crime to occur three elements are required, namely 1) a likely offender, 2) a suitable target and 3) an absence of a capable guardian (a person with the ability to protect the crime target). Cohen and Felson (Cohen & Felson, 1979; Felson & Cohen, 1980) argue that these three elements are necessary conditions for a crime event to occur. Two conditions have later been added to this 'chemistry for crime'. Firstly, influenced by Hirschi's (1969) control theory,

Felson (1995) added the necessary condition of the absence of an 'intimate handler'; a person with direct personal influence over the offender. Eck (1994 in Eck & Weisburd, 1995) further extended routine activity theory to add the role of 'place managers', that is individuals who can control crime by managing and controlling specific locations. Routine activity theory seeks to demonstrate that social change alters the balance between these necessary conditions that encourage or discourage crime leading to fluctuations in available opportunities for crime. As such, crime becomes a 'consequence of the products we design and use, and the business and social arrangements to which we subscribe' (Pease, 1997, p.233). Cohen and Felson provided a number of examples of social changes which alter the supply of crime opportunities, they included: an increase in the availability of expensive consumer goods constituting an increased range of suitable crime targets, an increase in activities conducted away from home, leaving dwellings 'unguarded' for longer periods, and the baby boom, providing an increase in young males of offending age (note an increase in the overall number of young males, not an increase in their degree of motivation). Felson (1988) argued that social changes have also affected the availability and influence of intimate handlers, through, for example, an increase in the size of the population who are single parent families, potentially reducing the number of individuals whose relationship is close enough to a young person to be able to exert an influence. Cohen and Felson's focus was largely on the informal processes which have an unintended causal influence on crime at a macro level. However, this perspective can and has been used to inform crime prevention that aims to deliberately manipulate everyday arrangements. Further, many of the examples of the application of routine activity theory have been on a micro rather than macro scale (see for example: Holt & Spencer, 2005; Marcum, Higgins, & Ricketts, 2010; Petrosino & Brensilber, 2003; Wang, 2002). Importantly, routine activity theory purports that changes to the environment can prevent crime without the need to intervene in individuals' motivation to commit it. While the motivated offender is a necessary condition in the 'chemistry of crime,' explanations for how such motivations are formed and how to stop them forming are not the object of concern (Felson, 1986).

### 2.1.2 Crime pattern theory

Crime pattern theory (Brantingham & Brantingham, 1993) seeks to explain the uneven distribution of crime opportunities at the local (e.g. city wide) level. This model highlights the spatial and temporal regularities inherent in routine activities (Groff & McEwen, 2008). Everyday routines are centred on popular locations or nodes such as home, school, work and leisure. These nodes are connected by paths across the environment. Routine journeys along these paths provide the meso level settings for the convergence of motivated offender and suitable target. Paths are themselves

shaped by factors such as local topography and urban form (Felson, 2009). The space covered by those paths that are most travelled in daily life provide an individual with an 'awareness space': for offenders this space allows them to become familiar with potential opportunities for crime. Thus the clustering of crime is governed by more than simply the distribution of targets (Hepenstein and Johnson, 2010). Crime will cluster around targets that are proximal to nodes and/or are made accessible by pathways. Certain types of crimes are predicted to cluster around the edges or boundaries of awareness spaces; although still within the offenders' awareness space, their distance away from regular locations means offenders are less likely to be known or recognised. Felson and Clarke (1998) argue that the potential for those who do not know each other to come together at edges increases the likelihood of crimes such as robbery and racial attacks. Crime pattern theory distinguishes between nodes which are crime generators, crime attractors and crime detractors. Crime generators are nodes to which offenders will travel in the course of their legitimate routine activities (for example, shopping centres or transport hubs), but once there, will take advantage of any crime opportunities encountered. Crime attractors are nodes which have a reputation amongst offenders for their available crime opportunities, consequently, these nodes act as 'magnets' to offenders who will travel there specifically because of the crime opportunities available (for example, crack houses, red light districts, entertainment districts). Crime detractors are nodes which deter and/or discourage<sup>2</sup> people from offending, relevant inhibitory factors will vary by crime type but might include locations with good natural surveillance or strong cohesive communities. Crime neutral areas tend to be low crime areas, they do not attract potential targets or offenders and the crime controls in place are regarded as adequate. Crime pattern theory, therefore, aims to map out the urban "patchwork of crime generators, crime attractors, crime detractors and neutral areas" (Felson & Clarke, 1998, p.15).

### 2.1.3 The rational choice perspective

The rational choice perspective provides a micro-level explanation of the decision making processes undertaken by offenders to assess risk, effort and reward once suitable conditions for crime (a crime opportunity) have been recognised. This has been extended by Wortley (2008) to consider situational triggers or precipitators that can, for example, prompt or provoke the motivation to exploit a crime opportunity. The rational choice perspective posits that individuals adopt a calculus in decision making to weigh up the cost and rewards of criminal behaviour (Clarke & Cornish, 1985;

---

<sup>2</sup> Felson (1995) distinguishes between factors that deter, those that increase perceptions of risk, and those that discourage, decrease the perceived reward.

Cornish & Clarke, 1986) Decision making takes place at two psychological levels, the decision to engage in criminal behaviour, generally (the 'criminal involvement' decision) and the decision to commit a specific offence (the 'crime event' decision). The criminal involvement decision entails an individual reaching the strategic conclusion that the benefits of involvement in crime outweigh the estimated risk and effort. Clarke and Cornish suggested that this decision will be made in relation to a specific range of offence types, placing the individual in a state of readiness to commit these crimes. For Clarke and Cornish, this decision 'implies that the individual has actually contemplated this form of crime as a solution to his needs and has decided that under the right circumstances he would commit the offence' (1985, p.167). The readiness to offend decision will be revisited and the set of crimes regarded as 'committable' regularly reviewed. Support for this criminal involvement decision can be found in research on a diverse range of crimes including burglary (Wright & Decker, 1994), robbery (Wright & Decker, 1997) and identify theft (Copes & Vieraitis, 2009a).

The criminal involvement decision is converted into the commission of a crime, at the time and place when the presence of the offender converges with an opportunity to commit a crime. At this point, the individual will engage in the second, tactical, level of decision making to determine the costs and benefits of committing that particular crime event. The offender's aim, at both levels of decision making, is to increase the potential rewards gained from the crime while reducing the effort and risks associated with offending.

From the rational choice perspective, the decision to offend is instrumental, goal-directed and structured by a rational assessment of the costs and benefits of the current criminal opportunity, 'The starting point of rational choice theory is that offenders seek to advantage themselves by their criminal behaviour. This entails making decisions amongst alternatives. These decisions are rational within the constraints of time, ability and the availability of relevant information' (Pease 2006, p.339). The decision to offend and the assessment of costs and benefits are influenced by salient elements of a situation. Cornish and Clarke refer to these conditions as choice structuring properties (Cornish & Clarke, 1987, 1994, 2002) and argue that these properties shape the offender's assessment of action alternatives at the time and place when the presence of the offender converges with a suitable opportunity to commit the crime. The action alternatives made available by choice structuring properties are dependent on the offender's distinct attributes although generalisations can be made for groups of similar offenders and offences (Hamilton-Smith, 2002).

The focus of the rational choice perspective is, consequently, not the social or psychological conditions that result in a predisposition to offend. These factors may guide individuals to consider crime as a course of action but importantly, these factors predispose rather than rigidly

predetermine and offenders should ultimately be regarded as having a choice about their involvement in crime (Clarke & Cornish, 1985). Within the rational choice perspective, the predisposition to offend is taken as given, with no attempt made to explain it or the supply of motivated offenders. Regularities in the nature of criminal events, the phenomenon of interest in this thesis, are not to be explained by changes in the motivations of offenders but by changes to situations<sup>3</sup>. For those tasked with understanding and preventing crime, the rational choice perspective offers the promise of systematic, explainable patterns of criminal behaviour and the potential to anticipate the routes through which offenders will adapt as a result of changes to their environment, including those that derive from security interventions.

There is a tendency for work stemming from the rational choice perspective to view the offender as passive receiver of objective opportunities, although, increasingly, there are attempts to introduce to the perspective some consideration of the processes through which offenders come to recognise, interact with and actively create opportunities (Wortley, 1997; Ekblom, 2007).

#### 2.1.4 Critiques of situational crime theories

##### *Rationality and decision making*

A true rational decision is one which selects the action with the largest expected utility after comparison of the probabilities and possible outcomes of all action options. Most writers working within the rational choice perspective fully acknowledge that this 'abstract rationality' – systematically conducted, free of emotional bias and social influences – is 'radically incomplete as a picture of any person we know' (Moore, 1984, p. 112 in Norrie, 1986). The perspective accepts that the rationality of offender decision making is limited or 'bounded' by factors including cognitive ability, the limited information available on which to base a decision at any given time and the emotional and social constraints on behaviour. Research has shown that the decision making of humans in general, and offenders in particular, lacks the systematic elements of abstract rationality. Rationality is therefore limited by factors including the following:

- Offenders make poor estimates of outcome probabilities (Carroll & Weaver, 1986; Coupe & Blake, 2006; Shover & Copes, 2010);

---

<sup>3</sup> Although changes to offenders' skills, knowledge and other capacities can themselves influence the dynamics of the crime situation and help to explain changes in crime.

- Decision making is always conducted under uncertainty, it is not possible to know with certainty what the optimal action would be (Wortley, 2014);
- Offenders do not fully consider the consequences of offending (Bennett & Wright, 1984; Walsh, 1986);
- Offenders do not systematically assess action alternatives, instead they tend to make standing decisions based on what has worked in the past (Weaver & Carroll, 2008);
- Offender decisions are biased by emotions, by the influence of alcohol and drugs and by the immediate pressures of their need for cash (Bennett & Wright, 1984; Wright & Decker, 1997).

The rational choice perspective acknowledges that individuals do not always select the optimum strategy and argues that the yardstick applied should be whether, from the actor's point of view, the decision was the optimal way to meet a goal. Cornish and Clarke argue that it is sufficient that actors make choices between alternatives to meet their goals and that the restrictions on human decision making mean that "decisions are likely to produce 'satisficing' – that is, satisfactory and sufficient – outcomes rather than optimal ones most of the time" (Cornish & Clarke, 2008, p. 25).

However, a more fundamental critique of the rational choice perspective stems from the view that many crimes are not driven by instrumental, goal-orientated behaviour (Hayward, 2007; Tunnell, 1990). Hayward argues that most criminal behaviour is not influenced by future rewards and not deterred by anything other than immediate and certain disadvantages. He asserts that the *majority* of human decision making is directed by normative and affective factors and that rational and logical analysis features rarely, if ever (Hayward, 2007). Even where decisions are rational, they are infused with normative and affective bias. He argues that within the rational choice perspective 'the human purposes and existential meaning of crime are thus literally banned from the equation' (p.234).

While this charge of total omission is not strictly true there are increasing calls for the perspective to more fully consider the role of emotions, values and subjective interpretations.

To fully understand the decisions and subsequent behaviour of offenders, it is essential to examine the social worlds which offenders inhabit, the values they regard as important and exactly what in a situation they find rewarding. A key example of this line of enquiry is the body of research concerning the influence of offenders' commitment to 'street life' on their assessment of the benefits and consequences of criminality (Shover & Honaker, 1992; Cherbonneau, & Copes, 2005; Jacobs, Topalli & Wright, 2003). A 'life as a party' sub-culture encapsulates the values of hedonism, neglect of responsibility, resentment of authority and disdain for conventional behaviour. These values, which importantly are enjoyed and actively played out in the company of others (Shover &

Honaker, 1992), are all served by the commission of spontaneous, reckless and seemingly 'irrational' offences. Not only are criminal values instilled within these street cultures, the lifestyle of drug use and partying creates short term financial pressures providing the incentive to commit crime while at the same time leading individuals into situations that are conducive to offending (Wright & Decker, 1997). In addition to shaping values, creating motivation and providing opportunity, the lifestyle serves to erode the resources that can help the offender avoid crime with the consequence of blocking alternative options making them appear invisible to offenders (Shover & Honaker, 1992). However, Shover and Honaker do not claim to 'falsify' the rational choice perspective; rather they call for it to be refined to enable the risk assessments made by offenders to be understood in the light of the personal and social contexts in which decisions are made. In agreement with this view, authors including Ekblom (2007) and Hochstetler (2001) call for a move away from an over mechanised and over-simplified understanding of decision making. To provide enriched accounts of the offender it is necessary to give consideration to their motivations for offending. Although crime involvement decisions may be distinguishable from crime event decisions, they are not inseparable. Coupe and Blake (2006) argue the characteristics and motivations of offenders are 'indispensable' to understanding the manner in which offences are carried out.

It is clear that in the rational choice perspective, a 'rational' choice is not necessarily a correct or well considered choice. It may not involve a full consideration of all alternatives, and may not produce the optimal solution for the offender. A rational choice is regarded as one taken in order to meet the required ends, as defined from the actor's perspective following a cognitive process. At the heart of rational choice perspective, and situational crime theories more broadly, lies the notion that actors assess the action alternatives within a setting based on their interpretation of the situation, affordance and capacity. In deciding which actions to take, offenders (and other actors in the crime situation) can be regarded as caused agents, they have the capacity to shape the situation they are operating within but they are also shaped by it (Ekblom, 2010). A central question for situational crime theories is whether, given due consideration to both the causal factors and agency, it is still possible to identify regular patterns in behaviour and understand how these patterns are shaped by factors in the situation?

### *Crime as opportunity*

The view that the changing occurrence and nature of criminal events can be explained and prevented, not through changes in the perpetrators or their motivations, but by changes to the environments in which offences take place, is not universally accepted within criminology and situational theories of crime certainly have their critics.

Sutton (2012) reflects many of these criticisms with his argument that situational crime theories provide no more than an 'excellent description of the data' and regards these perspectives as truisms not theories of crime. The three key props of Sutton's critique are that 1) a scientific explanation should not vary and, according to Sutton, situational crime theories allow the three key elements of opportunity to vary; 2) that this perspective conflates the data to be explained with the theories that seek to explain them; and 3) that the opportunities central to these theories are not objective phenomena. These criticisms will be discussed in turn. Firstly, the argument that an explanation should not vary, should not preclude basing an explanation on factors that *can* vary. Sutton takes issue with what he views as the elasticity of explanations emerging from situational theories of crime. Thus, if a crime is committed in the presence of an apparently capable guardian, it can be argued that the level of offender motivation must have been extremely high. If the target was not very attractive or vulnerable, perhaps the offender was particularly motivated and so on. Sutton, therefore, claims that situational crime theories are unfalsifiable<sup>4</sup>. However, there is an important distinction between the *elements* of the explanation, which are constant, and the *state* of those elements, which can vary. It is not unusual for a scientific explanation to be based on a range of factors which must be present in given ratios for an outcome to occur. The rational offender decision agenda can be regarded as 'interchangeable currency' (Ekblom & Sidebottom, 2008): offenders adapt, so if crime is made more risky for them they may respond, if the reward is high enough, by taking more effort. The challenge is that it is difficult to directly observe or manipulate these factors and we do not have records or data relating to un-committed crimes i.e. instances when the combination of required factors was too low. We do not have the evidence to inspect whether the exact same conditions have existed somewhere else but did not result in a crime. Sutton argues that there are insufficient attempts from within the discipline to find disconfirming evidence. This latter point is valid, although, there have been a limited number of attempts to explore crime opportunities not exploited by offenders (Mativat & Tremblay, 1997).<sup>5</sup>

As noted above, situational crime theories focus on the role of opportunity in crime causation. Many writers from within situational criminology argue that opportunity is the most important cause of crime (Felson & Clarke, 1998; Felson, 2008; Laycock, 2005; Tilley & Laycock, 2002; Wortley, 2011). Sutton argues that these theories are 'truisms' offering little more than an after-the-event

---

<sup>4</sup> Sutton's argument that opportunity based crime theories are unfalsifiable relates to the notion of a just-so-story, a derogatory term originating in Rudyard Kipling's children's tales used to describe unfalsifiable explanations such as 'how the elephant got its trunk', and nowadays applied by critics of 'adaptationism' in evolutionary studies.

<sup>5</sup> An interesting way to shed light on this would be to identifying predictions of future crime trends that proved inaccurate.

description of a successfully completed crime event. He states that situational criminology confuses its data with the phenomena to be explained. A similar criticism is raised by Wikström (2006a), who argues that criminological explanation frequently confuses symptoms with causes. This is perhaps a fair observation and it would be more appropriate and less derogatory to refer to these perspectives as models rather than truisms. Clarke (2012) reminds us that routine activity theory and the rational choice perspective were originally conceived as perspectives and not theories (and this remains his preference). Models have a long established and valued role in science. A model organises knowledge about a given system, this can include providing a representation of sets of relationships. Tukey (1980) argued that data analysis needs to involve both exploration and theory building; the two being intimately connected. A model is descriptive, but has the aim of aiding understanding and can guide further exploratory and confirmatory research.<sup>6</sup> Models can help to organise data, for example, situational crime models have highlighted that some environments consistently experience more crime than others (Bernasco & Block, 2011). Environments identified as being associated with high crime levels include, for example, areas with high levels of deprivation (Hirschfield & Bowers, 1997) and areas with a specific mix of licensed premises (Newton, 2014). Studies such as these could be regarded as 'descriptions of data' but they do more than this, they identify regularities and identify relationships that in conjunction with theory can help to explain the patterns. There is a predictive quality to models which enables the inductive prediction of the kind of crime, classes of places, targets, and types of crime that might occur (examples include Johnson & Bowers, 2004; Kershaw & Tseloni, 2005). Models have a clear role in scientific explanation. This includes the provision of frameworks on which causal explanations can be built and theories developed and tested.

Although this criticism of opportunity based crime theories is not entirely justified, Sutton is perhaps correct that the impression given through the language used by some writers in the field is unhelpful. Some situational crime theories may be better labelled as models. Bouhana (2013) argues that opportunity theories are axiomatic statements rather than explanations. She argues that opportunities only cause crime through the dynamic interaction with offender identification and perception. Opportunities are, therefore, better described as ecological. Bouhana argues that

---

<sup>6</sup> Sutton (2011) argues that 'Saying that opportunity is a cause of crime is like saying hydrogen and oxygen are the cause of water.' However, a model of water would identify that hydrogen and oxygen are essential components of water. An explanation of the production of water would require an explanation of the mechanisms through which energy rearranges the atoms so that they fit together. Physics has yet to fully explain this process, but identifying the essential chemical elements is an essential foundation for this explanation.

situational crime theories do not attempt to *explain* how offender decision making takes place (rather than what decisions are made), therefore, the full process of crime causation is not explained. Bouhana further adds that this is despite the availability of evidence and theory from potentially informative disciplines including behavioural, cognitive and neurosciences.

In the third element of Sutton's critique he stresses that crime opportunities<sup>7</sup> cannot be 'identified' as entities that are objective and certain in their own right and argues that what routine activity theory describes as crime opportunity, can only exist after an offender has recognised, interpreted and decided to act upon favourable conditions and a crime is successfully completed. Instead, Sutton argues, criminology should seek to explain the process through which subjective interpretation of favourable conditions is converted by offenders into a crime opportunity. Elements of this critique appear justified, and are recognised by some proponents of situational crime theories. Ekblom (2010) has argued that the rational choice perspective has glossed over the distinctions between perceived, attempted and realised opportunity for the sake of simplicity. Ekblom (2010), Bouhana (2013) and Wikström and Treiber (2009) regard opportunity as ecological; created by the dynamic interaction of the individual and the situation. For Ekblom, opportunity is the conjunction of causes that are sufficient for a crime to occur, both those residing with the offender and those in the situation. Opportunity is co-determined by the presence of favourable conditions in the situation. A crime opportunity does not materialise unless it matches the offender's available capacity, resources and purpose (an open window is only an opportunity for someone with a ladder, agility and courage (Ekblom, 2010)). Thus although descriptions of situational crime theories may describe the offender as 'converging' with an opportunity for crime, this is misleading as the opportunity cannot be created in the offender's absence. Conditions that are conducive to crime may not be recognised or exploited if they do not match the actor's purpose or resources. The call for situational theories to acknowledge factors that lead to individual differences in the perception of crime opportunity is also made by Ekblom (2007), Garwood (2009), Hochstetler, Copes and DeLisi (2002) and Wikström (2014). Thus, factors including experience, emotions, 'nerves' and prejudices can all influence whether an opportunity is recognised and how it is interpreted and acted upon. An approach that aims to unify explanations of offender based and situational theories of crime by exploring how

---

<sup>7</sup> Sutton replaces 'opportunities' with criminal explanation based on contingencies. Offenders benefit or are hindered by a series of accidents, or contingencies. The concept of contingencies does not provide scope to consider the offender as conscious free-thinking agent or consider how their beliefs, desires and reasoning will interact with the situation, referred to as adopting the intentional stance (Dennett, 1996). The concept of contingency also ignores the means through which offenders can actively see, boost or create conditions conducive to crime commission.

situational opportunities interact with motivational and moral factors is situational action theory, the following section will consider this theory in more detail.

### 2.1.5 Situational action theory

The sections above have highlighted the important role of the situation in shaping offenders decisions, but have shown that there are a number of limitations to common approaches to understanding the role of the situation. A development in situational theories which helps to address some of these problems is situational action theory.

Wikström (2006a) argues that goals, experience, emotions and morals shape actors' perceptions of the action alternatives available in a setting, and that these perceptions are central to understanding the actions that individuals take (or do not take). In common with the rational choice perspective, Wikström's situational action theory is concerned with individuals' assessment of action alternatives. However, of more central import within situational action theory are the mechanisms through which individuals perceive and identify the range of action alternatives that are available to them, 'The crucial question in explaining an act of crime is to explicate why an individual comes to see an act of moral rule breaking as an action alternative and why he chooses to act upon it in order to realise a particular motivation,' (p.94). This addresses the criticism from Sutton and others that situational theories treat crime opportunities as objective entities viewed by all actors in the same way.

Wikström explains that motivations are formed from individuals' desires and their beliefs regarding whether and how these desires can be satisfied. Consequently, an individual's desires, beliefs, knowledge, and morality, shape which action alternatives they identify and act upon within a given setting. Thus, while settings exist independently of individuals, actors interpret these settings and the opportunities within them subjectively. At the heart of situational action theory there is an acknowledgement of interactions between the real world and those who are observing and experiencing it. A crime opportunity is irrelevant for an unmotivated individual, therefore, the risks, effort and reward associated with it will not be assessed and the choices of action alternatives associated with the opportunity will not be deliberated. Alternatively, whether and how a motivated individual perceives the action alternatives that are associated with a crime opportunity will depend upon individual differences including capability, knowledge, experience and morality. The perceptions and choices of the actor, therefore, form an integral bridge between the individual and their motivations and the setting and the opportunities it presents. 'It is the interaction between the individual and the setting which determines the individual's course of actions' (p.87). The acknowledgement of interaction between the actor and setting is often lacking in explanations

stemming from the rational choice perspective. Cornish and Clarke (1994) do argue that actors evaluate the key features of a setting, referred to as choice structuring properties, before deciding their course of action. However, the concept of choice structuring properties is insufficient to explain the interaction between-actor and setting. Mativat and Tremblay (1997) argue that the definition of choice structuring properties is unclear with Cornish and Clarke initially referring to them both as factors perceived by the offender (Cornish & Clarke, 1987, p.935) but later widening the definition to include factors of which offenders may not be aware (941). Where researchers have produced lists of choice structuring properties for various crime types some have included factors relating to the offender, others list factors relation to the setting and others a combination of the two. Cornish argued that choice structuring properties should be understood as the 'interface between offence scripts and offender repertoires' (Cornish, 1994, p.183) and that offenders compared their own motives, needs, skills preferences and values with the properties presented by the situation. The use of choice structuring properties risks depicting the situation as a fixed entity which the offender can react to but not manipulate. Situational action theory offers some improvements to this argument by moving away from portrayal of factors in the environment as objective attributes and describes, more explicitly the specific mechanisms that operate at the interface between the individual and the situation.

The moral perspective of the actor is central to situational action theory. Situational perspectives are frequently criticised for ignoring morals, or at least acting on the assumption that all actors in the environment are immoral (Hayward, 2007), but this is not strictly true. Within the rational choice perspective, the offenders' morals will influence the selection of offences which form part of the 'stage-one' involvement in crime decision. However, the perspective pays less attention to the interaction between morals and favourable conditions, within the specific setting, in influencing the second stage, offence commission decision. In the proximal decision, morals can influence what the offender finds rewarding, or (self-)punishing from situational action theory, individuals' understanding of a given situation includes their moral understandings of what rules apply to the setting and whether their own values support or reject the adherence to those rules. Individuals who fully accept the moral context of a setting are unlikely to perceive rule breaking as a possible action alternative, even if they have desires that could be satisfied via criminal actions.

In earlier writings that show an affinity with situational action theory, Gibson (1979) introduced the concept of 'affordance' to consider how interactions between actors and their environments govern the range of action possibilities perceived by an individual. Affordance relates to the perception of opportunities in an environment which is based on how individuals perceive that they can use

objects in the environment to meet their needs (Ekblom, 2012a). Perceptions of action alternatives are, therefore, shaped by an actor's goals, abilities and morals. Affordance is therefore, a relationship between the attributes of the actor and the attributes of the environment, which work together as a system to establish the parameters of action possibilities. 'An affordance is neither an objective property nor a subjective property; or it is both if you like... [It] points both ways, to the environment and to the observer' (Gibson, 1979, p.129). Bloomfield, Latham and Vurdubakis (2010) highlight that the actor and the environment are not operating in isolation, the environment is shaped both by broader causal factors and by the actions of the actors within it. At the same time the actors, themselves influenced by distal causes are also influenced by the immediate environment.

Also of relevance to this thesis is the way in which situational action theory makes two distinctions. Firstly, between actions that are the basis of deliberation versus actions which are force of habit and secondly, between actions that are the result of prior motivation versus those that result from in-situ provocation<sup>8</sup>. This supports Wortley's modification of situational crime theories to add situational precipitators (Wortley, 2001). In contrast to the rational choice perspective, which emphasises the role of deliberation in decision making, Wikström highlights that actions can also arise from habit, defined as automated action born from repeated routine behaviour. Actions are more likely to be guided by habit when a situation is familiar or emotionally charged, 'the more emotionally charged a situation is, the more a person's thinking is dominated by the experiential system' (Carver & Scheier, p.333 in Wikström, 2006b).

### 2.1.6 Summary

Situational crime theories have provided models of how crime is patterned across environments. The models have encouraged academics to interrogate new types of data in the explanation of crime. Criticisms of situational crime theories highlight some need for refinement, with suggestions for these improvements already being put forward by authors including Ekblom, Wortley, Wikström and Bouhana. The language of opportunity as a cause of crime can create the impression of a lack of interest in the causal factors behind opportunities, an impression which Bouhana (2013) argues has discouraged potentially fruitful collaboration with other disciplines.

---

<sup>8</sup> Situations can provoke individuals to commit crime e.g. through prompting anger but they can also provoke feelings that help prevent crime e.g. guilt.

It is also a fair observation to suggest that opportunities are not the cause of crime, they are the consequence of a conjunction of causes and it is the (ecological) explanation of how these causes become conjoined within a crime situation that can explain crime. Central to the crime causation process is the offenders' recognition of how the situation presents potential opportunities, recognition that is dependent upon the characteristics, resources and purpose of the offender. Situational action theory helps to address the criticism that crime opportunities are not objective. It offers improvements to the consideration of offenders' assessments of action alternatives and makes it clear that actors interpret situations subjectively, and stresses the need to explain the individual differences that influence different assessments of the situation.

Of relevance to this thesis, situational crime theories demonstrate that understanding pre-crime situations is necessary in order to explain, and prevent, the occurrence of crime events. MO descriptions contain information pertaining to the situation and offender's actions within it, for example in exploiting vulnerabilities and coping with hazards. Extracting and successfully analysing this information has the potential to provide intelligence to further the development of theory and practice in this field.

## *2.2 Crime Causation: The Conjunction of Criminal Opportunity*

### 2.2.1 Introduction

The above sections have highlighted the role of the situation in understanding actions taken by offenders and also the need to integrate situational and pre-dispositional theories. Before introducing a framework designed for this task, it is appropriate to review how the notion of causality is construed in criminology currently, and the limitations of this stance.

### 2.2.2 Causality and contingency

Theories of crime causation should help understand why crimes are committed and why they are being committed in the way that they are; and it should offer proposals for crime prevention interventions. It should also, in combination with empirical observation, form a central part in anticipating how relationships might change.

Criminology has plenty of theories but the (perhaps over) proliferation of diverse criminological theories has been labelled as unhelpful. Wikström (2006a) argued that the vast range of biological, psychological and social factors that have been identified as correlates of crime 'has led some observers to question if "everything matters" and others whether "anything matters"' (p.117).

Wikström contends that this is a consequence of criminology's search for correlations between crime and other factors and the incorrect assumption that these associations equate to causation. Criminological research has tended to identify 'risk factors' for both offending and crime targets. These are used to predict but may not be direct causal factors, rather symbols or markers of causes. Wikström calls for those working within the discipline of criminology to develop an improved understanding of the nature of causality. Cause implies some form of power or agency to change something, whereas, association merely identifies factors that have a tendency to occur together.

Several authors from within criminology (Ekblom, 2011; Pawson & Tilley, 1997; Wikström, 2006a) argue that causation should be understood through a scientific realist ontology. Realism shares many features with classical positivism: most importantly it posits that the world exists independently of the observer, although it acknowledges that there is interaction between the real world and the observer (Wikström, 2006a).

Adopting a realist perspective can help to elucidate the problems experienced with the treatment of causation in criminology. Much of the frustration with criminological theory would appear to arise from attempts to find universal, one size fits all explanations. Universal laws are rejected by realism. Causes do not operate uniformly in all circumstances, rather they are dependent (contingent) upon the operation of specific mechanisms in interaction with given contexts. Therefore, a causal explanation requires the identification of the conditions necessary to trigger mechanisms that produce particular outcome patterns (Pawson & Tilley, 1997). Extending this Byrne (2009) and Harvey (2009) assert the potential of multiple-causation; that different configurational routes can produce the same outcome rather than there being one causal model that is the 'best fit.'

Secondly, as a divided discipline, criminologists have sought to explain crime either through offenders, or victims, or environments. This ignores the interaction between the elements within the system. Cilliers (1998 in Byrne, 1998) asserts that complex systems cannot be understood by studying their parts in isolation from the whole.

A further problem with the treatment of cause in criminology is the assumption that change is linear. Although systems spend much of the time in apparent equilibrium they have the potential to change in a fundamental, nonlinear, qualitative manner. These phase shifts are analogous to a number of concepts often used to describe qualitative changes in the nature of crime (for example crime waves, sprees, harvests, adaptations etc.).

### 2.2.3 The Conjunction of Criminal Opportunity

Ekblom (2001; 2011) has proposed a framework, the CCO, to aid the integration of diverse theories of crime causation in order to consider the combined operation of causal factors relating to the offender and those relating to the situation. Connecting with the discussion above the framework is not a theory, but a model which aims to systematically identify and classify the causes proximal to crime events which shape their occurrence and nature. Ekblom's model focuses on the immediate precursors of a crime or disorder event; distal causes of crime are not directly located in the CCO framework, although the framework can be used to consider how distal causes shape and influence a crime situation via proximal causes.

Although routine activity theory purports to cover the necessary conditions for crime to occur, the CCO offers a number of potential improvements. The CCO covers causal factors in much greater detail including identifying causes of offender motivation, rather than taking the presence of motivated offenders as a given. It allows for flexibility of description of the roles played by various actors in the crime situation: although additional roles have been added to routine activity theory this has not been systematic. The framework aims to demonstrate the range of factors that are required to cause a crime event and the dynamic interaction between those causes. Ekblom is not suggesting that crime is caused by a mere coincidence of structural factors. The framework aims to accommodate the fact that while the actions of offenders, victims and others are shaped by causal factors, these human actors also actively make decisions and implement actions with the aim of manipulating the environment in pursuit of their own goals.

This framework is compatible with Wikström's situational action theory. Within the framework there is clear space to consider the role of offender morals (which would predispose offending); their goals – readiness to offend, habits; deliberations (their anticipation of risk, effort and reward) – but also the framework shows how the situation can shape the action alternatives available to the offender. The configurations of circumstances in a given situation both motivate and enable people to achieve particular goals. Thus the CCO identifies both causes that influence the crime situation and causes that influence the offender and allows for the consideration of the interactions between the two.

Thus actors have a central role in the framework and are modelled as both caused and causing (Ekblom, 2011). All of these factors are conjoined via the offender's presence in the situation and his or her perception and anticipation of the risk, effort and rewards available. As earlier argued, it is not the case that objective factors create a crime opportunity which is presented to the offender: to the contrary, offenders can play a proactive role in searching for, engineering and perceiving crime

opportunities, while at the same time, victims and crime preventers can be equally proactive in their attempt to stop them. It is particularly important to understand that actors occupying the different roles in a given situation can pursue goals that are in conflict. These points will be further explored in the consideration of theft scripts in Section 2.3 below. This section will provide a summary introduction to the eleven elements of the CCO each of which presents a distinct cause of crime.

The CCO framework stipulates that a criminal event occurs when

An offender who is criminally predisposed,

- **Lacks the skills** to avoid crime,
- Has particular motivational/emotional state and is thus **ready** to commit crime,
- Is properly-equipped with **resources** including tools and perpetrator techniques,
- Perceives and anticipates an acceptable risk of **harm, effort/cost/time, reward,**

Is present in the crime situation with

- A **target** (a property, item of property or person) which they perceive as valuable and vulnerable,
- In the *absence* of people who can act as ready, willing and able **crime preventers**,
- In the *presence* of people who can act as **crime promoters**,
- In a wider **environment** and perhaps a **target enclosure** (building, vehicle, car park) which contains attractive targets, which routinely brings offenders and targets together and whose properties favour offender over preventer.

Eckblom stresses that in order to understand why a crime took place, or to anticipate and prevent a future crime, it is necessary to consider the full set of elements within the framework and how they operate together. This stands in stark contrast to the compartmentalised offender based vs. situation based approaches to crime causation that have dominated criminology. In addition, while the CCO presents eleven distinct sets of causes, none of these causes operate in isolation; on the contrary all elements interact dynamically. For example, individuals with specific criminal predispositions are more likely to socialise and interact with likeminded individuals. This can influence other components of the crime situation, such as the availability of resources, the move to a state of 'readiness to offend' and the assessment of risk and rewards. Thus, the CCO seeks to understand crime events as the consequences of a complex system containing interacting causes.

To allow flexibility in the different roles that can be performed, actors are included in the CCO according to the roles they play at a given time. This flexibility allows one individual to perform different roles: for example an individual could be crime promoting at one instance and preventing the next or an individual's crime promoting actions could result in them subsequently becoming a crime target (victim). The CCO focuses additional attention on attributes of the offender that contribute to the occurrence of crime events thus five nodes in the CCO describe the offender(s)' predisposition to offend, lack of skills to avoid crime, readiness to offend and (psychological) resources for crime. The framework also includes causes that relate to inanimate objects i.e. material targets and resources, wider environment and target enclosure. All of these factors are conjoined via the offender's presence in the situation and his or her perception and anticipation of the risk, effort and rewards available.

Eklblom (2002) suggests that the CCO can also be used to consider how innovation or change, be that political, economic, social or technological, will influence and manipulate the causes of criminal events, this includes considering tracing the impact of distal causes through to the immediate causes of the criminal event.

The nodes of the CCO will now be described in turn, in later chapters the CCO will be used to consider theft from the person offences. The discussion below focuses on crimes occurring in physical environments, however, the CCO has also been successfully applied to virtual environments and cyber-crime (Collins & Mansell, 2004).

### *Predisposition to offend*

The presence of a motivated offender is a necessary condition for a crime to take place. There is a key difference in the use of this terminology between the rational choice perspective and the CCO framework, although the two are related. In rational choice theory, predisposition to offend refers to the *status* of an offender once a 'criminal involvement decision' has been made. This predisposition is limited to a specific crime or set of crimes, although with 'mundane' generalist offenders this set of crimes may be fairly broad. For Cornish and Clarke (2003) predisposition is established prior to the crime event and once made it is made for the long term, representing an 'always on' status. However, other authors, notably Wortley (1998) have argued that factors within the crime situation can influence this decision; producing instances where opportunity precedes, and sometimes provokes, a criminal involvement decision.

In contrast to the rational choice perspective, Eklblom's use of the term 'predisposition to offend' refers to *attributes* that predispose an individual towards offending behaviour. These include social

and psychological influences that have been the focus of 'traditional' criminology that formulate, for example: individual differences in levels of aggression and social control, anti-social attitudes and prejudices, drug or alcohol dependencies. Other important influences include the assimilation of social and sub-cultural norms that encourage and/or enable particular behaviours. In the CCO, predisposition to offend relates to permanent or enduring attributes in contrast to the short term states described under the CCO node 'readiness to offend' (Ekblom, 2011 and see below). The origins of predisposing attributes are generally distal; situated away from the crime event both in time and/or geography. However, in the CCO distal causes are of interest according to the way in which they shape proximal causes. It is useful here to revisit situational action theory's proposition that all crime is moral action; an individual's internalisation of a moral code, the rules which help individuals navigate through action alternatives that they perceive to be right or wrong, are central to the predisposition<sup>9</sup> to offend. The fact that offenders regard theft and violence as morally acceptable actions shapes the action alternatives available to them when deciding whether and how to commit theft (Wikström & Treiber, 2009).

Thus while traditional criminology has sought to explain the distal causes of offending predispositions, these causes have largely been examined with the aim of understanding which individuals will offend and not how these predispositions shape the crimes that are committed. On the other hand, situational crime theories have taken the presence of predisposed offenders as given and undertaken theory building and analysis from the presumption that any situation may be penetrated by motivated offenders. The CCO encourages a consideration of how predispositions which may have formed distal to the place and time of a crime can help to explain, not simply whether or not an individual committed a crime, but why a crime was committed in the way that it was.

### *Lack of skills to avoid crime*

The CCO node 'Lack of skills to avoid crime' becomes relevant where to an impartial observer, there appears to be a legitimate alternative that would meet the offenders' needs. This area is extensively covered by social and psychological explanations to offending behaviour. Included in this group of causes are educational and skill barriers to earning a legitimate living. Also relevant are broader social and psychological capabilities that can aid an offender in the avoidance of crime these include impulsivity and low self-control (Gottfredson & Hirschi, 1990; T Hirschi & Gottfredson, 1995).

---

<sup>9</sup> Wikström uses the term propensity

Wikström argues that self-control only becomes relevant when an individual faces a conflict between their moral rules and a motivation to act in breach of these rules. Thus self-control is defined as “the inhibition of perceived action alternatives or the interruption of a course of action, which conflicts with the agent’s own morality” (Wikström & Treiber, 2009, p.80). The ability to exercise self-control is influenced by relatively stable personal characteristics (executive capabilities); it is also susceptible to transient influences including intoxication, heightened emotions, stress and time of day. Therefore, for Wikström, self-control itself is a situationally-variable factor defined through the interaction of the individual with the features of a setting. This is part of a wider set of executive functions which influence the action possibilities of which the individual is aware and the decisions made in relations to these possibilities.

Personal and social skills can enable an individual to avoid crime, for example having the interpersonal skills to prevent a minor confrontation escalating into a full blown incident. It is notable that Topalli (2005) found that offenders were more likely than a comparison group to (incorrectly) interpret the behaviour of others as violent.

### *Readiness to offend*

In the crime situation the traces of the long term and/or distal causes interact with short term influences on mood and motivation that make offending in a given situation more or less likely. Thus, while the CCO node *predisposition to offend* is concerned with the potential consequences of long term causes, *readiness to offend* centres on an activated motivational/emotional state and includes causes situated in current life circumstances and those in the immediate situation, which may *awaken and direct* particular predispositions. Readiness can stem from desires, but also from frictions, provocations and (the desire to fulfil) commitments (cf. Wikström, 2006a).

There are differences in the use of terminology between Clarke and Cornish (Clarke & Cornish, 2004) and Ekblom (2010). Clarke and Cornish adopt the term ‘readiness’ to refer to the offender’s status once the crime involvement decision has been made. Unlike Clarke and Cornish, Ekblom does not suggest that the combined influences of predisposition and readiness to offend are restricted to a specific list of acceptable offences. Rather it is feasible that the short term nature of influences on offence readiness allows for a degree of flexibility in offence alternatives. Different types of readiness will result in the adoption of different methods as the rewards and risks considered in any assessment of cost benefits will be qualitatively different for differently ‘readied’/motivated offenders.

### *Resources for committing crime*

Offenders need to be suitably equipped with resources in order to meet the challenges set by specific offences. Offenders with restricted resources must either accept a limited range of offending possibilities, or develop strategies to accommodate for resource gaps. These strategies may include adopting a different method of offending, taking active steps to skill and/or tool up, or enlist the cooperation of co-offenders who are equipped with the relevant resources. Resources should, therefore, not be regarded as static as an offender can actively seek to improve his or her access to resources. Ekblom and Tilley (2000) categorise the resources required for crime commission including personal characteristics such as strength and agility, cognitive resources including learnt techniques or awareness of targets, moral resources, physical resources including tools, weapons and transport, and collaborative resources including membership of criminal networks. These resources are required throughout the crime commission stages. Ekblom and Tilley point to Cloward and Ohlin (1960) to highlight the unequal distribution of resources for offending which parallels the unequal distribution of resources for many legitimate activities. Access to resources is regulated by a combination of structural factors, the offender's strategic efforts to generate capacity and crime preventers' strategic attempts to restrict capacity. It is important to note that the selection of MO is not entirely at the disposal of the offender, availability of resources (along with the other CCO components considered here) will set the parameters of available choices. Despite the limited direct investigation of offender resources (Ekblom & Tilley, 2000 and Gill 2005 are amongst the exceptions); available studies on offending and MO do shed light upon the resources required in a range of offence types, as will be explored in the context of theft from the person in Chapter 5. Mativat and Tremblay (1997) explain how membership of ethnically defined networks (a human resource) facilitated the development of/taking up of crime scripts, while presenting a barrier to non-members. Offenders engaged in pure counterfeit credit card fraud (which is dependent upon the manufacture of cloned cards) were of Asiatic origin providing a distinctive link with the countries where cards were manufactured.

### *Offender presence in the situation*

For a crime event to occur the motivated and ready offender or offenders must be present<sup>10</sup> in the situation. Crime pattern theory (Brantingham & Brantingham, 1993) provides a model for understanding how offenders come into contact with crime opportunities, based on the patterning of their routine activities, the nodes they visit and the pathways between them. Therefore, offender

---

<sup>10</sup> Or virtually present in cyber-crimes.

choices are bounded by the geography they inhabit with distance and familiarity important criteria for target selection (Hakim, Rengert, & Shachmurove, 2001). Studies of offenders' journey to crime have revealed a distance decay pattern (Wiles & Costello, 2000). In a quantitative model of residential burglary Bernasco and Luykz (2003) found that 60 per cent of the variation in burglary levels across neighbourhood could be explained by proximity to the offender's home. Acquisition of other resources, notably a car, can extend criminals' awareness space and enable them to conduct a wider search (Wright & Decker, 1997).

### *Crime targets*

Newman and Clarke (2003) identify four categories of crime target. Prime targets are those that directly meet the ends of the offender. Convertible targets are those that can be exchanged in order to reach the desired end such as stolen goods to convert into cash, or cash to convert in drugs. Transitional targets are targets that constitute stepping stones towards the end goal, for example stolen credit card details that provide the personal data for mortgage fraud. In the case of theft from the person the 'victim' may constitute a transitional target that needs to be overcome in order to obtain a convertible or primary target. Ekblom uses the term target person to refer to the victim. For ease of expression the term victim has been adopted in the analysis within this thesis.

Ekblom and Sidebottom (2008) highlight the often overlooked and over-simplified distinctions made between target value, target attractiveness and target vulnerability. They distinguish between these concepts and use value to refer to the potential of a product to gratify the motivation of the offender be that for enjoyment, status, misuse or the ability to transfer it into/for another valuable item. Target vulnerability relates to inherent and permanent (or very long term) characteristics of targets that are inseparable from the object or person that increase the probability of attack. Vulnerable targets are those that have features that enable the offender to take it or attack it. Therefore vulnerability relates to features of the target rather than the motivation it/they engender in the offender. For Ekblom and Sidebottom a target's attractiveness relates to features which cause the offender to form the intention to steal it, either due to its perceived value, or some combination of vulnerability and capacity for realisation of value. Attractiveness should be considered as a subjective attribute from the perspective of the offender, in terms of what they themselves value and how well equipped they are to take the product in-situ, and then realise its value. Attractiveness is therefore situationally defined, and is a composite of value and vulnerability.

Target attractiveness has been considered by rational choice theorists. Cohen and Felson introduced the VIVA acronym (Value, Inertia, Visibility, Accessibility) and later Clarke (1999) introduced CRAVED

(Concealable, Removable, Available, Valuable, Enjoyable, Disposable) in a bid to explain why some products are more commonly stolen than others. These acronyms highlight how factors inherent to the product can define value/enjoyment<sup>11</sup> can increase the value of a product to offenders but understand this it is necessary to consider the specific motivations of the offender – what does s/he find enjoyable. The same target, without changing in value can increase or decrease in attractiveness to the offender based on changes to the situation, changes in motivation or the resources available to the offender; in fact, based on changes to any of the other CCO nodes. Thus to usefully apply these acronyms requires rather more work than simply thinking about the product/target in isolation.

### *Target enclosure*

Target enclosure refers to structural features which surround the target and may include buildings, gates, safes etc.; and which are characterised by interiors, boundaries, entry/exit points and peripheries. Research consistently demonstrates that the nature of an enclosure influences a target's attractiveness to offenders. This finding is consistent for research investigating different crime types (burglary (Wright, Logie, & Decker, 1995), car crime (Michael, Hull, & Zahm, 2001), shoplifting (Carroll & Weaver, 1986), robbery (Wright & Decker, 1997), vandalism (Smith, 2003)) and using different methodologies. Nee and Taylor (2000) and Coupe and Blake (2006) show how cues relating to target attractiveness (which include the nature of the target enclosure) are considered relative to one another. For example an offender may accept the greater effort/risk of a stronger target enclosure if it offers greater value targets.

### *Wider environment*

Target enclosures are located within wider environments such as shopping centres, public spaces, housing estates or transport hubs. Although Ekblom makes a distinction between target, target enclosure and wider environment, dependent on the crime type these elements may overlap (for example a car can be a target of theft of vehicle, and an enclosure for theft from vehicle). It is also difficult to consider them separately when offenders consider them together (the attractiveness of a target will be influenced by its enclosure and the wider environment in which it is situated).

---

<sup>11</sup> If Ekblom and Sidebottom's definition of value is adopted then valuable and enjoyable are overlapping elements, although enjoyable related primarily to the 'prime target.'

Consequently some elements discussed under wider environment are equally relevant to targets and target enclosures.

Wortley highlights that criminals can both be motivated and demotivated by immediate environments (Wortley, 2008). The recognition that the environment can influence criminal behaviour emerged in the 1960s with the term Crime Prevention through Environmental Design (CPTED) later coined by Jeffrey (1971). The key principles of CPTED outlined by Cozens, Saville and Hillier (2005) Cozens, Saville and Hillier (2005) help to illustrate the features of environments that can be criminogenic/criminocclusive. Central to CPTED is the concept of **defensible space** (Newman, 1973) which relates to the design of buildings/enclosures in such a way as to help occupants, owners and users keep criminals out. Closely related is the concept of **territoriality** where space is designed to communicate ownership, and to distinguish between public, private and semi-private space, for example through the use of fences, pavement treatments, artwork, signage and landscaping. Design can also communicate the intended use of space. Good designs increase the motivation of those responsible for, or using, a space to ensure control over it (poor design can also encourage negative territoriality such as turf wars between gangs). Territoriality can be facilitated by **access control** to keep certain people out of buildings/enclosures and structures, while still admitting those people with a right to be there. Natural access control aims to seamlessly guide users in entering and exiting a space through the careful placement of signs, entrances, exits, fencing, landscaping and placement of features in the environment. More formal measures aim to physically prevent users from entering spaces that are out of bounds. The degree of natural or formal **surveillance in the environment or around the target enclosure** render environments more or less conducive to crime. Surveillance can be technological, most commonly with CCTV, but also the placement of features in the environment (buildings, columns, vegetation) can influence the extent to which people are able to see and be seen by others. Perhaps more relevant to target enclosures, the extent to which physical structures such as walls, windows and doors are resistant to attack and/or penetration by criminals (**Target hardening**) influences the availability and attractiveness of targets. The perceptions that actors, including offenders, potential victims, crime promoters and crime preventers hold about an area are shaped by the aesthetic appearance and social reputation of a building, place or neighbourhood. Thus the **image** and **maintenance** of a location shapes who uses a location and how it is used. Maintenance further ensures the effectiveness of security systems and acts as an additional expression of ownership. The number of people spending time in or passing through a particular location conducting, routine, honest activities like shopping or dining provide activity support which can prevent, deter or discourage offenders from committing crime. Increasing the number of people passing through or working within a location can help to increase the supply of crime preventers

who can interrupt the actions of offenders either directly or indirectly. That is unless these individuals are unable to perform this role in which crowds can become cover for offenders.

### *Crime promoters*

Felson (1995) notes that crime events cannot be understood by focusing on only one actor within the system. Crime promoters are people who either inadvertently or deliberately play a role in increasing the likelihood (and harm) of a criminal event. This may be through carelessness, encouragement, or provocation or through the provision of practical contributions, as seen above through the provision of resources. Within the CCO it is possible to distinguish between proximal and more distal promoters, examples of the latter include the manufacturers of criminogenic products. The role of crime promoters is important, but again it should be stressed that where they are needed but not directly available offenders have the option to actively seek them out.

### *Crime preventers*

Crime preventers are active roles in which people make crime less likely to occur, or reduce the impact of a crime. Again the actions of crime preventers may be deliberate, or accidental, the role may be formal (police, street wardens, security guards) or informal (retail staff, passers-by). The role of crime preventers can operate prior to a crime event (locking away valuables), during it (sounding the alarm), or following (identifying perpetrators, which may lead to prevention of *subsequent* crime events).

The role of crime preventer is closely related to the capable guardians, place managers and intimate handlers described in routine activity theory but is more flexible and allows the inclusion both of individuals who exert considerable influence and those with only a minor influence.

The definition of a crime preventer as a 'role' rather than a 'person' allows for the possibility that one individual may perform different, sometimes quite opposing, roles within the course of crime commission/attempted crime commission. For example a target victim may initially exhibit crime promoting behaviour but on recognition of the threat posed by a the offender may switch to adopt preventing behaviour such as putting valuables out of site, walking quickly or selecting a less secluded route. The broader scope of crime preventer also avoids the problem that seems to have been encountered with routine activity theory with the addition of preventative roles each time a role is identified that does not fit.

### *Offender anticipation of risk, effort and reward*

From the rational choice perspective an offender undertakes an (albeit imperfect) calculation of the costs and benefits before committing a crime, and this is incorporated within the CCO. Importantly from this perspective these costs and benefits are 'not enduring characteristics of persons' (Cornish & Clarke, 1987) but vary from one potential crime situation to another. Salient conditions of a situation will influence an offender's decision to commit an offence and shape the way in which it is committed. Nee and Taylor (1998, 2000) emphasise the importance of the interaction between the various situational cues that inform an offender's decision making. They highlight the dangers of assuming a symmetry of cues, for example that the presence of a CCTV camera carries the same but opposite weight as the absence of one. Moreover, as highlighted above, the decision agenda of risk, effort and reward comprises interchangeable currencies – the variability on which Sutton bases much of his critique of opportunity based theories.

### 2.2.4 Summary

The CCO has been presented as a method to organise available theories and concepts and use it to understand how interactions between the actors and the environment shape and structure crime events. CCO elements cannot be considered in isolation. The operation of each element in a given situation is defined by the status of all of the others. Although the focus of the CCO is on proximal causes, the framework helps to consider how distal causes exert their influence on the immediate crime situation. This allows for a fuller consideration of the nature of offenders, their motivations, resources and constraints and can therefore be used to address some of the criticisms raised regarding choice theories.

The CCO is comprehensive and encourages a more detailed exploration of crime causation than alternative models (such as the crime triangle/problem analysis triangle (Clarke & Eck, 2005) largely derived from routine activity theory). However, there is a risk that the elements of the CCO are interpreted as static, and they need to be interpreted as dynamic and interacting with each other. A side-effect of this interaction is that there are many overlaps between the elements meaning it can sometimes be difficult to locate information in a suitable place within the framework. The following section will describe a framework for interpreting the crime commission process, namely crime scripts. While the CCO has the potential to describe dynamic interactions between the elements *crime scripts* directly address the wider, sequential dynamics of the crime situation, which might lead the offender through a *series* of individual conjunctions.

## *2.3 Criminal Methods and Crime Scripts*

### 2.3.1 Introduction

This section introduces the concept of cognitive scripts and explains how they function to facilitate decision making. This is followed by a demonstration of the use of scripts as a methodological framework to study the decision making of offenders and to inform the design of crime prevention interventions. Cornish (1994) proposed that cognitive scripts could be used to make sense of the crime commission process by viewing each crime as a sequence of temporally ordered stages or 'scenes.' In describing the steps taken to complete a crime, crime scripts are a similar concept to MOs. However, a distinction will be made in this thesis between MOs, which represent a simple description of methods (either of one offender or a methods commonly adopted by many offenders) and crime scripts, a description of methods which has been structured, using a specific methodology, into a clear sequence.

### 2.3.2 Cognitive scripts

The concept of a cognitive script was first developed in the 1970s by a multi-disciplinary team of cognitive psychologists and computer scientists led by Robert Abelson and Roger Schank at Yale University. Schank and Abelson's (1977) original formulation of scripts was concerned with their function as structures for memorising knowledge. However, Abelson (1981) argued that because scripts function specifically to organise 'procedural knowledge,' i.e. knowledge that is exercised in order to perform a specific task, scripts can be equally useful for understanding the behaviour that is directed by this kind of knowledge.

Scripts provide actors with templates (Schank & Abelson, 1977) that organise knowledge and guide behaviour in familiar or conventional activities; in doing so scripts provide a form of 'perceptual shorthand' that allow the actor to operate successfully within a complex world through simplified, stereotypical knowledge structures. In other words, they provide a selection of sub-routines that can be performed without thinking too much about them. The most frequently quoted example is that of the restaurant script, which contains the scenes: finding a seat, reading the menu, ordering drinks from the waiting staff and so on (Schank & Abelson, 1977). Similarly, Morr, Serewicz and Gale (2007) examined first date scripts, identifying a regular pattern of scenes that start with 'GET READY' and (hopefully) end with 'FUTURE PLANS'. Script frameworks have been applied extensively to several realms of social life including shopping (e.g. Leigh & Rethans, 1983), sexual behaviour (e.g. Krahe,

Bieneck & Scheinberger-Olwig, 2007), management studies (e.g. Goodhew, Cammock & Hamilton, 2008) and education (Monteiro, Carillo and Santiago, 2010).

Scripts enable the actor to rapidly retrieve past action experience that is relevant to the current situation, sometimes even before the actor is in possession of full information about the circumstances (Seifert et al., 1994). To aid their implementation, scripts contain cues that prompt the actor to select the most relevant script, for the current goal and present circumstances. These prompts guide both decision making and physical behaviour.

For complex processes the decision making is broken down into stages. These stages are organised in a hierarchical structure in which lower level goals are the means through which higher level goals are achieved. Thus, lower level, tactical goals are performed *'in order to'* achieve higher level goals. The existence of *in order to* relations imposes inherent causal, logical and clear temporal structures to a script. In such cases certain script scenes will have a strict temporal order with later stages being strictly dependent on earlier steps of the sequence, while in contrast, other scripts will allow greater flexibility in their ordering. Initial script scenes can shape action in later scenes with early scenes restricting or expanding the range of choices available to perform subsequent scenes.

Scripts can contain cues that indicate that action possibilities exist for a given set of circumstances. Abelson (1981) argues that actors will have a rule-based policy for the initiation of a script that is influenced by situational conditions. In concurrence with the rational choice perspective, relevant conditions governing script initiation might include costs, effort or incentives. Abelson cited several research studies investigating the willingness of individuals to engage in behaviour to assist others, these studies revealed how very small variations in situational context can result in significant differences to the probability of entering into a script. Abelson concluded that 'the situational lability of behaviour is theoretically troublesome for explanations based on abstract values such as altruism or social responsibility. But from a script orientation, such lability is more or less what one would expect' (p.719). If an individual's probability of engaging in helping behaviour is open to change based on small differences in situational context, it would be reasonable to expect that situational context would similarly influence the initiation of criminal behaviours. These observations would appear to support Wikström's situational action theory, and provide an explanation for the lability, or variation, in offender decision making which Sutton finds unscientific. For example, Eck (1993) found that an offender's familiarity with a situation influenced his commitment to crime commission. Eck argued that an offender will persist when, places, or targets are familiar to him and that he will be less likely to persist when they are not.

Abelson (1981) argues that an individual must not only recognise that the possibility to perform a script exists, they must also commit themselves to performing it. He states that 'starting a script performance usually entails a commitment to finish it' (p.719). Notwithstanding the latter, it is possible that crime scripts, performed as they are in hostile circumstances, may contain a greater number of readily embedded prompts and pathways for 'script abortion' than other types of behaviour.

In addition to prompts that guide an actor's behaviour, scripts (as cognitive representations) contain details of the resources or 'props' that are required to complete a scene. Since the majority of tasks are performed in social situations, script will also include an outline of the way in which other actors are expected to behave. Script templates can contain tags that flag people who have particular skills or resources to assist in accomplishing script goals. Scripts can, therefore, help handle collaborators, unwitting assistants (such as crime promoters) and adversaries and can include strategies for handling both the complementary and clashing scripts of other actors. It was noted above, that scripts help to narrow the actor's choice down to the most familiar sub-goals that seem to fit with the current conditions. By relying on this shortlist of pathways to goal, actors are guided to the most appropriate path for their given circumstances and heuristically avoid the need to make a comprehensive search of all options. Studies in artificial intelligence have shown that sequential searches through memory are inefficient (Abelson, 1981) therefore, such a strategy for script selection would defeat the distinctive benefits of scripts in serving to simplify elements of decision making. This conception of decision making fits well with the conception of 'bounded' or 'situated rationality' within rational choice theory, discussed earlier.

On the surface, the notion of predefined pathways to goals can paint an overly mechanical picture of thinking and behaviour. However, it is important to note that scripts provide a structure for *organising* decision making and realising it, not replacing it. In this way, scripts are distinctly different from habits, with habits directing strictly automated behaviour. Wikström (2009) distinguishes between action that follows deliberation and action that is force of habit. Scripts offer a means to organise both sorts of action. Examination of scripts reveals that one obvious way in which these frameworks facilitate mindful behaviour is the necessary inclusion of scenes that require an explicit decision. For example, the classic restaurant script includes the scene 'decide what to order' as Abelson notes 'while it is possible for ordering food to become habituated to a single food item, there is typically a place in the script specifically marked, in effect, as "Now think"' (1981, p.723). As will be seen below, these observations can be extended to scenes within crime commission scripts, including 'target selection.' Wikström stipulates that it is possible for habitual behaviour to be

interrupted by interferences; again, the script framework can be used to understand how interferences are managed with action being diverted from the automatic to the scripted.

Like habits, unusual situations may not be scripted. Langer, Blank and Chanowitz (1978) found that scripts were not used when circumstances were so novel that the situation has to be processed in detail. Copes, Hochstetler and Cherbonneau's (2012) research demonstrated that carjackers did not always stick to their scripts and were at times forced to divert from scripts to actions that required greater consideration and planning. Thus there is, therefore, a continuum of actions from reflex, habit, scripted /prompted deliberation and novel behaviour.

Scripts can facilitate the processes of learning. Once a new action is performed, or even observed, novel behaviour can become embedded as a new script or incorporated into an elaboration of an existing one. Importantly, Eckland-Olson, Lieb and Zurcher (1984 in Hochstetler, 2001) argued that script scenes are interchangeable, allowing elements of scripts to be 'decoupled' in order to provide the 'cognitive building blocks' for the performance of new, non-scripted, tasks. This recombination of existing script elements represents an important form of innovation; if successful, a new combination is remembered for the future. This is reflected in the observation that individuals who have experience and expertise in a given task, will tend to have more thoroughly developed scripts. Examining what novices do poorly can also reveal what has to be learnt to successfully perform a given script, as demonstrated by Carroll and Weaver (1986) in their comparison of expert and novice shoplifters.

Experience in performing a given script can also lead to its refinement. When established scripts are applied within a situation they are reformulated based on feedback resulting from their performance. If one means to the end is blocked, then actors can select an alternative route. Where familiar, scripted, actions do not produce expected outcomes and pathway alternatives are unsuitable, then more conscious attention is triggered. The experience of overcoming obstacles can then be incorporated as part of a refined version of the script. This was demonstrated in Copes, Hochstetler and Cherbonneau's (2011) study of carjackers in which robbers described how bad experiences caused them to build revisions into future performances – for example, ensuring a less conspicuous approach, and watching victims carefully in case they reach for a weapon. Graesser, Gordon and Sawyer (1979 in Abelson, 1981) call this incorporation of new experiences "tagging," and Piaget terms the process "adaptive learning" (Piaget, 1954 in Ekblom, 2011). During this process, new information is explicitly encoded and tagged into a generic script structure. Positive experiences can also act to refine scripts, although, Bower, Black and Turner (1979) found that obstacles and errors are remembered particularly well.

The facility to incorporate newly learnt information into scripts is central to the ability to adapt behaviour in the light of changing circumstances, either to short term changes in the environment or more permanent 'phase shifts' requiring substantial adaptation. Cornish and Clarke (2002) distinguished between first order and second order elaboration of scripts. In first order elaborations, new pathways are identified for accomplishing existing script scenes. Second order elaborations result in more significant script changes, for example, the insertion of new scenes or the reordering of existing ones. Investigation of these elaborations will be particularly important for the present research into script change.

Of particular import to the present research are the patterns that are produced as a result of similarities in the ways offenders perform crime scripts. While individuals construct, store and adapt their own scripts, Greene (1989 in Woodhams & Toye, 2007) proposed that 'because some things work in achieving goals and others do not, people will tend to establish similar representations of action - outcome relationships,' (p.202). Still, this does not lead to completely uniform patterns of behaviour amongst those with similar goals (or different occasions in which the same individual achieves the same goal). Abelson (1981) outlines the concept of *equipfinal actions*, to explain the observation that different actions can be performed in order to accomplish the same result (i.e. there is more than one way to peel an orange).

Greene's observation that certain steps are always necessary to get particular jobs done applies to physical elements of task performance, but due to cultural norms, it also applies to social elements of tasks. The 'first date' scripts described above are shaped by cultural norms and expectations and within these cultures, script availability may be shaped by social status (for example gender, age, or religion). From the perspective of situational action theory all action is moral action; therefore, the moral outlook of the actor will restrict the scripted actions available to individuals. The script pathways of offenders may violate mainstream cultural norms but there is still room for them to be shaped by the values of offending sub-groups. Other script pathways may be shaped by access to resources such as money or transport.

A distinction can be made between strong and weak scripts. As a result of restrictive physical and/or social constraints, strong scripts display a high level of agreement in how script pathways are performed and therefore, most people perform the script in the same way. This compares to weak scripts that allow greater freedom and display less agreement and greater variation in the way people perform them. Consequently while the nature of scripted behaviour leads to high levels of between-actor consistency this will always be diluted to a degree by inter-individual variation (Mischel, 1999 in Woodhams, Hollin & Bull, 2007).

The above discussion has argued that while scripts do not replace inference processes they do serve to facilitate rapid and efficient decision making and action in a number of ways. Script frameworks contain several shortcuts to decision making including: alerts to potential action possibilities in a given situation, cues about how to act in certain circumstances, prompts and resources required to complete all or part of a task and guidance to the anticipated behaviour of other actors.

### 2.3.3 The utility of the script framework in understanding behaviour

In addition to viewing scripts as a cognitive facility that guides decision making, scripts are also used as a research method to understand decisions by inspecting actions that have been performed or considering actions that could be executed in given circumstances. Borrion (2013) distinguishes between potential scripts and performed scripts. Potential scripts refer to ways in which it is theoretically possible to commit a crime. Performed scripts recount historical information about how crimes have been committed. The use of both potential and performed scripts offer useful methods for understanding crime. The findings that are presented in Chapter 7 and 8 are examples of performed scripts drawn from historical police data. The inspection of patterns in these performed scripts has the potential to allow inferences about causal processes that govern decision making and action. Consideration of potential scripts could help to identify ways scripts might change in the future including in response to deliberate attempt to block them.

Analysis of collections of scripts encourages the systematic consideration of all stages of behaviour and the identification of the necessary characters (people) props (resources) and actions required to accomplish a task. Comparisons between scripts can inform explanations for differences in behaviour in different contexts, between different individuals and changes over time.

For example, Morr Serewicz and Gale (2007) found that dating scripts are gendered, female first date scripts contained the pre-date scene of calling friends for reassurance, male scripts did not. The research identified that first date scripts had remained relatively stable over recent decades and that despite moves towards female equality, the gendered elements of dating had changed little. However 'dating' research has been concentrated on Northern European and Western European samples. Cultural comparisons could reveal these Western scripts to be wholly ineffective, inappropriate and perhaps offensive in other societies.

In addition to a detailed description of 'what happened,' the inspection of scripts reveals the challenges that are accomplished and the resources that were required for the successful execution of a script. This can help to shed light on the actor's decision to act, and his decision to perform a specific script permutation over others. It may be possible to draw inferences from exploration of

the 'routes not taken.' Given that the factors which shape an actor's selection of script can be related to the situation, the CCO offers a potential means to identify and explore the relationships between an actor's goals, his or her efforts to obtain their goals and the crime situation. There is also a potential to use script analysis in efforts to predict behaviour. Yun and Roth (2008) identified a set of distinctive hostage taking scripts that each led to different outcomes for the hostage (release, execution); identification of early script stages could be used to anticipate likely outcomes for different scenarios.

While scripts appear to provide a useful tool, Borrion (2013) and Ekblom and Gill (2015) have highlighted how researchers have used the concept inconsistently to describe both the behaviour of actors, and events. The latter entails more than a description of one actor's decision and actions, it also includes descriptions of other actors' behaviours and situational conditions.

### *Scripts as a methodology*

Borrion (2013) observes that researchers have used scripts at different levels of abstraction. To handle this hierarchy, various researchers have devised terminology to describe the various script elements. However the terminology has been rather inconsistently applied – in many cases researchers have devised their own terminologies to provide clarity to their audiences (Chainey & Tompson, 2008). This is understandable but can add to the confusion when attempting to compare findings.

Cornish (1994) identified four levels of script specificity: tracks, scripts, protoscripts and metascripts. More recently, Tompson and Chainey (2011) have revised this typology adding levels that allow scripts to be additionally divided into acts, facets and scenes. Following typologies produced by Cornish and Tompson and Chainey, the following terminology will be adopted to describe scripts, broader families of script and detailed script components. *Metascripts* represent procedures for undertaking the broadest groups of offences e.g. theft of property and *protoscripts* represents the level at which sub-categories of offences can be distinguished e.g. theft from the person. *Script* refers to specific offences that are subdivided by dimensions with the situation (Leclerc, Wortley and Smallbone 2010), this might include aspects of the victim, the location or methods used by the offender. Scripts can be further divided into *acts*, which represent the main process stages required to complete a script and *scenes* which are the processes required to complete each act. There is little specific guidance regarding how to divide crime scripts into acts and scenes, leading to considerable variation in how these terms have been applied (Borrion, 2013). Extending the dramatic analogy, throughout this thesis, acts will be defined as requiring a specific action resolution for completion

and/or notable changes in location or time between acts. Scenes will be demarcated by changes to the characters present and/or small changes in setting or time, scenes that are part of the same act generally take place in the same locale (Dunne, 2012). The division of scripts into acts and scenes is particularly helpful because it encourages consideration of the causal factors that shape each of these different stages, and for some scripts these can vary considerably.

Finally, tracks correspond to different permutations of actions that can be utilised to accomplish scripts and can be viewed as the different routes that are available through the acts and scenes of a crime script, such as robbery with or without the use of physical force. Leclerc, Wortley and Smallbone (2011) observe that it is at this level at which most situational crime prevention is implemented. Although tracks are distinctive, Tompson and Chainey state that tracks should share sufficient salient conditions to keep them as part of the same script. With complex crimes there may be a myriad of different permutations through which goals can be achieved. This has led a number of authors to realise the question of balance between appropriate levels of generalisations and allowing for the exploration of meaningful variations in scripts (Wortley in Ekblom & Gill, 2015).

Leddo and Abelson (1986) suggested that all scripts, regardless of their function, could be generalised into a universal script containing the scenes '**Preparation, Entry, Preconditions, Initiation, Actualisation, Doing, Post Condition, Exit.**' However commentators have argued that this generic framework is difficult to apply in practice (Ekblom & Gill, 2015). Tompson and Chainey (2011) reduced the universal script to four categories and argued that rather than include pre and post conditions as distinct and separate scenes, analysts should consider the prerequisites and facilitators within *each* scene. Prerequisites can, therefore, be considered as milestones, sequential dependencies on which future steps are reliant. The inclusion of prerequisites promotes the consideration of how of causal factors in the situation shape and are shaped by actors performing scripts. In the case of crime scripts, this allows the consideration of how offenders exploit and cope with the conditions in the situation described by the CCO; in fact, it is possible to identify a separate conjunction of opportunity that must be exploited/cope with at each scene (Ekblom, 2003).

### *Crime scripts*

Cornish (1994) first advocated the use of scripts in the analysis of crime, the fact that by their very nature, scripts are context-specific, makes them particularly suited to the call from situational perspectives for a crime specific analysis, "The script concept offers a useful way of developing better accounts of crime commission. By 'better' is meant fuller, more systematic, appropriately crime specific, and helpful for crime prevention purposes." (Cornish & Clarke, 2002, p.48).

Script frameworks have been applied to the analysis of a number of types of crime including car crime (Copes et al., 2012; Jacobs, Wright, & Topalli, 2003; Tremblay, Talon, & Hurley, 2001), Fraud (Tremblay, Cusson, & Morselli, 1998; Tremblay, 1986) Sexual offences (Beauregard et al., 2007; Brayley, Cockbain, & Laycock, 2011; Leclerc & Tremblay, 2007; Leclerc et al., 2011) Robbery (Copes et al 2011) and illegal waste activity (Tompson & Chainey, 2011) further examples can be found in Leclerc & Wortley (2014) (Leclerc & Wortley, 2014). Research findings that did not originally adopt script frameworks can also be revisited and interpreted with the lens of crime scripts (as recommended by Cornish & Clarke, 1994 and Copes et al., 2011).

Using a script framework, crimes can, therefore, be regarded as a temporal sequence, containing scenes that the offender performs to achieve lower level goals that will ultimately lead to a higher level goal of crime commission. These goals can include objectives that positively further overall aims e.g. to transfer property from the victim, but also include so called 'hygiene scripts,' these are procedures designed to avoid undesirable outcomes such as sustaining injury or being recognised (Ekblom & Gill, 2015). The script covers all stages of the crime from preparatory steps, through actualisation of the offence to subsequent post crime phases (Cornish & Clarke, 1994, 2008). Crime scripts can vary from the straightforward, crimes with only a small number of acts and scenes and only a few options for how each scene can be performed, to more complex crimes, requiring the accomplishment of multiple acts which, in turn, are comprised of scenes, each of which can be performed in a variety of ways. Other crimes may have only a few scenes but a large range of options for carrying them out. Research into offender MOs has found little variation in the scripts for certain crime types (e.g. bag theft, Gamman, 2001<sup>12</sup>) and significant variation for other crime types (e.g. vehicle theft, Tremblay, Talon & Hurley, 2001). Within one crime type, variation can range from high tech to low tech scripts. For example, techniques of identity theft range from the theft of utility bills from rubbish bins to the creation of cloned financial websites (Copes & Vieraitis, 2009b). The diversity or uniformity of crime scripts employed for specific crime types may have implications for the identification of newly emerging crime scripts.

The analysis of crime scripts, not only, outlines each of the necessary stages in the commission of a crime, but also, identifies the required props and cast members. Systematic analysis also uncovers the interplay of offenders' scripts with those of suitable victims, those promoting crime (either consciously or unconsciously) and those with the potential or desire to prevent crime. A comprehensive application of script frameworks highlights that it is not just the scripts of offenders

---

<sup>12</sup> Gamman uses the term perpetrator techniques rather than crime script.

that require consideration, attention must also be afforded to the scripts that reveal the mechanics and logistics of the routine activities of everyday, honest life (Leclerc, 2013) and the direct script clashes between victims and offenders, and crime preventers and promoters (Ekblom, 2012b).

Analysis of scripts can reveal the process of offender decision making and therefore, has the potential to inform those tasked with preventing crime, of the stages or scenes in which they can intervene. A script approach embraces the whole process of crime commission and can (and should) be applied to all stages in the commission of crime, from the planning and resource gathering stages through to the concluding steps, such as disposal of tools or selling on of goods. These are all stages in which crime can be made harder, riskier or less rewarding through intervention. Script analysis can identify points of track change and/or failure – where the offender is forced to take a different, more difficult or higher risk track or where the script is aborted without the final ends being achieved. Analysis of performed scripts can identify the interruptions, obstacles or missing resources that lead to changes to or abandonment of script tracks (Cornish & Clarke, 2008). This can help to understand how and where to block scripts, and potential offender counters to blockages in crime scripts.

Ekblom and Gill (2015) argue that *empirical* descriptions of offender methods are insufficient to inform crime prevention. They argue that crime prevention intelligence requires *explanatory* descriptions, such descriptions relate actions to the functions they aim to achieve and information on the causal mechanism that mediate behaviours. Explanatory scripts can also contain information on how behaviour changes over time, such as with age and experience and how scripts evolve in response to longer term environmental changes.

Crime scripts can also be used to anticipate likely offender innovation in response to attempts to reduce criminal opportunity (including displacement). Analysis of crime scripts over a longer timescale can expose the evolution of criminal techniques. Tremblay et al (2001) argue that the structure of available criminal opportunities translates into identifiable offender behaviour patterns. Changes to the opportunity structure should, therefore, be identifiable in changes to aggregate behaviour.

Cornish (1994) describes different modes through which scripts are elaborated. Scripts are adapted in response to opportunities, obstacles (including crime prevention interventions) and aggregate offender learning. Cornish further argues that, the more hostile an environment is to crime commission, the more likely that adaptations will focus on *risk minimisation*, in more favourable environments adaptations will be more attuned to *maximising rewards*. Therefore, the type of

change identified could be indicative of the cause. This type of adaptation has been highlighted, for example, in relation to drugs markets. If drugs markets are squeezed by law enforcement they are more likely to become closed and risk adverse making it harder to identify and access ring leaders. Markets, where operators are less fearful of intervention, may be easier to police, at least until the operators become aware of the intervention (Haracopos & Hough, 2005).

#### 2.3.4 Summary

Scripts were originally presented as a cognitive facility that aids effective and rapid decision making to guide goal-directed behaviour. However, the role of cognitive scripts is to organise decision making, they do not replace it. Scripts are also used as a methodological tool that aims to reveal decisions made and the factors that shaped them through an inspection of performed actions. Inspection of scripts can reveal the temporal sequences within actions, and help to suggest the lower and higher level goals that actors aim to accomplish. The fact that scripts contain cues relating to different situational conditions is directly relevant to situational explanations of crime and can facilitate analysis of the proximal causes that shape offender decisions and actions. Leclerc (in Ekblom & Gill, 2015) believes that as script “is far more than a methodology – it is rather, a journey into the head of the offender during the commission of a crime. It can reveal the offender’s motives and other previously unseen situational aspects, some of which may inhibit the script’s execution” (see Ekblom & Gill, p.5). However, the extent to which scripts ‘reveal’ such intelligence will depend upon what data is available about offender actions: making inferences about offender intentions, decisions and actions based on limited data is risky.

Scripts have been used to understand offender behaviour but there is also potential to explore crime promotor and crime preventer scripts. In particular, analysis of victim scripts offers the potential to consider the situational lability of target vulnerability.

The hierarchy of different levels of script specificity enables the examination of behaviours with similar purposes that display different actions performed in order to reach them. Finally, the use of script methodologies can help to anticipate potential changes to criminal methods. This can include the identification of crime prevention interventions that can disrupt offender scripts and the consideration of the consequences of blocking specific actions. Script methodologies can also be used to explore the potential impact of changes to routines activities, through, for example, social changes or technological changes, and whether these will facilitate or hinder offenders in reaching their goals. The consideration of potential scripts can, in turn, help identify potential offender adaptations in response to both deliberate and accidental disruptions to crime scripts.

## *2.4 Previous Studies Analysing Offender MO*

The section above introduced scripts as a conceptual and methodological tool that has the potential to reveal the actions performed by offenders and the factors in the environment which shape them. The current section now moves to practical issues of conducting analysis of offender methods, with a focus on studies that have utilised MO descriptions that are included within police recorded crime data. These include studies which utilise text analysis techniques similar to those being explored in this thesis (and discussed further in Chapter 4). However, while these studies highlight the potential of both the data and available methods in providing insights for criminology, to date the questions addressed have been rather narrow and largely focused on crime detection rather than crime prevention, suggesting that this potential is currently under-exploited.

### **2.4.1 Linking offenders to offences**

The majority of research projects utilising MO data are concerned with aiding the detection of crime by improving the linking of offenders to offences. This includes projects that have utilised data and text mining techniques to create shortlists of the offenders most likely to have committed specific offences. Adderley and Musgrove (1999, 2001a, 2003,) used neural networks, a form of unsupervised clustering, to draw up a short list of offences likely to have been committed by active networks of offenders. The analysis compared known aspects of offenders' MOs with aspects of the recorded offences, including the MO, time and location. The techniques were applied to a set of bogus official burglaries (1999) sexual offences (2001a) and burglary (2003). Adderley and Musgrove concluded that data mining of MO descriptions was more successful at linking offenders to offences than manual searches by police personnel, with data mining identifying a higher proportion of matches in a shorter period of time.

In a series of publications, Canter, Bennell and colleagues used a method of classification, namely multidimensional scaling, to extract details of MO data and convert it into a series of dichotomous variables which were used to classify similar offences and link crimes committed by the same offender. Classification exercises were applied to commercial burglaries (Bennell & Canter, 2002), stranger rapes (Canter, Bennell, Alison, & Reddy, 2003) and commercial and residential serial burglary (Bennell & Jones, 2005). Yokota and Watanabe (2002) applied naive-Bayes classification to MO data, again with the aim of linking offences to a given offender. The Bayesian technique took into account the rarity of different MOs to modify probabilities. Similarly the EPSRC<sup>13</sup> iMOV Project

---

<sup>13</sup> Engineering and Physical Sciences Research Council (EPSRC)

(Bache, Crestani, Canter & Young, 2008; Bache, Crestani, Canter & Young, 2010) utilised language modelling and Bayesian probability methods of classification with the aim of predicting offender characteristics (age, gender, ethnicity and occupation (employed/unemployed)) from an offence description that included free text MO data. The model produced a shortlist of offenders to prioritise for investigation. The analysis was conducted for domestic burglary and used MO free text and feature codes. With similar aims, Cocx and Kusters (2006) conducted cluster and association analysis of police data, including MO descriptions and crime scene documentation, to assess the similarity of different offences and thereby determine which crimes may have been committed by the same offenders.

The COPLINK programme, based in Arizona (Hauck, Atabakhsb, Ongvasith, Gupta, & Chen, 2002), utilised police databases amounting to 1.5 million crimes. Published papers relating to COPLINK suggest the analysis has been dominated by efforts to detect committed crimes rather than prevent them. Although the project also demonstrates some promising developments in the analysis of free text police fields such as the ability to automatically identify 'named entities' in free text such as names and addresses, mentions of narcotic drugs (in all their diverse nomenclature) and references to vehicles (Chau, Xu, & Chen, 2002).

## 2.4.2 Comparing MOs

Other studies analysing MOs come closer to the aims of this thesis by seeking to link crimes committed in a similar way. Bowers and Johnson (2004) used MO fields in police recorded crime data to compare the MOs of domestic burglaries. Bowers and Johnson's analysis revealed that burglary pairs that were committed close together spatially and temporally shared similar MOs, lending support to their hypothesis that near repeat burglaries are committed by the same offender. The police data used in this analysis included a set of fields representing different elements of MO which are completed by the police using standardised categories. The availability of pre-coded nominal variables aided the comparison of MOs but the intensity of the analysis, which required the comparison of over six million burglary pairs, meant that the investigation was restricted to only two MO elements. The authors selected means of entry and point of entry for inclusion in the analysis as previous literature had shown these two to be the most relevant MO elements for domestic burglary. This study shows the potential of MOs as intelligence about crime but also highlights the key challenges to using this data. MOs are multi-dimensional and the use of traditional forms of analysis can necessitate limiting an investigation to a restricted set of these dimensions, this could potentially cloak important findings. Pre-coded categorical variables are more suited to statistical analysis than free text fields but these fields may not be available in all police datasets nor for all

crime types. Moreover, categories in these data will be based upon what is already known about the crime type in question, which may prevent the identification of changes in the nature of crime commission.

Other studies have compared MOs that are recorded as free text fields. However, the analysis is conducted in order to detect crimes (i.e. determine which crimes may have been committed by the same person). The Regional Crime Analysis Program based in Richmond, Virginia (Lin & Brown, 2006) used an outlier-based approach to link similar robbery offences. The OVER Project based in Oxford, UK used a range of clustering and classification techniques to link similar crimes (Oatley, Ewart, & Zeleznikow, 2006; Oatley & Ewart, 2003). Cluster analysis was also utilised by Ma, Chen and Huang (2010) to identify similar property crimes committed in China based on crime type, location type, MO, victim type and offender characteristics, while Yu, Shao and Luo (2008) applied cluster analysis to free text descriptions of crimes to devise measures of similarity and distance between offences committed by serial murderers. These studies show the potential to identify patterns and similarities in crimes using a combination of text analysis and statistical analysis. However, the studies' collective focus on crime detection means that they are based on the assumption that crimes committed in a similar way are committed by the same person. The studies, therefore, do not provide any insight into situational factors that provide opportunities for crime or causes that could potentially be manipulated to reduce the likelihood or consequences of crimes occurring in the future.

Although they did not apply data mining or related techniques Donkin and Wellsmith (2006) demonstrate the ability to extract information from free text fields to learn about the changing nature of the commission of crime events and provide intelligence for crime prevention. Donkin and Wellsmith investigated trends in car theft over time by extracting data from MO descriptions. Other than using 'find' functions within software the extraction was largely conducted manually. Their task may have been made easier with the application of computer-assisted techniques to extract specific information/text from free text fields. This would also increase the feasibility of repeating their analysis on a regular basis to investigate the progress of changes in crime commission. In addition, Donkin and Wellsmith's research was deductive in that it sought confirmation and clarification of a trend of which police officers were already aware. The use of computer-assisted techniques can also facilitate a more inductive approach which allows for a more open-ended exploration of data and identification of patterns that the analyst was not aware of at the outset.

Previous studies have, therefore, demonstrated that MO data has the potential to be used as a source of data to inform crime prevention and that available methods of content analysis of natural

language have the potential to draw out patterns from this type of data. However the two aims of providing intelligence for crime prevention and the interrogation of textual data using computer-assisted techniques have not previously been brought together. It is notable that a large proportion of the data mining-orientated studies in this field have been conducted by computer scientists and published in the journals and proceedings of computing and engineering disciplines. These studies seem a little divorced from criminological research and specifically from situational crime theories. Greater crossover between these disciplines could see these techniques being used to address a broader set of questions, including a sharper focus on crime prevention.

In addition to reviewing data analysis techniques it is also important to consider the systems that deliver analysis. The studies, discussed above, have generally concluded that computer-assisted techniques are 'better than human decision makers' at reaching conclusions from volumes of data. However, Adderley and Musgrove (2001b) report that most of the 'bespoke' data mining systems developed for police forces remain relatively underused. They suggest, that, although these systems may be better at answering certain questions, the information is limited when set against the effort and cost of their development and implementation (initial purchase, roll-out, training and supervision). They also argue that the tests against which systems are designed do not reflect the problems addressed by analysts in real life. These findings support the decision made in this research to utilise open source tools wherever possible to prepare and analyse the data. These systems are not only economical for use in a practical context but they also offer considerable flexibility in adapting the analysis to different questions.

## *2.5 Conclusion*

This chapter has introduced three core conceptual tools that aim to direct and facilitate the analysis for crime prevention. Situational crime theories show, that, in order to prevent crime, it is necessary to understand and intervene in the causal processes operating in the immediate situations surrounding potential crime events. Although Section 2.1 highlighted that there is a degree of debate surrounding the relative of influence of proximal causes compared to distal, it has been argued that situational crime theories are richer when they pay greater attention to distal causes. Situational theories should avoid depicting actors, be they offenders, victims, crime promoters or crime preventers, as slaves to their environment. Further, it should not be assumed that the same environment will produce the same interpretations and action possibilities for all actors. Even actors who share the same criminal intentions will interpret the situation differently based on factors such as their executive capacities, moral rules, experience and resources.

Thus while crime prevention involves blocking opportunities for crime, both prevention theories and practice should recognise that opportunities are not objectively 'out there': they are ecological, the product of interactions between the offender and other elements of the situation. This includes the need for offenders to recognise opportunities and perceive them as relevant to their goals and within the scope of their self-perceived resources.

The CCO offers a conceptual tool which helps to comprehensively assess the role of proximal crime causes. This includes those causes traditionally associated with situational crime theories but also encompasses the role of causes pertaining to the offender. This helps to consider how distal causes operate to influence the immediate situation by shaping offenders' *perceptions* of the situation. Methods of categorising the near causes of crime, derived from situational crime theories, already exist, most notably routine activity theory's trilogy of vulnerable target, likely offender and absence of capable guardian. The CCO builds on these categorisations by further dissecting what makes the target suitable, the offender likely and so on.

The chapter introduced scripts. Schank and Abelson's (1977) originally discussed scripts as a cognitive facility available to actors to facilitate efficient goal-directed decision making in a complex world. Researchers using scripts as a methodology have used data on performed actions to make inferences about actors' stored action templates and the factors that created and shaped them.

The CCO and scripts offer complementary approaches, the CCO provides a static picture of proximal causes of criminal events which sets the scene for understanding the dynamic interactions of causes as they are influencing actors' behaviour but at the same time are shaped by actors' perceptions, choices and enactments within scripts. In later chapters the CCO framework will be used in conjunction with scripts methodology to analyse data from previous studies (Chapter 5) and from police data (Chapters 7-9). Ekblom and Gill (2015) call for a better link between script methodologies and causal explanations, this thesis will explore whether explicitly using the two together with help to improve this link. It is feasible that the elements of script descriptions (actors, props etc.) could be mapped onto the elements of the CCO bringing the two frameworks closer together, it is their **joint** contribution to the interpretation of crime problems that is explored in this thesis.

As a description of offenders' actions MOs offer a source of intelligence on offenders' scripts, intelligence that could be used to explore how these could be interrupted through crime prevention measures. The final section of this chapter described previous studies that have analysed crime scripts and highlighted some of the practical challenges of analysing MOs as a central element of the study of scripts.

## Chapter 3 Research Aims and Objectives

Chapter 2 explored crime as a routine activity and argued that crime events are shaped by situational factors. The chapter introduced script frameworks as a method that can be used to describe and classify each stage in the commission of an offence. Through these descriptions, it is possible to identify the situational elements that interact to create opportunities for crime and for crime prevention. While scripts provide a descriptive tool, the CCO was introduced as an analytical framework which allows the identification of the causal mechanisms influencing the described script elements at the various stages of crime commission. Identification of explanatory causal mechanisms can then inform preventative action.

Applying these conceptual frameworks to the understanding and prevention of real crime problems requires the availability of empirical data. These data need to provide details of how offences were committed and the situations in which they occurred. Previous research that has adopted script frameworks has tended to focus on a relatively small number of complex crimes for which detailed data (such as that found in investigation and case files) has been available. This thesis is focused on crimes that occur in significant volumes which makes the manual analysis of detailed case files, if they are available, an unrealistic proposition. A possible alternative to detailed case files are the short descriptions of the method of commission that are recorded by police forces and attached to crime records. This MO field potentially provides the empirical data required to conduct analysis of patterns in the commission of volume crime. However, these data have limitations which can make analysis difficult. The problems with MO data are explored more fully in Chapter 6 but the central issue is that, in their original form they are free text fields that are not directly amenable to statistical analysis.

The current thesis poses three broad questions:

1. *Are police recorded MO data a potential source of actionable intelligence to inform crime prevention?* Answering this question will require a review of police data to examine the nature of the records and determine whether they contain information relevant to crime prevention, whether they contain sufficient detail and whether any inherent characteristics present problems for their analysis.

If this first question can be answered affirmatively a subsequent question will be:

2. *Can techniques drawn from content analysis, natural language processing and statistical analysis be used to identify meaningful patterns in MO data for high volume crimes?*

Answering this question will require an assessment and application of the analysis techniques available for use with textual data.

Again dependent on answering question two in the affirmative, the final broad question to be posed in this thesis will be:

3. *Do script and CCO frameworks add value to the analysis and interpretation of patterns in MOs?* Answering this question will require the application of frameworks to the analysis, the interpretation of the empirical data and a critical reflection on the extent to which frameworks have helped to guide, organise and interpret findings.

In line with these three questions, the research objectives are organised into three phases, data review, data analysis, and interpretation. The objectives for each of these phases will be detailed in following sections.

The data used in this thesis are police recorded crime data relating to thefts from the person and robberies of personal property. These crimes have been selected because they are reported with sufficient frequency to explore the issues of analysing the MO patterns of high volume crime. While legally, theft of the person and robbery of personal property are treated as very different offences, in the practicalities of their commission these crimes have much in common. Analysing two quite similar crimes allows the research to test whether computer-aided textual analysis techniques can help discriminate between crime types. This sets up a more challenging test than analysing two distinctly different crime types. Theft from the person and robbery of personal property are both crimes that involve some level of interaction between the offender(s) and the victim(s) which presents further challenges for the analysis of text in attributing actions and intentions to actors.

The following three sections will describe each of the research phases in turn. Although presented as sequential, in reality there was considerable feedback and iteration between these phases.

### ***3.1 Data Review***

The data review phase will assess whether police recorded crime data contains useful information regarding the nature of MOs. In line with the overall aim of the thesis, this assessment is concerned with whether police MO descriptions have the potential to generate actionable intelligence for the prevention of crime.

Ratcliffe (2008) distinguished between *data*, records of observations; *information*, data that has relevance to a specific purpose, *knowledge*, which is generated by the interpretation of information

and *intelligence* as knowledge that has the potential to guide or inform action. In line with this definition, the review will first consider the observations as *data* to assess whether they are sufficiently reliable to form the basis for further analysis and interpretation. The data review will consider the extent to which factors such as recording errors, missing data and consistency of information across records, present threats to the quality of any subsequent analysis. The data review will not simply be searching for errors, there are likely challenges with analysing free text fields even when they have been accurately completed. These can include similar crimes being described using a very different vocabulary but can also mean that crimes that are quite different will appear similar because vague terminology has been used. The review will also assess the degree of confidence that can be attributed to MO descriptions as valid accounts of the crimes that occurred and identify factors that may reduce this confidence. In other words, how closely do MO descriptions reflect the events that occurred? This will include a focus on the social processes that shape crime reporting and crime recording and a consideration of the bias that these factors may introduce.

Chapter 2 explored the potential of situational crime prevention to block the situational conditions which cause crime, the review will, therefore, consider whether the MO descriptions contain sufficient data relating to these conditions for them to be regarded as a source of *information* for crime prevention. This thesis is concerned with the generation of intelligence for crime prevention but MO descriptions are traditionally recorded for the purposes of detection. Consequently, the review will consider the extent to which there is overlap between the information requirements of these two purposes.

In summary the data review will pose the following questions:

- What are the data like?
- To what extent do recording errors, missing data and recording inconsistencies reduce the reliability of the data for analysis?
- Which characteristics of free text data present challenges for data extraction?
- To what extent can the MO descriptions be regarded as valid and sufficiently complete accounts of crime commission?
- What factors shape and potentially bias the nature of MO descriptions?
- What are the implications of the identified data characteristics for analysis?
- Do MO descriptions contain information that is relevant to crime prevention?

To fully answer the question of whether MO descriptions contain *actionable intelligence* for crime prevention requires a consideration of methods of analysis and interpretation. These questions are considered in phases two and three of the research.

### *3.2 Data Analysis*

Assuming that the data review concludes that MO descriptions contain data that is sufficiently reliable and valid to continue with analysis, the analysis phase will begin by preparing the data and converting the text fields into a format amenable to statistical analysis. In the event that data review reveals that the data has many errors, is patchy or poorly structured, the challenge will be to determine whether these problems present insurmountable problems for data processing and analysis.

An important issue is that of simplifying the data in the MO descriptions. Free text data fields give those recording information considerable flexibility in the choice of vocabulary used to describe an event. Humans are well equipped to interpret this flexibility but it poses problems for computer automated approaches. Data preparation will aim to reduce this flexibility through the application of techniques drawn from content analysis and natural language processing<sup>14</sup>. The research will assess whether these methods provide time savings or other advantages compared to 'manual techniques' of data cleaning. This phase will also seek to identify whether any of the data issues identified during the data review remain after the data processing is completed, for example, does the presence of multiple actors in an MO description pose problems for the analysis.

Data preparation will be followed by data analysis. The analysis will help to determine whether the reduction of flexibility in the data has retained sufficient detail to reconstitute a picture of what happened within each crime. One test of the degree of information loss is to identify whether statistical analysis can successfully discriminate between crimes that are known to have different methods of commission because they have been differently classified, i.e. to distinguish between theft from the person and robbery of personal property.

The data analysis will aim to identify patterns in the MO descriptions and to identify groups of crimes with similar MOs. Given that the starting point for this analysis is free text data, it will be important to ascertain whether the vocabulary used to describe MOs is a sufficient basis to make these distinctions and comparisons. As noted in Chapter 2, Green (1989 in Woodhams and Toye,

---

<sup>14</sup> Natural language processing is a branch of computer science that deals analysing and understanding the languages that humans use.

2007) proposed that some uniformity is found in patterns of action because of a close relationship between goals and outcomes. If the data analysis reveals variation in the ways in which theft from the person and robberies of personal property are conducted, this will show that they are equifinal actions (Abelson, 1981) with different options available to commit the same goals. This is an important finding from the point of view of crime prevention as different routes to the same goal may require or be amenable to blocking via different crime prevention methods but at the same time the availability of different routes may provide channels through which crime prevention can be deliberately circumnavigated by offenders (tactical displacement). In Ratcliffe's terms, the identification of patterns can therefore be seen as elevating the status of the data to *information* for crime prevention.

In summary, the analysis phase will apply techniques from natural language processing and content and statistical analysis in order to answer the following questions:

- Can techniques from natural language processing and content analysis convert free text MO fields into a format amenable to statistical analysis?
- To what extent are semi-automated methods of data preparation effective, efficient and reliable compared to manual processing?
- To what extent does the pre-processing of data reduce its reliability and validity?
- To what extent does the pre-processing of data result in data loss?
- What residual data problems remain following the processing phase?
- What patterns can be identified in the methods used to carry out offences of theft and robbery of personal property?
- Can the analysis techniques successfully discriminate between crimes that have been differently classified?
- Can the analysis techniques identify groups of crimes committed using similar MOs?

### ***3.3 Interpretation of Findings***

Sutton (2012) made the observation that crime opportunity theories provide nothing more than after-the-event descriptions of crime. MO descriptions, from their very conceptions, are exactly that. To describe a crime that has been committed is central to the point of recording this data. However, this research seeks to explore whether MO descriptions can be used for broader purposes. Can the analysis of MO data provide models that elucidate the common patterns in crime commission? Then, with the application of conceptual frameworks, can these models provide the basis for explaining the nature and shape of these crime events and for identifying preventative interventions

through which similar crimes could be blocked in future? This phase of the research, therefore, aims to examine whether script and CCO frameworks aid the interpretation of methods of crime commission.

As discussed in Chapter 2, proponents of script frameworks argue that scripts can provide a methodology to classify and give structure to information about the process of crime commission. The findings from the crime analysis will be mapped to script frameworks with the aim of identifying the scenes that are performed to accomplish theft from the person and robbery of personal property and to reveal the actions taken to accomplish each 'scene', the props (e.g. tools and weapons) used to complete the crime and the cast members (victims, accomplices) engaged in the offence. Using scripts to interpret MOs can help to identify which scenes are essential to the completion of a script, and which conditions, actions, props and cast members are apparent prerequisites or facilitators for the completion of each scene. This phase of the research will seek to establish the degree of diversity that is available to offenders committing these crimes by determining whether several alternatives to crime completion exist or whether the majority of offences appear to be committed via similar steps.

Crime scripts are not committed in isolation. It was been noted above that cast members may include crime victims, and these actors have opposing goals to the offender(s). Systematic analysis of scripts should help to reveal how offender(s) interact with other actors, for example to take advantage of a crime promoter's own scripts, and to forestall or handle interruptions from crime preventers.

Scripts represent goal-directed action (Cornish & Clarke, 1994), script frameworks, therefore, aim to interpret an offender's goals and decisions through his/her actions. Dissection of scripts into their individual scenes aims to reveal offenders' lower as well as their higher level goals. From the perspective of situational crime theories a key set of higher level, overarching goals are those to reduce effort, reduce risk and increase reward. Lower levels goals direct the behaviour through which these higher level goals are achieved, and this behaviour will include the offender's manipulation of elements of the crime situation. The examination of offender actions should, therefore, reveal which elements of the crime situation have facilitated the commission of crime.

The CCO was also introduced in Chapter 2 as a framework that aims to help with the consideration of the proximal causal factors that shape the crime situation and in turn, shape the occurrence and nature of criminal events. Through comparisons of different crime scripts, the research aims to

demonstrate Abelson's observation that small differences in situational context i.e. in the proximal causal factors, can lead to significant differences in script tracks.

By applying script and CCO frameworks together, the research aims to identify and classify offender actions and the elements of the crime situation that have directed and shaped offender decisions, stage-by-stage; these elements can then be related to proximal causes in the crime situation(s). The identification of relevant causal factors provides the *intelligence* required for blocking crime opportunities and preventing crime. This phase of the research also aims to explore the impact, on the nature of offending, of potential changes to the elements in the crime situation, including changes that are an indirect result of other factors and changes that have been made intentionally for the purpose of crime prevention. The research will then critique whether and to what extent, this integration of script and CCO frameworks has been helpful in the understanding of crime causation and the identification of opportunities for crime prevention.

It is anticipated that the availability of a large volume of data will proffer advantages in being able to identify the range and diversity of scripts and the range of crime situations in which they are committed. At the same time, the data may present limitations in the level of detail provided about the crimes and the extent to which all crime scenes are comprehensively covered in the descriptions. A central aim of this thesis is, therefore, to assess whether high volume crime data contains sufficient detail to enable the systematic application of script and CCO frameworks. Applying causal frameworks to the data raises further questions about the MO descriptions. These are more conceptual than those posed in Section 3.1. In applying the CCO framework the following will need to be considered. How much information about the range of causes in the crime situation is available within the MO descriptions? Given their brief nature, how likely is it that they will include information relevant to understanding causation?

The extent to which goals can be interpreted through empirical traces of actions alone may be limited. The same may also be true for causal mechanisms. There is a potential risk of jumping to conclusions about offender decisions and the causes that shape them, from secondary data that does not directly involve the offender, in contrast to research that questions offenders directly. To provide context for the findings, the research will use findings from previous research into theft from the person and robbery of personal property to provide insights into the MO patterns. The literature to be consulted includes research that has explicitly applied script frameworks and studies that have not, but none of the studies applied the CCO framework. Script and CCO frameworks will be applied retrospectively to studies in order to systematically classify the knowledge that is

currently available about the nature of crime commission and the factors that shape offender decisions and actions.

Applying the frameworks to the literature will provide a further test of whether the joint application of CCO and script frameworks helps to improve understanding of the nature of crime commission. The existing literature includes a wealth of in-depth studies, if scripts and CCO frameworks are not helpful in structuring the detailed and rich information presented by these studies, it may seem unlikely that they will be helpful where the information is brief and patchy. Analysis of the literature will also provide support for the analysis of police data in enabling the researcher to make 'reasonable assumptions' about missing information.

Theoretical frameworks will be used to provide a link between the existing literature and the new empirical findings and therefore, relate the findings to broader knowledge about crime causation and prevention. This will help to make sense of the empirical findings and to make inferences about offender decision making with greater confidence. It will allow the triangulation of the findings emerging through empirical analysis.

In summary, the interpretation phase aims to apply script and CCO frameworks to the data in order to answer the following questions:

- To what extent do frameworks enable the identification of the scenes performed in the commission of theft from the person and robbery of personal property?
- Which script scenes are essential for the commission of these crimes?
- Can alternative routes to crime commission be identified?
- Do the frameworks assist with the identification of the resources, props and cast members involved in each script scene?
- Do the frameworks help to identify and understand how offenders handle script clashes, including direct interferences into their crime commission?
- Do MO descriptions contain sufficient information to inform a consideration of crime scripts and crime causation?
- To what extent do the frameworks help to extrapolate offenders(s)' high level and low level goals?
- To what extent do the frameworks reveal the elements of the crime situation that have facilitated or hindered crime commission and/or been manipulated by the offender?
- To what extent do frameworks inform an understanding of how variations in situational conditions affect the nature of crime commission?

- Do script and CCO frameworks add value to the interpretation of empirical data relating to crime commission?
- Do script and CCO frameworks add value to the interpretation of existing literature relating to crime commission?
- Do script and CCO frameworks enable the triangulation of empirical findings with those drawn from previous research?

### *3.4 Summary*

This chapter has outlined the research aims and objectives in detail. The data analysis will be conducted on police recorded crime data relating to theft from the person and robbery using a combination of computer-aided content analysis and statistical procedures. These crime types have been selected because they occur in sufficient volume, allow the testing of tools to distinguish between related crime types and provide an insight into the challenges of analysing interaction between different actors. The conceptual frameworks will be applied to MO data to consider their use in organising findings, identifying causal factors and extracting intelligence that can inform crime prevention. To triangulate the findings emerging from the empirical data analysis the frameworks will also be applied to the body of existing research on these crime types.

The final sections of the thesis will reflect on the implications of the findings. This includes identifying directions for future academic research and theoretical development but also considers the findings' applicability to practice. The thesis adopts an applied practical perspective throughout and an overarching principle is to produce findings that have practical applicability to inform both the production of intelligence via crime analysis and the implementation of appropriate interventions on the basis of that intelligence.

## Chapter 4 Methodology

As outlined in Chapter 3, this thesis aims to assess whether police recorded MO descriptions are a potential source of actionable intelligence to inform crime prevention and whether script and CCO frameworks add value to the analysis and interpretation of patterns in MOs. These aims were broken down by a series of objectives requiring a review of the data, data analysis and the interpretation of and critical reflection of the findings.

These research questions were addressed using a large dataset of textual data extracted from police recorded crime data. The textual data were processed through several stages to make them amenable to statistical analysis. Statistical analysis was then conducted and finally the results were interpreted through the lens of conceptual frameworks. This chapter will begin by describing the data that were used and then outline the challenges that are commonly encountered when analysing textual data. This will be followed by a description of the techniques that were used to prepare and analyse the data and the rationale for selecting these approaches. The key stages in processing and analysing the data are described in the sections below.

### *4.1 An Introduction to the Dataset*

The data consisted of 29,140 cases of robbery and theft from the person committed in a single police force over a period of 36 months (January 2007 to December 2009). The analysis considered theft from the person and robbery of personal property. There are three advantages to the selection of these crime types.

Firstly, these crimes occur with sufficient frequency to attempt to answer the research questions with reference to volume crimes. Secondly, these two crime types can be considered scripts within a broader meta-script (theft of property) and pro-script (thefts from the person). However, the distinctions in how the two crime types are carried out produce two distinct crime scripts. Despite these differences the distinction between theft from the person and robbery is not always clear cut and it can be challenging for crime recorders to determine which classification to apply. This was evident from the inspection of MO descriptions, which included the documentation and justification by police or civilian personnel of classification decisions and on occasion the adjustment of assigned classifications, an issue discussed further in Chapter 6. For this reason, one of the analysis aims was to investigate whether text analysis could successfully discriminate between crimes that had been differently classified. It was felt that distinguishing between two fairly similar crimes would be a more challenging test than attempting to differentiate between two very different crimes.

Finally, these crimes involve direct interactions between different actors as victims and offenders are co-present during the course of the offence. Consequently, MO descriptions for these crimes include details of both offender and victim actions, and in many cases interaction between the two. These features present interesting challenges for analysis of the text.

Access to the data was negotiated via a data processing agreement. The agreement stated that the name of the force should not be published in any research outputs; in line with this agreement the force will be referred to as Force A throughout the thesis. The selection of police force was a convenience sample, based on an existing working relationship with the force. However, the convenience of selection is not thought to have biased the results of the study or to have affected the ability to generalise findings to other UK forces. Based on previous experience with working within MO data, Force A's recording of this information is representative of that found in other UK police forces. In addition, Force A is amongst the largest in the UK (both geographically and in terms of population) and includes large urban centres as well as suburban and rural areas which should ensure that a full range of different MOs are included in the data.

Data were received for all crimes of robbery and theft from the person committed across the force area during the three year period (January 2007 to December 2009). The database included the following variables:

- A ten character unique crime reference number.
- Home Office Crime Classification (2009 codes<sup>15</sup>).
- Date and time of commission – comprised of two sets of variables, 'first' and 'last' to accommodate cases where exact time or date of commission was unknown or uncertain.
- Location of the crime – indicated by postal address, ordnance survey grid references (1m resolution) and administrative boundaries (ward, police beat, and police division).
- Location code – brief text field. This consisted of one or two words to describe the nature of the crime location e.g. road, licensed premise, public footpath. This field was merged with the longer MO description field for textual analysis.

---

<sup>15</sup> Note that Home Office codes were updated in 2012.

- Brief MO Free Text Descriptions averaging 50 characters and reading up to 200 characters.

The analysis reported in Chapters 7 to 9 has predominantly utilised the MO free text description and the location code. However, Chapter 10 discusses the potential to use additional data fields to explore the relationships between MO patterns and other aspects of crimes. This includes comparing MO patterns with fields found directly in the data (e.g. data and time) and by using geo-referenced data to link to other variables such as the socio-demographic characteristics of the crime location.

Table 4.1 shows that the majority (60%) of the 29,140 cases in the data were robbery of personal property, just over a quarter (27%) were theft from the person. The remainder of cases were assault with intent to rob personal property (6%), attempted robbery of personal property (6%), attempted theft from the person (0.4%) and six cases of conspiring to rob personal property. The higher numbers of robbery offences in the data may reflect the higher reporting levels for these crimes. Based on survey estimates, the police nationally come to know about an estimated 45% of robbery of personal property offences and 33% of theft from the person offences (Chaplin, Flatley, & Smith, 2011). Of the robbery of personal property offences in the data, 9.5% of cases were attempts. This compares to 1.5% of all theft from the person cases. In the case of theft from the person, the low proportion of attempts may represent the fact that a victim will be unaware that an offender has attempted to take their property.

**Table 4-1 Home Office crime classifications included in the data**

<b>Home Office Classification</b>	<b>Crime Classification</b>	<b>N</b>	<b>%</b>
<b>34/3</b>	Robbery Personal Property	17634	60.5
<b>39/0</b>	Theft from Person	7949	27.3
<b>34/4</b>	Assault with Intent to Rob Personal Property	1751	6.0
<b>34/3</b>	Attempted Robbery Personal Property	1687	5.8
<b>39/0</b>	Attempted Theft from Person	113	.4
<b>34/3</b>	Conspire to Rob Personal Property	6	.0
	Total	29140	100.0

MO descriptions had been completed for the majority of cases with only 20 missing a description. Cases were inspected to identify crimes with multiple victims. This included identifying cases with completely identical MO descriptions and cases that were cross referenced with other cases in the file. The National Crime Recording Standard (NCRS, (Home Office, 2015c)) stipulates that police forces take a victim-orientated approach to recording crime and instructs that forces must record one crime per victim, therefore, although each 'linked case' had a unique crime reference number they related to the same event (maximum seven times). As noted in Chapter 3, the unit of analysis

for this research was the event and not the victim: consequently any 'duplicate' MO descriptions were removed<sup>16</sup> so that only one case for each event was retained. After the removal of duplicate cases and those cases missing an MO description, 22,193 cases remained in the analysis file.

The sections below describe the approaches used to analyse these cases. This begins with a discussion of the challenges of analysing textual data (a more detailed review of specific problems in the police records analysed follows in Chapter 6). This is followed by an introduction to the disciplines of content analysis, natural language processing and text mining that have informed the analysis of these data. Within these three disciplines, it is established practice to refer to a database of text entries as a 'corpus' and each of the individual entries (in this case each MO description) would be regarded as a distinct document; the additional fields (date, crime type etc.) in the dataset would be regarded as attributes of each document. However, for the sake of simplicity and consistency with other studies in criminology, the analysis will use terminology more frequently adopted in crime analysis. Therefore, the terms 'data' and 'database' will be retained in place of corpus and the term 'case' will be retained in place of document.

## *4.2 Text as Natural Language*

In order to identify distinct patterns of MO descriptions it was necessary to use the vocabulary used to describe offences as a means of distinguishing between different types of MO. However, the nature of the language used in free text descriptions can pose numerous problems for analysts, particularly where a large volume of data are being considered. MO fields can be regarded as an instance of 'natural language.' Crowston, Allen, and Heckman (2012) state that "Natural language provides a nearly infinite variety of ways to structure and convey meaning using differing syntactical structures, synonymous terms and embellishments (adverbial and adjectival clauses)" (p.539). Natural language contrasts with structured and formal languages such as computer programming languages. Structured languages are deliberately constructed to remove any ambiguity and to provide a one-to-one correspondence between vocabulary and meanings (e.g. commands and actions). This is not the case with the natural language used in MO descriptions, where, not only are there many ways to commit an offence, there are clearly many ways in which the same offence can be described.

---

<sup>16</sup> While clear duplicates could be removed, where the MO descriptions related to the same event but used different words, or where cross reference information was not available, they could not be easily be identified or removed.

The richness of natural language used in free text fields makes automatic matching of similar text difficult. Subtle differences in description can relate to broad differences in activity. Conversely, activity that is broadly similar can be described using a wide selection of vocabulary. Different terminology is not a matter of inconsistency, it is a central feature of natural language: for example, the act of theft within robbery can be described as 'removing,' 'pulling,' 'taking' and 'grabbing' and so on. As experts in natural language, human interpreters are better equipped to deal with this flexibility and subtlety than automated systems— although often the different uses of a term can still be unclear. In particular the human is better attuned to whether variations matter for the analysis: for example, are synonyms stylistic choices or are they substantive differences? In relation to MO descriptions, the challenges lie in identifying that two crimes relate to a similar method when there are semantic differences in the terminology used to describe them, and identifying that two crimes have been committed through different methods when they may have been described using generalised/vague terminology.

The flexibility in natural language presents several problems for automated approaches to the analysis of vocabulary including:

- The option for the author of the text to choose between different words with closely related meanings using:
  - Synonyms
  - Hyponyms: Words that designate membership of a class (e.g. *scarlet* is a hyponym of red)
  - Hypernyms: Words with a broad meaning that more specific words fall under (e.g. a *colour* is a hypernym of red)
- The use of words with ambiguous meaning that can only be understood within their context. This includes homographs, words which share the same spelling but have different meanings.<sup>17</sup> For example, '*stalk*' can take the form of a verb meaning to follow or harass a person but it can also take the form of a noun referring to part of a plant.

---

<sup>17</sup> Often also referred to as homonyms, technically words which have the same spelling and same pronunciation but different meanings.

- Multi-word phrases which must be analysed together, particularly where these phrases are idioms for which all words must be analysed together to understand their meaning (e.g. to pick up a passenger).
- Phrases where the co-presence of a negating phrase or word, significantly changes the meaning e.g. no injury.
- Proper nouns (personal names, street names, brand names etc.). The use of proper nouns can cause problems with matching similar descriptions if one case uses a noun and another a proper noun. For example the use of 'Euston' or 'Lime Street' without mention of 'train station'. Further, determining the meaning of some proper nouns can require local knowledge.
- Misspelled words- which introduce unintended variability to the data.
- Instances where the same word is used in both singular and plural forms.
- Instances where the same word is used, but in different tenses (the victim walked compared to the victim was walking).
- Commonly used English words which do not add to understanding. These words, such as 'and', 'a' and 'the' are commonly referred to as stopwords.
- Words and phrases that are not relevant to the purpose of the analysis. These can produce 'noise' which interrupts the analysis.

Police data represent what has been termed a 'local grammar' a subset of natural language which represents the communication of a specific community with a common purpose (Crowston et al., 2012; Liddy, Jorgensen, Sibert, & Yu, 1993). The specialised vocabulary utilised within a local grammar can reduce the flexibility of language and reduce some of the ambiguity encountered by analysts. For example, homonyms will exhibit a more restricted set of possible meanings that are relevant to that community, in this case to crime and policing. While ambiguity is bounded by the local grammar, understanding and analysing this text requires a comprehension of the shared set of meanings and concepts applied within the domain. Further, this comprehension must be built into any automated methods designed to assist the human analyst. While it is possible to tailor systems to local grammar it should be borne in mind that too close attention to local terminology could restrict the broader applications of an automated system to other contexts.

### *4.3 Methods of Analysing Natural Language*

A number of analytical techniques were applied to the MO text fields in order to answer the research questions. These comprised:

- content analysis tools to normalise the text in free text fields, classify segments of text and identify the frequency with which these segments occur in the data;
- cluster analysis to identify patterns in MOs, to discriminate between crimes committed using different methods and to identify groups of crimes committed using similar MOs.

The rationale underpinning the selection of these methods is underpinned by a number of academic disciplines which have addressed the challenges of free text and natural language analysis, notably content analysis, natural language processing and data/text mining. This section will discuss the approaches offered by these disciplines and describe how they have been employed in this thesis. In the main, these disciplines provide statistical approaches to textual analysis. Early explorations of the data and approaches for analysis revealed that qualitative analysis would be very time consuming. A thorough qualitative analysis of a sample of 100 cases took over a day, meaning that analysis of the full dataset would likely take several months. This time scale may have been feasible for a one-off academic analysis, but was not thought to be a practical approach for practitioners, particularly if the analysis is to be routinely conducted. Consequently, quantitative methods that provide more automation were sought. The strategies reviewed do not represent an exhaustive list as the landscape of textual analysis is both broad and ill defined. Commonly used methods for analysis of free text or natural language include:

- Content analysis (which can be quantitative or qualitative);
- Natural language processing;
- Text mining;
- Narrative analysis (again, this can be quantitative or qualitative).

These methods are far from being mutually exclusive. The lines between different domains of text analysis are blurred and many functions are commonly carried out under the umbrella of more than one analytical genre (although, different terminology is often used to describe the same function). It should be noted that, although all of these techniques are designed to be applied to natural

language, the specific term 'Natural language processing' is more usually reserved for applications of computational techniques (e.g. machine learning algorithms) to natural language.

Distinctions can also be drawn between manual and automated techniques and between symbolic and statistical approaches to text analysis. The distinction between manual and automated processes should be regarded as a continuum which ranges from completely manual through assisted manual, human automated to fully automated. Most manual analysis tasks are now facilitated, to some extent, by technology even if the technology does little more than organise the results of the analysis. Computer-assisted approaches are increasingly being adopted to assist manual procedures, such as with Computer-Aided Qualitative Data Analysis (CAQDAS). At the other end of the scale there are few, if any, analytical tasks which do not require some level of human input. For example, even automated spellcheckers require the users to make choices between alternative suggestions.

Symbolic approaches are knowledge-based and apply human developed rules to the data and require an understanding of the meanings within the text. Statistical methods of text analysis stem predominantly from within the fields of text mining and natural language processing. Such methods identify patterns within the text but do so without understanding the meaning of the text.

A distinction can be made between analysis that uses text as an indicator of other processes and the linguistic traditions that explore text as the *object* of analysis. Disciplines such as conversational analysis, qualitative narrative analysis and discourse analysis, focus on how and why text is put together in a particular way. These distinctly differ from the tradition, adopted in this thesis, of treating text as a window into the social world experience (Tesch, 1990).

### 4.3.1 Content analysis of natural language

Content analysis aims to systematically classify segments of text (the unit of analysis may be words, phrases or other units) according to substantive themes that are, usually, pre-determined by the researcher(s). Coding is designed to ensure that the segments coded to each category share the same or similar interpretation (Weber, 1990). This thesis aimed to analyse volume crime data: the number of cases in the data necessitated automated or computer-assisted analysis. However, as noted above, the richness and flexibility within natural language presents a number of complications to automated processes, notably in determining whether two segments of text genuinely have the same or similar meaning. Content analysis provides a number of tools that can assist with these complications. Content analysis can include both symbolic and statistical approaches to data analysis. The allocation of codes can be determined by the meaning of the text or, as is increasingly

the case with computer-assisted methods, the process can be regarded as the application of a set of rules without the need for the computer to recognise the meanings on which the coding frame is based. These rules are embedded within a coding frame which stipulates the conditions for assigning a code to any given segment of text. Coding frames can be regarded as dictionaries (or ontologies) with a definition provided for each conceptual code. Standard, “off-the-shelf” dictionaries are available which can considerably reduce the burden of coding frame development; however, the analysis benefits when standardised tools are revised and tailored to suit the data and research questions in hand. Any work undertaken to refine dictionaries (as in this thesis) can then act as an improved starting point for future work on similar databases.

A commonly used approach applied to large volumes of data, which formed the starting point of analysis for this thesis, is to identify the frequency of words in a segment of text based on the underlying assumption that terms used most frequently will reveal the subject matter of the text. This approach, known in natural language processing as N-gram analysis, is also referred to as a ‘bag-of-words’ technique as the analysis can be conducted without a concern for word order. In this approach, the ‘codes’ bear a close resemblance to the words in the raw data. In an automated analysis, words will only be counted as ‘similar’ if they are exactly the same. To cope with the inherent flexibility of natural language it is, therefore, necessary to standardise and simplify the text (a basic forms of coding) prior to analysis. The processes applied to standardise the text prior to analysis are depicted in Figure 4.1 below.

Initial steps in standardisation might require the identification and correction of spelling errors, for which there are readily available English language checkers. It is also established practice to omit stop words, i.e. words that occur so frequently in the language that they are not expected to add anything to the analysis. Stopword lists commonly include ‘function’ or ‘structure’ words such as *the*, *is*, *at*, *which* and *on*. Open source, validated lists of stop words are readily available for use, although, these lists may require inspection and modification to the specific research question and data sources used.

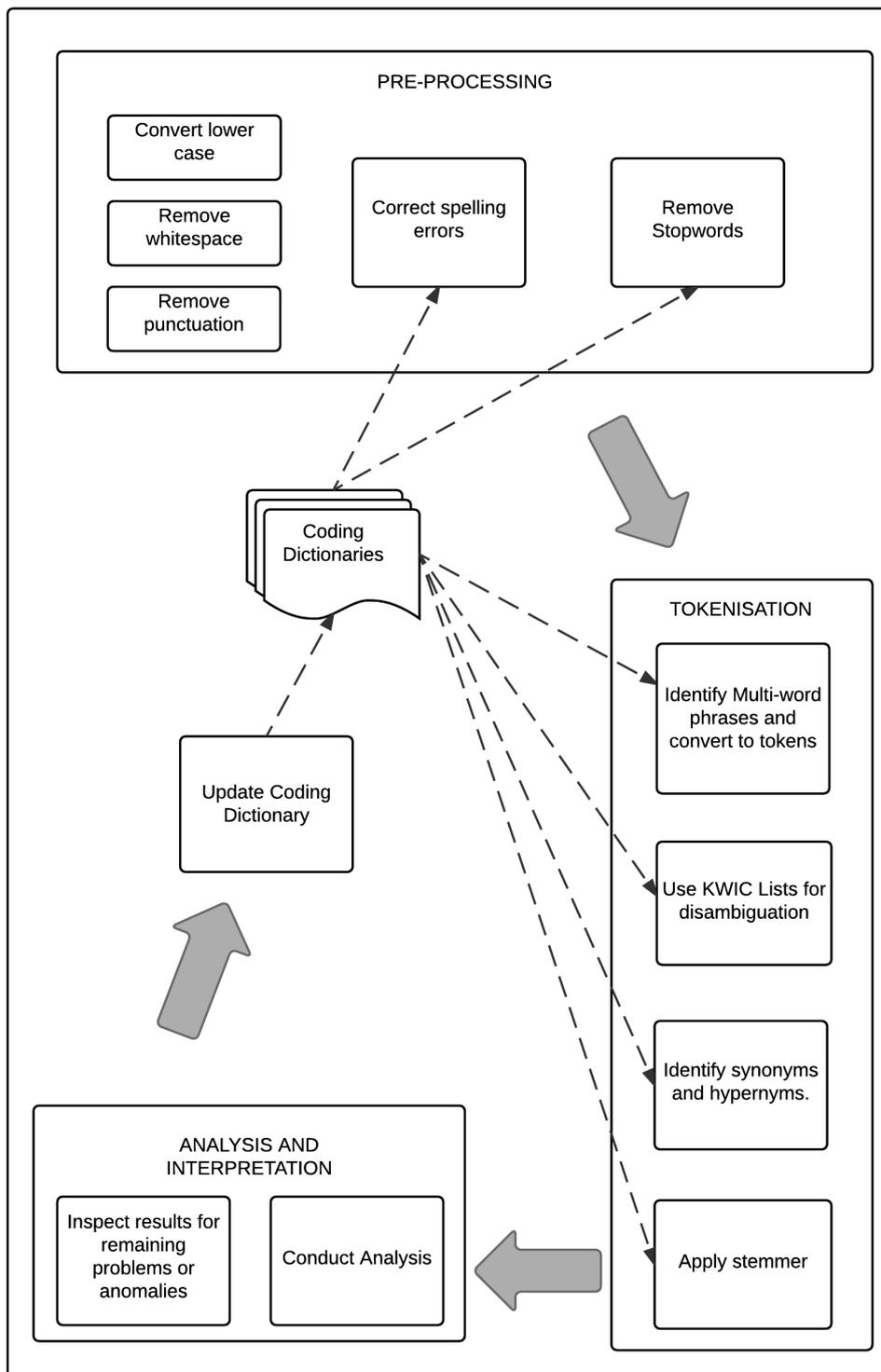


Figure 4-1 The data preparation process

The next essential step in standardising the text is a process known variously as stemming or tokenisation. This thesis will adopt the terms tokenisation and tokens in acknowledgement of the fact that the standardisation of text included steps additional to basic stemming. In stemming, a

word is reduced to its stem, or root form. A lemma is the canonical form of a set of words (Leetaru, 2012). For example, walk, walks, and walking are defined as different lexemes that share the same lemma i.e. walk. In the analysis of text, words are frequently reduced to a stem that represents this lemma, by removing affixes, although this approach can be crude and does not successfully standardise all lexemes (run and ran, for example). Fortunately, algorithms for stemming have been studied in computer science since the 1960s and there are a number of sophisticated algorithms available for use. The analysis in this thesis utilised the widely used English language version of the stemmer 'Snowball' (Porter, 2001) which has become regarded as the standard approach to stemming in a wider range of languages (Willet, 2006).

Individual words may not be the most appropriate unit of analysis as it is not always possible to derive the meaning of phrases and idioms from their component words, for example to "change my mind" constitutes a single unit of meaning that does not directly relate to the individual words. In a similar vein, the presence of negating words can radically alter meanings within text: for example, in the case of police data it is important to distinguish 'no force used' from 'force used'. Again, the development of dictionaries has provided flexible approaches for handling phrases, but these approaches are imperfect and dependent on the availability of an appropriate dictionary for the language/domain in question. An alternative approach is to create a list of the most common multi-word combinations within a dataset. The advantage of this approach is that the list is based on the corpus to be analysed and is therefore of greater relevance to the data in hand. This thesis utilised the online software TerMine (Frantzi, Ananiadou, & Mima, 2000)<sup>18</sup>, developed at the University of Manchester. This program can be used to automatically identify frequently used multi-word phrases in a dataset. Although such software offers considerable time savings, ultimately, the inspection of this list and the decision regarding which phrases need to be treated as a single unit of analysis must be conducted by the analyst. Once the multi-word phrases relevant to the dataset have been identified, they can be treated as if they were a single word simply by conducting a 'find and replace' to swap the spaces between words with an underscore (e.g. 'no force' becomes 'no\_force'), these phrases are then subsequently treated as one token.

As noted above, natural language is flexible, meaning different vocabulary can be applied to describe similar events. To treat different, but related vocabulary, as the same requires extending the process of tokenisation to identify and group together synonyms, hypernyms and hyponyms and other words and phrases that indicate the same or similar class of object or action. This process is more

---

<sup>18</sup> <http://www.nactem.ac.uk/software/termine/> (last checked 6/08/2015)

akin to traditional coding, as applied during qualitative analysis of text, in that groups of words that present a similar meaning are labelled under the same 'code.' Throughout this thesis, the term 'token' refers to a code applied to a group of words, terms or phrases that share a similar meaning. Again, open source thesauri are available to assist with the identification of synonyms. This includes WordNet, an online lexical database of English developed at Princeton University<sup>19</sup>. Alternatively, researchers can modify existing dictionaries or develop entirely new dictionaries that are specific to their research concerns. The identification of words with similar meaning is not an objective process and is guided by the analyst's interests which should themselves be informed by the research questions. For example, one research project may simply be interested in identifying physical assaults, whereas, another may need to draw distinctions between different types (kick, punch, use of weapon etc.) or severities of physical assault. These distinctions need to be reflected in the coding frames. Although the WordNet database was used to assist in the identification of synonyms, much of this work was accomplished via inspection of the data: creating a bespoke dictionary was tested and refined through a number of iterations.

A more challenging stage of data preparation is the correct identification and standardisation of homonyms, defined above as words with the same spelling but different meanings which, under a basic frequency count, could be subject to misclassification. This requires a process known as disambiguation to clarify whether similar text really does warrant the same interpretation. This process is harder to automate, although software can produce a list of known ambiguous words and highlight them for further inspection. These potentially ambiguous words or phrases can then be displayed in the context of the surrounding text. This is important as the meaning of some words can only be determined from the rest of the phrase in which they are contained. Key Words in Context (KWIC) lists, also known as a concordance view, assist in the removal of ambiguity and the accurate categorisation of text segments that can be replaced with an unambiguous tokens. Such amendments can be time consuming and although, once again, dictionaries can provide rules for disambiguation for use in automated processes, there is still a risk of error. The disambiguation of homonyms is aided, to a degree, when working with a local grammar, in this case crime reports, rather than truly natural language, as the potential number of meanings for a word is restricted by the context of being recorded by the police as part of a crime report. For example, within police data, the potential meanings of 'stalk' are restricted and so there is a greater probability, but no

---

<sup>19</sup> <https://wordnet.princeton.edu/> (last checked 6/08/2015)

certainty, that it will refer to the verb form. When the crime type of a case is known, then the meanings within the text of an individual MO field are likely to be yet more constrained.

The aim of quantitative content analysis is to convert coded segments of text into numerical variables which represent features within the text. These numerical variables can then be subjected to traditional statistical analysis. This allows the generation of inferences about the importance of certain themes in a text, and for comparison between different texts. This can include the analysis of differences between texts produced at different points in time. When numerical variables are produced based on tokens, care needs to be taken about inferences made based on multiple mentions of the same token; a token does not necessarily carry the same weight each time it is used. It is important to consider the unit of analysis, and in this research, the occurrence of a token within each case is important – repeated tokens within the same case are not treated with any greater importance. While frequently used tokens are important, the majority of tokens, in any document, are used sparsely; they may occur only once (hapaxes) and for some research questions, including the questions posed by the current research, important insights may be gained through a consideration of infrequently used tokens or absent tokens. This relates to Information Theory (Shannon, 1948) which maintains that rarer occurrences contain more information than commonly repeated occurrences.

Word frequencies do not provide any information about the relationships between words. It was noted above that, for N-gram analysis, consideration of word order is not necessary; however, it is possible to extend this technique through the inspection of the sequencing of words within a segment of text. Words that are used together (directly) or in proximity may help to inform research hypotheses and practice-orientated hypotheses (Chainey, 2014). The inspection of words that most frequently precede or succeed a specific word further help to provide context and interpret the meaning of word occurrences in the data. Data can be queried to identify collocations and concurrences of words within texts. As with individual word frequencies, it may be the rarely occurring or unusual combinations that are of interest.

### **4.3.2 Data and Text Mining**

The analytical techniques used in this thesis could be described as methods of data mining. Data mining aims to ‘discover’ knowledge from raw data, and has been developed to exploit the growth in large volume datasets. Text mining applies algorithms to text in order to extract patterns from the data. The main functions of data (and text) mining have been identified as:

- Summarisation – to provide a succinct description of key aspects of the data; this can include techniques of visualisation.
- Classification – to sort data into categories that are predefined by the researcher.
- Clustering – to sort data into categories that, in contrast to classification, are determined from the data themselves and not predefined by the researcher.
- Prediction – to use existing patterns in the data to predict future changes.
- Association – to explore relationships and associations between attributes or cases in the data.
- Deviance detection – to identify cases that deviate significantly from the dataset norms.
- Sentiment analysis – to monitor the degree of positive or negative sentiment in texts.
- Sequential analysis (an emerging field within data mining) – to conduct on-going analysis of data streams updating the analysis in real time as new cases are added to the database.

These approaches can be regarded as statistical rather than symbolic methods as the algorithms applied do not consider the meaning of the text. Instead, interpretation is later imputed upon the findings by the human analyst. In fact, the distinction between data mining and traditional statistics is a blurred one and there is considerable debate as to whether data mining constitutes a different discipline, or is merely a branch of statistics (Marshall & Townsley, 2006; Smyth, 2001). The difference most frequently touted is the focus of data mining on inductive exploration of patterns rather than deduction, inference and explanation. Estivill-Castro and Lee (2001) argue that the automated approaches offered by data mining allow the investigation of a greater number of variables and avoid restricting analysis to a confirmatory approach, thereby allowing unexpected patterns to be uncovered. However, most of these elements easily fit into the statistical paradigm as part of exploratory data analysis and hypothesis generation. Glymour et al. (1997) stress that data mining should not be divorced from statistics. They hold that statistical theory and methods can help to improve data mining, particularly by offering 'methodological morals' in the form of procedures for evaluation, measures of reliability and the adoption of a realistic attitude to the inherent uncertainty in all models. The adoption of Glymour et al.'s recommendations could help to fend off many of the criticisms frequently aimed at data mining which is often referred to negatively as 'data

dredging' (Selvin and Stuart, 1966 cited in Glymour et al., 1997). Fayyad, Piatetsky-Shapiro and Smyth (1996) highlight that these criticisms are justly applied to 'data mining' when it is applied blindly but not when it is embedded in a wider knowledge discovery process. Fayyad et al. stress that data mining should be regarded as part of a wider field of Knowledge Discovery in Databases (KDD). Application of KDD aims to ensure that, before analysis gets underway, analysts utilise prior knowledge of the field to define the purpose of the model (ensuring that it is asking the right questions) and that post analysis, the results are carefully interpreted and evaluated before being incorporated into the existing knowledge base.

Data mining has been discussed because it is a growing field in relation to the analysis of data. Although the analysis technique conducted for this thesis, cluster analysis (discussed in section 4.4 below) is adopted within data mining, it can also be classified as a traditional statistical technique which would allow the field of data mining to be ignored. However, a number of projects from the field of data mining have offered insights into the analysis of MO descriptions. Marshall and Townsley (2006) provide a useful summary of data mining techniques and their applicability to criminology. A number of studies have used data mining techniques to interrogate MO descriptions, and these are discussed further below. To date, the range of questions posed by these studies has been rather narrow, with an almost exclusive focus on the identification of offenders for the purposes of detection and notwithstanding the importance of detection, this somewhat narrow focus means the neglect of other questions, particularly relating to the prevention of crime, which these methods could address.

In addition, there are continuing developments in the data mining of 'big data', including the development of techniques applicable to large volumes of brief text fields beset with recording problems, not least the analysis of Twitter data. Therefore, developments in the field of data mining should be monitored to ensure any relevant innovations in the analysis of data can be applied to police data.

### **4.3.3 Narrative Analysis**

The above section has discussed how natural language, represented in textual data, can be standardised to make it amenable to statistical analysis. Quantitative narrative analysis (Franzosi, 2010) has also been proposed as a method for analysing descriptions of actions. The approach seemed to offer potential for the analysis of police recorded MO data, but on further investigation this proved not to be the case. The approach of quantitative narrative analysis is discussed here to demonstrate that it would not have been feasible with the large volume of cases to be analysed, but

also because the contrasts between quantitative narrative analysis and content analysis and natural language processing highlight some potential limitations of the methods that were adopted in this thesis.

Content analysis, and in particular its more automated forms, analyses instances of tokens in isolation from the surrounding language. However, this wider context can be enlightening in the analysis of text. In particular, 'bag-of-words' approaches to the analysis of text can sever the connection between actions and the actors performing them. The discussion of conceptual frameworks, Chapter 3, demonstrated the importance of understanding the performance of roles for the understanding of causes of crime, in particular the roles of crime promoter and crime preventer were highlighted. Further, the discussion of scripts established that the latter are not performed in isolation, multiple-scripts can produce favourable collaborations and script clashes. It would seem advantageous, therefore, to keep track of who is doing what to whom within the analysis of MO descriptions.

An alternative approach to the analysis of text, which seeks to keep together component elements of context, is quantitative narrative analysis. Labov (1972, p.359 in Franzosi, 2010) defined narrative as a "method of recapitulating past experience by matching a verbal sequence of clauses to the sequence of events which [it is inferred] actually occurred." MO descriptions represent a narrative account of the process of crime commission, although partial and imperfect in their nature.

Franzosi (2010) argued that narrative accounts should be analysed with the use of a 'semantic grammar' which codes the text in a hierarchy of subject (actor), action and object. The analysis should identify each element in the data – every actor and their linked actions and objects. Each of these elements can then be described with a modifier, or set of attributes. The benefits of this approach are that the contextual links which clarify who is doing what to whom remain intact. The downside of non-narrative approaches is that they produce a list of actions that are detached from the relevant actors. An example of an MO description coded according to the Franzosi semantic triplet is provided in Appendix 1. This coding is notably very detailed and complex. It requires even small fragments of text to be segmented into distinct clauses and requires some understanding of grammar to accomplish the task. Although Franzosi et al. recommend software to record and organise coding decisions (Franzosi, Doyle, McClelland, Putnam Rankin, & Vicari, 2013), this approach still requires manual coding decisions. There have been recent developments in natural language processing which include the automated identification of parts of speech (verbs denoting

actions, nouns denoting actors and so on), although these remain in their infancy<sup>20</sup>. The challenges of linking related parts of speech together i.e. linking actions to actors, has been identified as one of the most difficult problems facing natural language processing (Jurafsky & Martin, 2008). Therefore, although narrative approaches hold promise for the analysis of MO descriptions, at the time of writing, it appears to be unfeasible for large volumes of data.

#### *4.4 Statistical Analysis*

Raw unstandardized text data is unsuitable for statistical analysis. For example, frequency tables would contain an unmanageable number of different values, each one occurring only a small number of times. This table would disguise the fact that many of the different terms counted relate to essentially the same thing, therefore any patterns would be masked. The standardisation process reduces the number of different terms used in the database which allows for the production of frequency tables which are not only smaller and more manageable but also more representative of the meanings in the data.

Following the standardisation of text, frequency tables of the number and spread of different tokens across the cases were produced. This constituted the first step towards statistical analysis of patterns in the data. In text analysis fields this type of frequency table is known as a document term matrix. This is a table in which each case is represented by a row<sup>21</sup>, each token found in the standardised data is represented by a column and the cell values indicate the number of times each token occurs in each case. In the current research, the analysis was only concerned with the occurrence of words in each case and not the number of times a word was used in each description: therefore this matrix was recoded to indicate simply the presence or absence of a token in each case as illustrated in the hypothetical example in Figure 4.2. From the matrix alone, it is not possible to identify word order. In this sense, this method represents a 'bag of words' technique, as described above. The analysis is not influenced by the meaning of the words or the order in which they occur. This represents a potential limitation of this approach because, as noted above, narrative and temporal links between actors, actions and objects can be essential to determining the meaning of a segment of text. Nonetheless, as will be seen, this approach can deliver useful information from practical and research perspectives, while avoiding the extreme challenges of semantic processing.

---

<sup>20</sup> A test of one of these algorithms on a sample of the police data showed it to be very unreliable with a high degree of error.

<sup>21</sup> Remember that in content analysis and natural language processing each case is referred to as a 'document.'

	TOKEN A	TOKEN B	TOKEN C	TOKEN D	TOKEN E	TOKEN F	TOKEN G	TOKEN H	TOKEN I	TOKEN J	TOKEN K	TOKEN L	TOKEN M
Case 1	0	1	0	0	0	1	1	1	1	1	1	0	1
Case 2	0	0	0	0	0	1	1	1	1	1	1	0	1
Case 3	1	0	1	0	1	1	1	1	1	1	1	0	0
Case 4	0	1	0	0	1	1	1	1	1	1	1	0	1
Case 5	1	1	0	0	1	0	1	1	0	0	0	1	1
Case 6	1	1	0	0	1	1	0	1	0	0	0	1	1
Case 7	1	1	0	0	1	0	0	1	1	0	1	1	1
Case 8	1	1	0	0	1	0	1	1	1	0	0	1	1
Case 9	1	1	0	0	1	1	0	0	0	1	1	1	1
Case 10	1	1	0	0	1	1	1	1	0	0	0	1	1
Case 11	1	0	0	1	0	1	0	0	1	0	0	0	0
Case 12	1	1	0	1	0	1	0	0	1	0	1	1	1
Case 13	0	0	0	1	0	1	0	0	1	1	0	0	0
Case 14	1	0	0	1	0	1	0	0	1	0	1	0	0
Case 15	0	1	1	1	0	1	0	0	0	0	0	1	1
Case 16	0	1	1	1	0	1	0	0	1	0	0	1	0
Case 17	0	1	1	1	0	1	0	0	1	0	0	1	0
Case 18	0	1	1	1	1	1	1	1	0	0	0	1	0
Case 19	0	1	1	1	0	0	0	0	0	0	0	1	0
Case 20	0	1	1	1	0	1	1	0	0	0	0	1	0
Case 21	0	1	1	1	1	0	1	0	0	0	0	1	0

Figure 4-2 A hypothetical document term matrix (highlighted cells show the presence of a token in that case).

#### 4.4.1 Cluster analysis

Following text standardisation, coding and the creation of a document term matrix, the data is amenable to traditional forms of statistical analysis. In order to identify patterns of text occurrence within the data, cluster analysis was employed to simplify the data in such a way that the results closely reflect the original raw data, while providing useful insights into the patterns with the data. Cluster analysis applies the three classification principles defined by Burt (1940) as bringing similar things together, distinguishing between those items that are different and identifying kinship between categories.

When applied to textual data, cluster analysis identifies groups of cases which display similar patterns in token occurrences and groups those with similar patterns as members of the same cluster. Examples of clusters might include a set of crime events that share the same combination of offender types and locations e.g. a group of crimes occurring on street corners involving young people. It should be noted that tokens are not necessarily uniquely characteristic of one cluster. A distinct combination of tokens will provide the characteristic of a cluster, the same tokens in different configurations can also typify different clusters. Each case is assigned to a cluster and this cluster membership can be saved as a categorical variable to be used in further analysis. This allows for the exploration of the distribution of different clusters over time, location and across different victim and offender groups.

### *Methods of cluster analysis*

There are two main forms of clustering: hierarchical and iterative partitioning (Aldenerfer & Blashfield, 1984); the method employed in this analysis is a hybrid of these. Hierarchical methods produce sets of clusters that are organised into a tree structure in which 'child' clusters are nested beneath 'parent' clusters. This hierarchical structure can be graphically represented by a dendrogram (Figure 4.3). Hierarchical clustering can either be conducted bottom up (agglomerative) or top down (divisive). Agglomerative hierarchical clustering begins by placing each case in its own unique cluster so that the number of clusters is equal to the number of cases. The distance between each cluster is calculated and the closest two clusters are merged. This step is repeated by successively merging the two clusters that are most similar until the number of clusters equals 1. Divisive hierarchical clustering reverses this process, beginning with one cluster and conducting successive divisions until the number of clusters is equivalent to the number of cases. In both agglomerative and divisive clustering, the researcher must decide the optimum number of clusters to retain in their final solution through the inspection of output, including statistics and graphics (principally the dendrogram).

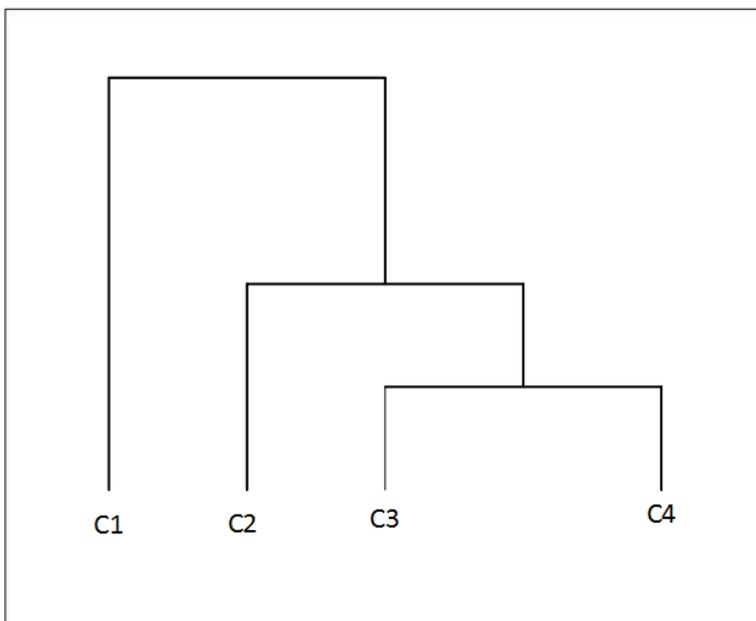


Figure 4-3 Example dendrogram showing four hierarchical clusters

In contrast to hierarchical methods, optimisation methods of clustering do not produce a hierarchical tree; instead they produce all the clusters at once. These methods, the most frequently used of which is k-means, aim to produce one optimal solution by rearranging the clusters so as to minimise the distances within each of the clusters while also maximising the distance between each

of the clusters. During the optimisation process, cases can move in and out of different groups until the optimal solution is identified. Although this provides more flexibility in identifying the optimal solution, it does mean that the result does not have a hierarchical structure. In k-means clustering, K represents the number of clusters to be obtained and it is necessary to specify K prior to analysis. In k-means analysis, it is not necessary to select the type of points from which distance will be measured (nearest neighbour and so on) because the algorithms explore all possible distances in order to select the optimum one.

K-means normally requires the specification of starting points or 'centroids', one for each cluster. These centroids are most commonly selected at random. It is from these centroids that the algorithm begins ordering and reordering the cases or documents into clusters. This has been highlighted as a potential problem with k-means, as the selected 'starting' centroids may influence the final solution, and different starting points could lead to different solutions (Everitt, Landau, Leese, & Stahl, 2011)

The current analysis adopted a method of cluster analysis known as Bisecting K-means (BKM). This method is a hybrid between hierarchical and optimising techniques in that it performs a repeated series of k-means analysis to produce a hierarchical structure. Thus BKM follows the following process (also depicted in Figure 4.4).

1. Begin with all cases assigned to one cluster
2. Initialisation: Select a case (usually randomly) to act as a centroid. Using the k-means (k=2) algorithm, all cases are compared to the centroid to determine the optimum place to partition the data to create two clusters.
  - a. The optimum split is determined by identifying partition(s) that ensure that all points within a cluster are closer to the centroid of that cluster than the centroid of the other cluster (in the case of BKM, only one partition, creating two clusters, is drawn each time)
3. Split the data into two clusters
4. Repeat steps 2-3 for each of the new clusters until the desired number of clusters is reached

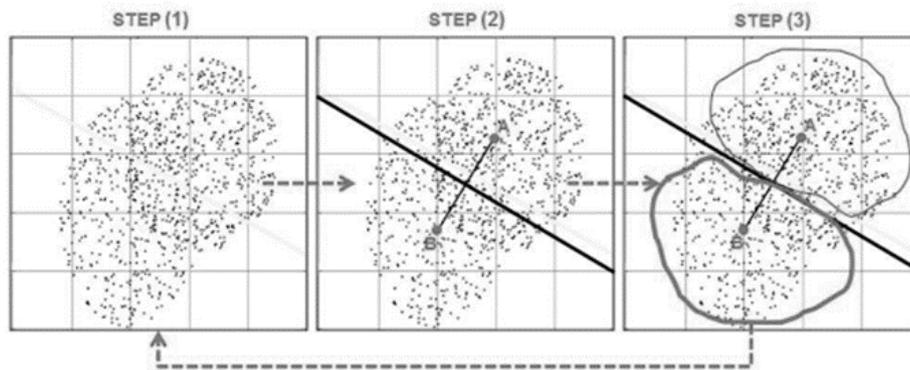


Figure 4-4 An illustration of how bisecting K-means operates – Reproduced from Lancia (2012b)

When applied to the analysis of text, BKM has been shown to work more efficiently and to produce more reliable results than other types of cluster analysis (Steinbach, Karypis, & Kumar, 2000). Steinbach et al. highlight that, in the classification of textual documents, there is a high potential for overlap in the word usage across clusters. It is common for all clusters to utilise a shared overlapping vocabulary for different classes of documents. Thus, it is possible for a document's nearest neighbour to use a lot of the same terms, but not to be from the same class of documents. Steinbach et al. argue that k-means and BKM are more suited to clustering in these circumstances and this is because these methods effectively compare the similarity of all documents in a cluster to all others and selects the optimal split. Further, Steinbach et al. argue that BKM performs better than standard k-means approaches due to the tendency for the latter to produce clusters of widely different sizes. Although small clusters tend to be of high quality, very large clusters tend to have a negative impact on a solution's overall validity. BKM further offers the benefit of a hierarchical structure which, as noted, is useful for the comparison of different solutions with different values of  $k$ .

When using BKM there are different criteria available to decide where to split clusters. Options include identifying the largest cluster or identifying the cluster with the greatest within group dispersion or selection may be based on a combination of the size and dispersion. However an experiment by Steinbach, Karypis and Kumar (2000) found that the use of different criteria to split clusters had little discernible difference in the overall cluster results. The implementation of BKM in the analysis for this thesis used the sum of squares due to error (SSE) as a decision point for dividing clusters. This identifies any part of the cluster with a large SSE and splits this into two new clusters.

As with traditional k-means, BKM requires the a priori specification of the number of clusters. However, it is possible to produce a set of cluster solutions covering a range of different values for  $k$ . This can be time consuming, but given the exploratory role of cluster analysis (emphasised by Byrne & Uprichard, 2009) it is valuable to explore alternative cluster solutions. The hierarchical nature of

BKM helps with the inspection of different cluster solutions as it is possible to trace the parents and children of individual clusters and explore the implications of different cut off points.

### *Interpreting clusters*

BKM cluster analysis requires that the analyst specify the desired number of clusters prior to running the procedure. Garson (2014) advises that solutions aiming to identify typical cases should aim to produce a small number of clusters, whereas solutions aiming to identify exceptional cases call for many clusters. This is helpful, although 'many clusters' is a relative term. In practice, the procedure can be repeated to compare the effect of requesting different numbers of clusters. Solutions can be compared by assessing the ratio of between-cluster dispersion to total cluster dispersion. BKMs, and k-means more generally, aim to minimise *within cluster dispersion* (i.e. maximise homogeneity within each individual cluster) while maximising *between-cluster dispersion* (i.e. maximise heterogeneity between clusters) ensuring that there are genuine differences between the identified clusters. An index score can be calculated which compares the ratio of between-cluster dispersion to total cluster dispersion, where an increase in the value of this index indicates an improved solution. This is calculated as:

$$Index = \frac{Between - Cluster Variance}{Between - Cluster Variance + Within - Cluster Variance}$$

Metrics can also be produced for individual cases to assess the degree of fit between a case and its cluster. A cluster membership score can be assigned to each cluster indicating the degree of certainty that a case belongs to a given cluster. A score of 1 indicates a 100% certainty that the case was accurately assigned to a given cluster, whereas a score of 0 represents an absolute certainty that this case does not belong in a given cluster. This score is proportional across all of the clusters: therefore, the sum of scores across all clusters for each case will always equal 100. It is, therefore, possible for cases that demonstrate characteristics appropriate to more than one cluster to have, for example, a score of 50% for each.

Cases with very low scores for their allocated clusters may not represent a genuine fit, but were no better aligned to any other cluster. It is important to consider how such outliers should be treated. In cluster analysis, outliers can be treated by disregarding outliers prior to analysis, forcing them to be assigned to a cluster (this is the outcome in a BKM solution), or creating an 'outlier cluster'. In this analysis, there are a number of reasons why a case could be an outlier – there may be something different in the way it is recorded, or they could represent genuinely unusual events that have been

committed in a different way from other cases in the data. These cases could represent the emergence of new offender methods, important intelligence in the bid to prevent crime. For this reason, it is important not to disregard unusual cases as these may carry useful information. There is a risk that interesting/important outlier cases are merged into other clusters and lost.

As Everitt and Dunn (2001) highlight, clusters are groups of cases based on rules: they cannot be falsified as with theories, only assessed for their usefulness. It is therefore, important to interpret the clusters. This requires the identification of their defining features, in the case of this thesis determining what class of events each cluster represents. A useful step to this end is the identification of tokens that are characteristic of each cluster. Using the chi-square statistic, it is possible to determine which tokens are represented in a given cluster at a higher frequency than would be expected if tokens were distributed evenly across all of the clusters. These tokens will be referred to as 'overused' tokens within each cluster. The dataset analysed was large and the chi-square scores generated had a tendency to be high, therefore, a high critical value threshold was imposed to determine statistical significance. Thus, tokens with a score of 10.83 or higher were deemed to be overused to a statistically significant level ( $p < .001$   $df=1$ ). A series of diagrams were produced in order to visualise the characteristics of each cluster.

Wordclouds were produced to highlight the tokens that were most characteristic of the each cluster (see Figure 4.5 for an example). In the wordcloud, the size of each token is representative of its relevance to that cluster. Size was weighted by the relevance chi-square score<sup>22</sup> (it is important to stress that words are weighted not simply the *number* of cluster cases, which included that token). Tokens represented by the largest font size, therefore, have the highest probability of occurring in the given cluster compared to a distribution where all tokens were equally distributed across the clusters. Tokens that are characteristic of a cluster, may be relatively infrequent across the database as a whole, but if those tokens largely occur within the same cluster, they have a high relevance in explaining that cluster. Tokens can also be relevant to more than one cluster, it is the combination of relevant tokens that determines each cluster. In Figure 4-5, the dominant tokens are *entertainment*, *pocket* and *clothing*. These tokens occurred in this cluster at a frequency higher than would be expected if all tokens were spread evenly across the data. The bottom panel of Figure 4.5 shows the

---

<sup>22</sup> The Chi Square score compares the frequency of tokens occurring in a given cluster compared to the expected frequency that would result from random distribution of tokens across all of the clusters. In the charts larger-font words have a larger chi square score. Only statistically significant ( $p < .001$   $DF=1$ ) chi squares are included in each figure.



Table 4-2 Methods for reporting and evaluating clusters

<b>Clatworthy et al. recommendations</b>	<b>This study</b>
The computer program used in the analysis	R – “Cluster” Package.
Similarity measure	Squared Euclidian Distance.
Cluster method	Bisecting K-Means.
How number of clusters was determined	Initial cut was determined by the ratio of inter and intra cluster variance. To explore substantive issues further, an extended solution containing 30 clusters was also inspected.
Evidence for the validity of clusters – for example, against research literature expectations and against other variables	The clusters are considered in the light of research evidence and conceptual frameworks. The clusters are also cross-tabulated with other relevant variables, including crime type and by comparing different sections of the sample i.e. by selected quarter.

#### *4.5 Summary*

The Methodology chapter has described the analytical approaches that have been applied in this research and has provided a rationale for their selection. A key factor in the selection of methods was their appropriateness to address the research questions with volume crime data. This has largely influenced the selection of quantitative methods over qualitative, although as highlighted above, this choice is not without limitation. The data to which these methods were applied is described more fully in Chapter 6, followed by a presentation of the results of the analysis in Chapters 7 to 9, including an exploration of the relationship between statistical results and themes drawn from conceptual frameworks. These chapters will also reflect further on lessons learnt as the methodology was applied and the suitability of the methodology for practice. Prior to these ‘data driven’ chapters, Chapter 5, now following, will further explore the applicability of CCO and script frameworks to the crime types of interest i.e. theft from the person and robbery of personal property.

# Chapter 5 Integrating and Applying Frameworks: Research Literature

## *5.1 Introduction*

Chapter 2 introduced the CCO and script frameworks that have been developed to understand environmental influences on offender decisions and actions. This chapter will use these frameworks as a structure for organising findings from previously conducted research relating to theft from the person and robbery of personal property. Applying these frameworks serves two purposes. The exercise will help to critique the utility of these frameworks as tools for organising the available knowledge about offence commission and the factors that shape offender MOs. The exercise will also provide a body of knowledge that can be used to contextualise and fill any gaps in the findings from the analysis of crime data in Chapters 7 to 9.

The use of frameworks to review and synthesise existing literature imposes more structure than normally found in a traditional literature review. However, the review should not be regarded as a full systematic review (Grant & Booth, 2009). In contrast to a systematic review, there were no predefined search or inclusion criteria, consequently the search strategy is not reproducible. There were also no formal attempts to assess the quality of studies prior to inclusion although all included studies are from academic sources and the majority are peer reviewed. The strength of the approach used was the use of theoretical frameworks to formally categorise the research findings. This process was facilitated by the use of the qualitative data analysis software NVIVO (di Gregorio, 2000) which enabled passages in research articles to be coded with relevant nodes derived from the frameworks. Figure 5.1, an NVIVO screenshot taken when the literature review was progressing, illustrates the codes that were created to organise the literature. When applying the CCO framework the codes were taken directly from Ekblom's CCO nodes, with a few additional codes added to record reflections and critique. For script frameworks the process of coding was more inductive. Codes were developed to reflect the tasks that the literature revealed were necessary or commonly performed during the commission of offences. These codes were then organised into a hierarchy of acts and scenes.

Tree Nodes		Sources
Name		
+ 00 Notes		23
+ 01 Authors		80
- Script - Strategy Scenes		64
+ 01 Readiness		13
+ 02 Preparation		23
- 03 Selection		33
+ Target search		8
+ Target Selection		16
- 04 Transfer		15
Approach		5
+ Engagement		6
Manage resistance		0
+ Transfer		4
+ 05 Post		18
+ Script definition		35
+ Script function		11
- Script Structuring		88
01 Definitions		20
- CCD		82
+ 00 Caused agent		21
+ 00 Dynamics		22
+ CCD - 01 Crime promoters		7
+ CCD - 02 Absence of crime preventers		7
CCD - 03 Wider environment		4
+ CCD - 04 Target enclosure		5
+ CCD - 05 Target		13
CCD - 06 Offender Presence		9
+ CCD - 07 Anticipation of risk, effort and reward		32
+ CCD - 08 Resources for Crime		54
+ CCD - 09 Readiness to offend		13
CCD - 10 Lack of skills to avoid		3
CCD - 11 Criminality (predisposition)		2
Walter - Situational precipitating inducements		10

Figure 5-1 Literature review in progress: a snapshot of the coding structure in NVIVO

The starting point for coding the literature into crime script categories was the legal definitions of theft from the person and robbery of person property. This is an unusual approach as legal definitions have been deliberately sidestepped by situational crime theorists. Cornish and Clarke (2008) argue that even the finer legal distinctions between crimes may be too broad to use as a basis for understanding criminal decision making. However, legal definitions may at least provide a useful starting point in providing an outline of the actions that must be accomplished to determine

whether or not an offence has taken place (actus reus). Section 1 of the Theft Act 1968 for England and Wales defines theft as 'the dishonest appropriation of property belonging to another with the intention to permanently deprive the other of it'. Theft, therefore, requires an unauthorised physical transfer of property and/or an abstract transfer of possession. 'Theft from the person' occurs when, at the time of the theft, the property in question is in the physical possession of the victim i.e. the property is held by, worn, contained in an item of clothing or physically attached to the person, for example, by placing a bag over a shoulder (Home Office, 2015a). Robbery of personal property occurs when the transfer of possession is achieved by force. Under Section 8 of the Theft Act 1968, "a person is guilty of robbery if he steals, and immediately before or at the time of doing so, and in order to do so, he uses force on any person or puts or seeks to put any person in fear of being then and there subjected to force." Current counting rules for robbery (Home Office, 2013) state that classification as robbery requires sufficient evidence that force was *used or threatened* or that the victim was under the impression, based on the offender's language or actions, that force was *intended*. Where force is directed towards the property but not to the victim 'e.g. a bag is taken cleanly from the shoulder of a victim or a phone is taken cleanly from the hand' (Home Office, 2013) the counting rules stipulate that the offence should be classified as theft from the person and not robbery of personal property.

Theft of property can be regarded as a meta-script, theft from the person as a protoscript, and theft from the person through robbery and theft from the person without force form two distinct scripts, each of which can be performed via a range of tracks. Official definitions can therefore be helpful in determining and describing the essential differences between scripts and outlining the essential actions (actus reus) that must occur for a given crime script to be completed successfully (even if this is only at the higher levels of script hierarchies). In the case of theft from the person, there is clear overlap between the legal definitions and basic script definitions.

However, legal definitions only include the fundamental actions that must be performed for a criminal offence to occur, these definitions are not concerned with the manner in which actions are performed, the sub-tasks that may form precursors to the central actions<sup>24</sup> or the potential variety of action/decisions routes by which an action is eventually performed. Existing research findings can provide an understanding of the more detailed script layers including the component script acts and scenes. The sections that follow, dissect thefts from the person into a series of script scenes and describe each one in turn, outlining the key tasks that must be accomplished in each scene and

---

<sup>24</sup> Perhaps with the exception of actions that relate to conspiracy to offend or going equipped.

considering the evidence relating to factors that influence offenders' conduct during each scene. The literature reviewed below includes a small number of studies that have utilised script frameworks to some degree (Beauregard et al., 2007; Copes et al., 2012; Nee & Meenaghan, 2006; Topalli, Jacques, & Wright, 2015; Wright & Decker, 1994). For other studies, script frameworks have been applied retrospectively. None of the studies had explicitly applied the CCO but this is also used to organise the material retrospectively. Following the discussion of each scene, Table 5.1 will provide a summary of an integrated script and CCO framework that has been populated, based on findings from existing studies.

In comparison with other types of crime, theft from the person offences comprise very few acts, with most of the decision making and action being compressed into a short time and taking place predominantly in one location. Each scene of the scripts will be articulated in order to identify 1) the key components of each scene, 2) the tasks that are imperative for the offender and 3) the situational causes that may shape offender decisions and actions. It should be acknowledged that articulating scripts by way of individual scenes, while intended to be analytically useful, may make the scenes appear more concrete and discrete from each other than they are in reality. In fact the literature reveals that there is a great deal of overlap and blurring between scenes, the ordering of scenes may vary and some scenes can occur simultaneously.

The discussion of scenes will include the identification of prerequisites for the initiation of each scene and those for the transition to the next scene. These conditions are considered through the lens of the CCO framework. The CCO presents a static list of proximal causes and while some preconditions are 'provided by the environment' the discussion below demonstrates that offenders can also dynamically manipulate situational conditions to optimise the creation of the necessary conditions to achieve their goals (Ekblom, 2012b). Offenders do not manipulate these conditions in isolation and a key factor to explore is the influence of script clashes and the opposing goals of other actors. The literature identifies global preconditions that operate across the whole script and initiate and then shape subsequent offending, this is in line with Leddo and Abelson's (1986) 'universal script' which identifies 'preconditions' as a distinct scene. However, Tompson and Chainey (2011) modified the universal script, calling for the identification of preconditions for *each scene*. This helps to highlight the changing influence of situational conditions as the script progresses and to demonstrate the requirements for initiating a scene, completing it and then transitioning to the next phase.

## *5.2 Exploring and Explaining Theft from the Person Scripts*

### 5.2.1 Precondition: Readiness

A clear precondition for entering into a theft from the person script is readiness to offend. This section will investigate what can be understood from the literature about the formation of a commitment to enter into a theft from the person script. The nature of this commitment and the circumstance in which it is formed can influence the script and the tracks taken. However, it is important to note that, in reality, the distinction between commitment to a script and acting it out can be very blurred, particularly in the case of the offences being considered here.

Important insights regarding proximal and distal sources of criminal motivation for theft from the person have been uncovered via a body of qualitative and ethnographic studies (Deakin, Smithson, Spencer, & Medina-Ariza, 2007; Jacobs, Topalli, & Wright, 2000; Jacobs et al., 2003; Jacobs & Wright, 2008; Jacobs, 2012; Topalli & Wright, 2004; Wright, Brookman, & Bennett, 2006). However, studies utilising script frameworks rarely explicitly examine the factors which cause an offender to enter into a crime script. The lack of integration between studies that explore how crimes take place and those offering broader causal explanations is unfortunate because the selection of a specific script track is shaped by the nature of readiness and predisposition. The CCO framework prompts the researcher to consider offender predisposition and the more immediate prompts of readiness (including those located within the situation) along with the availability of resources and factors restricting the actor's access to alternative forms of action, including their executive function (Wikström, 2009). Rational choice perspective and situational action theory, respectively, explain that offenders will only act on opportunities which match a previously made readiness to offend decision, or have been formulated as acceptable action alternatives in line with their moral code. Situational action theory provides the further assertion that alternatives afforded by the situation are unlikely even to be *recognised* by individuals who either do not have relevant goals, have moral rules against the course of action or have sufficient executive function to avoid offending.

In parallel with the CCO's distinction between predisposition and readiness to offend, Jacobs (2010) stresses the distinction between 'static factors' that predispose an individual to offend and those factors which shape crime in the moment. A prevalent theme in robbery research has been the influence of street culture in providing both distal and proximal motivations for theft. Assimilation into subcultures takes place at a distance spatially and temporally from the specific crime event but directly shapes the goals, decisions, life options and behaviour of offenders within the crime

situation (Shover & Honaker, 1992)<sup>25</sup>. Particularly for younger offenders, a predisposition to robbery develops through the internalisation of violence, risk taking and dominance of others, as not just acceptable action alternatives, but valued goals in their own right (Deakin et al., 2007). Thus, from the perspective of situational action theory, robbery offenders, not only have weak personal moral rules *against* violence, but also, strong personal rules and emotions supporting violent actions:

*“Acts of violence ...occur when people with weak personal moral rules and emotions opposing the intentional harming of others ..... or strong personal moral rules and emotions supporting the intentional harming of others .... encounter opportunities or frictions which may tempt or provoke them to act violently in settings which have a violence-conducive moral context”*  
(Wikström & Treiber, 2009, p.82).

Research has identified a number of short term motivations causing readiness to commit theft from the person including, a specific need for cash, the need to fund the continuation of ‘life as a party’ (Shover & Honaker, 1992), to access drugs (either directly or cash to purchase drugs) the disinhibiting influence of drugs or alcohol and specific emotional states such as anger and jealousy (Jacobs & Wright, 2008; Topalli & Wright, 2004). Ekblom discusses the role of revenge and righteousness as a motivation of crime, including violence (Ekblom, 2007). Jacobs and Wright (2008) explore the influence of revenge as a motivation with specific reference to street robbery. In such cases, robberies are an act of retaliation committed with the practical and symbolic purpose of maintaining honour and respect. This desire to maintain image and status with others is a longer term predisposition; actions of others that risk tarnishing this reputation create short term readiness to offend. To meet these ends, revenge robberies must be swift, visible and often excessive. The motivations and emotions behind ‘readiness to offend’ are essential, not only, in understanding why offences take place, but also, in explaining how. For offenders, robbery as revenge is morally justified. The nature of street culture, with tensions generated by a combination of exhibitionism and jealousy, means that prompts, provocations and pressures (Wortley, 1998) for readiness are ever present. The influence of offenders’ impression management, identified through qualitative

---

<sup>25</sup> There are suggestions that initiation into street or gang culture may in itself require acts of theft and/or violence (Asthana, 2008)

studies, appears to be neglected in work that adheres strictly to a narrow conceptualisation of rational choice.

Not only do criminal networks provide the values supporting involvement in crime, they can provide tutelage and resources to support engagement in crime scripts while simultaneously restricting access to alternatives to criminal behaviour. Wright and Decker (1994) cite Lofland's (1969) concept of 'psychological encapsulation' to explain a phase during which the individual will feel socially isolated and unable or unwilling to seek help from others including help that could provide an alternative course of action than crime. It is apparent that criminal subcultures do more than provide and shape motivation, they also provide crime promoting relationships and erode fiscal and social capital, and restrict access to alternative courses of action to the extent that those options may appear invisible to offenders. It is not easy to maintain a job while participating in an offending lifestyle (Shover & Honaker, 1992). Thus, in CCO terms, the culture of the street can exacerbate an offender's 'lack of skills' to find an alternative opportunity to robbery and theft.

As already suggested, the nature of predisposition and readiness has a direct impact on the nature of robbery committed, for example the nature of motivations were shown by Deakin et al. (2007) to govern the timing of offending. For example, where the decision to offend was driven by the need to fund a drug fix, robberies tended to be committed shortly after waking. Methods of offending that require patience and persistence will not be available to offenders who lack these executive functions, particularly when faced with desperation brought on by drug dependencies (Copes & Cherbonneau, 2006). Wright and Decker (1994, 1997) describe the emotional pressure under which many offences are committed, those offenders who have the ability to control or limit this pressure are freed to be more selective (and 'rational') in their selection of crime opportunities.

As will be reiterated throughout this chapter, the presence of distinct types of predisposition and readiness does more than provide an on/off impetus for offending: it also shapes the choices of scripts and script tracks executed by offenders which in turn, changes the way in which CCO elements form and conjoin.

### **5.2.2 Act: Preparation**

A pervasive theme within the theft from the person literature is the absence of any formal planning/preparation phase, although the literature does provide some exceptions to this rule. This act is, therefore, far from essential to the completion of offences of this nature. Hochstetler (2001) found that most robbers and burglars either did not plan their crimes or spent only a few minutes planning. Some authors have suggested that experienced offenders and those whose need for

rewards are less urgent are more likely to plan (Copes & Cherbonneau, 2006). However, other research refutes this. For example, Feeney (1972) cited in Jacobs (2012) argued that experience tended to take the place of planning – an assertion that is important for considering the role of scripts as a mechanism for retaining and readily accessing knowledge gained through experience (Seifert et al., 1994).

Research investigating the methods used by burglars has shown that their preparations can involve creating a bank of suitable targets for future offences (Nee & Meenaghan, 2006). The mobility of victims targeted in thefts from the person restricts the potential to store targets for future reference and therefore this option is rarely available for theft from the person offences. Overall, research has shown that street robbers do little planning or preparation; exceptions include cases where a specific victim is targeted, which was more common when the victim was another criminal or when the attack was motivated by revenge (Jacobs, 2010). The higher risks involved in targeting other street criminals warranted greater prior consideration. Copes, Hochstetler and Cherbonneau (2012) found that carjackers going for a riskier target such as a drug dealer, would embark on this with greater consideration of the timing, location and methods of attack, as if this type of target prompted a cue to ‘think tactics through carefully.’ It should be reiterated however, as discussed in Chapter 2, that the lack of planning should not be taken as an absence of purpose or rationality.

It would be anticipated that any preparation would involve the gathering together of required resources and although this is clearly plausible, theft from the person studies do not report evidence of a specific ‘get resources’ scene. Research has shown that the availability of resources and nature of motivation interact to influence the selection of MO (Cherbonneau & Copes, 2005; Kriven & Ziersch, 2007). However, robbery and theft from the person are not overly technical crimes, and few script tracks are resource specific. When considering scripts it is important to consider not just the presence or absence of resources, but the offenders’ perception of whether they are adequately resourced which may be a more important influence over whether and how they will offend.

Weapons can be important resources for robbery scripts. Smith (2003) reported that a third of robberies in the UK involved weapons. Both Deakin et al., (2007) and Barker, Geraghty, Webb, Key, & Britain (1993) found that offenders preferred to use weapons to communicate a threat rather than to inflict an injury. The intended target can influence the decision to carry a weapon and given the higher risks when targeting other criminals, offenders are more likely to carry a weapon in these cases (Deakin et al. 2007). In addition to physical resources, personal resources for robbery include the nerve to perform actions under the pressure of possible apprehension, and the ability to stomach violence (Gill, 2000; Wright & Decker, 1997).

Collaborators may also be important in theft from the person. Deakin et al. found that street robbers frequently work in pairs or groups while snatch thieves were more likely to operate alone. The authors also noted that co-offending was more prevalent with young offenders. Divisions of labour employed in robbery offences include employing an accomplice to act as look out while the offence is carried out (Hallsworth, 2005; Wright & Decker, 1997). Robbery offenders interviewed by Deakin et al. reported that the benefits of co-offending amongst street robbers included pooling knowledge such as the locations of CCTV and the availability of suitable targets. Copes et al. (2012) found that offences were more likely to be collaborative if they took place in riskier locations such as drug dealing areas. Again, the research demonstrates interactions between the scripts available, the characteristics and intentions of offenders and the situations in which offences take place.

Thus preparation concludes when the offender perceives him/herself to be appropriately resourced. Although it is plausible that preparation may involve conscious steps to plan and get resources, such as selecting a location and time to offend, meeting with co-offenders or accessing weapons with the intention of committing thefts, research studies suggest that these offences are more likely to unfold organically. Offenders may be 'prepared' by virtue of already being with potential co-offenders at a time when other conditions conjoin. Nevertheless, being readied and resourced are important preconditions for theft from the person scripts and these conditions influence, not only whether a script is initiated, but also, which track will be taken. Neither do these scripts begin in a consistently neat "if this, then that" linear sequence. The formation of commitment to enter a script may be prompted predominantly by the presence and recognition of appropriate opportunities, affordances in the environment or the presence of co-offenders or the absence of crime preventers and as shall be discussed below, readiness to offend and recognition of opportunity, may be and frequently are simultaneous.

### 5.2.3 Act: Target selection

That there is a target, and this target is a person, are essential preconditions given the legal definitions of theft from the person and robbery of personal property. In thefts, the victim, or the target person, does not represent the offender's ultimate target, they are the host which carries the final target. The target person needs to be located, recognised and selected by the offender. The selection of a target takes place in a wider environment that must be conducive to the offender's available scripts for conducting an offence. The wider environment is of central importance here because features of target attractiveness are situationally defined. Suitable targets, motivated offenders and conducive environments do not always drift together coincidentally. Target selection

may, but does not always, follow an active search but even where target discovery is fortuitous, selection requires a recognition of the target's suitability.

The act of target selection is described below in two scenes, the first covers the search for a target which may include the selection of an appropriate location and the second covers the recognition and selection of an attractive and vulnerable target (person). The scenes describe the steps through which offenders become located in and take advantage of a conducive wider environment which contains a supply of suitable targets.

### *Scene: Search for target*

The offender must select a target that is located within a suitable wider environment; indeed, the offender can actively influence the role of wider environment by making an explicit decision to travel to a conducive location. Other actors performing roles as crime promoters or crime preventers can also influence the degree to which an offender perceives an environment to be suitable. However, an active search for a target location is not essential and offenders can also take advantage of environments which other 'routine activity' scripts bring them into.

As noted in Section 2.2, features of the wider environment that have been identified as generally conducive to crime include: low levels of surveillance, the ability to access and importantly to escape, and the location of the site within the offenders' awareness space (Brantingham & Brantingham, 1975). These factors have been shown to be relevant to robbery. Research demonstrates that robbers prefer to offend close to home (Deakin et al., 2007; Feeney, 1986; Wright et al., 2006; Wright & Decker, 1997) and this is the case – even where there is access to transport, which would provide a means of extending the search ground. It should be noted that Deakin et al. argued that robbers could quickly become familiar with previously unknown areas particularly through sharing knowledge with collaborators; a further example of how offenders can shape opportunity as well as be shaped by it.

Attributes of offenders and their lifestyles can place limits on the search areas available to them. Cherbonneau and Copes' (2006), Copes and Vieraitis (2009) and Wright and Decker (1994) describe how offenders aimed to blend in to the environment, during both the target selection and exit stages of an offence. Some offenders felt their skin colour restricted the areas in which they could search for targets without being noticed, other offenders actively sought to overcome this by dressing to blend in (Feeney, 1972; Wright & Decker 1997). Note that in most cases personal qualities did not determine whether or not an offence was committed, rather they shaped the

nature of offending that took place. Although physical appearance is relatively fixed, additional resources can be drawn upon to adapt to the challenge of specific offence scripts.

Robbers will operate in areas where they anticipate there will be a ready supply of attractive targets; therefore target search is linked to the routines (scripts even) of potential victims. This does not necessarily mean affluent neighbourhoods, or neighbourhoods where affluent people spend their time: Wright and Decker (1997) found that some offenders preferred poorer neighbourhoods where the population depend more on cash than credit cards. In a similar vein, those locations where cash flows can also attract robbers – including cash machines, cheque cashing outlets (Holt & Spence 2005 in Bernasco & Block, 2009) and bars and licensed premises (Roncek & Maier, 1991 in Bernasco & Block, 2009).

The CCO also suggests that absence of crime preventers may influence the location of robbery and theft from the person. This is confirmed by the literature reviewed although, in line with routine activity theory, the term guardian is frequently used rather than crime preventer (as noted in Section 2.2.3 preventer has a broader meaning than guardian). Both Cohen and Felson, with the original conception of capable guardian and Ekblom, in the conception of crime preventers, distinguish between formal and informal guardians/preventers. Harper, Khey and Nolan (2013) found evidence that the presence of formal guardians, in the form of concentrated police patrols, could deter offenders from committing robbery. However, they also found that concentrated patrols could displace robbery to other areas. Winge and Knutsson (2003 in Hollis-Peel, Reynald, Bravel, Ellfers & Welsh, 2011) found that increasing guardianship through actively monitored CCTV can be effective in reducing robbery.<sup>26</sup> In relation to informal crime preventers research has shown that sparsely populated environments, for example public transport stations at off peak times, offer little supervision and provide few effective crime preventers, (Smith & Clarke, 2000). However De Haan and Vos (2003), found that most robbery offenders did not consider informal preventers/guardians (i.e. normal bystanders) to be much of a threat. Hollis-Peel et al. (2011) argued that guardianship operates through the availability of people watching who are able to detect problematic behaviours. Thus, availability of lots of people may not be sufficient to provide guardianship if these people are unable or unwilling to act. Research has shown that crowded environments can be conducive to certain types of crime, particularly theft from the person. In a study of theft of purses from shopping bags in Birmingham, UK, Poyner and Webb (1997) found that the crime was concentrated in the

---

<sup>26</sup> Ekblom (2011) views crime preventers as active human roles, however Hollis Peel et al (2011) conceptualised CCTV as human guardianship when it is actively monitored.

most densely populated market areas, and less so in markets with a more spacious layout. Tabangin, Flores and Emperador (2008) also point to the effects of the physical environment in creating crowded environments. By mapping crime incidents against features of the physical environment they demonstrated that crowding is a major factor contributing to theft from the person. Places where people are confined in constrained environments, e.g. shopping centres, busy pedestrian walkways and places where different groups converge together e.g. busy pedestrian junctions, pinch points around escalators or entrances were found to experience higher levels of theft from the person. In crowded places crime preventers are unable to distinguish offenders from other individuals in the space and actors may be too distracted to detect any suspicious activity.

Factors within the environment can influence both the availability of guardians/preventers and their ability to detect and respond to suspicious behaviours. Land uses and urban topography can influence the supply of crime preventers (Johnson & Bowers, 2013). Research has also indicated that the weather can influence the supply of potential crime preventers (Tompson & Bowers, 2015). Factors such as the level of illumination can influence the ability of crime preventers to detect suspicious activities (Coupe & Blake, 2006). However, in relation to artificial lighting the research evidence is mixed and most recently Steinbach et al. (2015) demonstrated that there was no association between reducing levels of lighting and increases in crime, including robbery.

A quantitative analysis of police records of robbery in Chicago by Bernasco and Block (2009 using data collected by Sampson, Raudenbush & Earls, 1997) found that structural community factors can influence offenders' decisions with the effect of either fostering or preventing crime. They found that 'poverty, cultural differences between community residents, and high residential turnover rates make it difficult for community residents to build up a socially well-organised community,' (p.102). Resultant low levels of informal control could translate into an absence of capable crime preventers.

A series of studies have sought to produce statistical models of the factors that influence offenders' choice of target location. In contrast studies which simply examine the characteristics of offence locations, studies utilising conditional logit models explore the characteristics of both offence locations and all of the alternative locations in which an offender could have offended but did not. Conditional logit models examine decisions made as a discrete choice against a set of potential alternatives. Models adopting this approach seek to understand the factors which caused an offender to select one location over the range of possible locations that were not selected. This approach has been used to study the target locations of street robbers (Bernasco & Block, 2009; Bernasco, 2010; Bernasco, Block & Ruiters, 2013) however a high degree of consistency can be seen from the findings of studies examining other crime types particularly burglary (Bernasco &

Nieuwbeerta, 2005; Clare, Fernández & Morgan, 2009) and theft from vehicles (Bernasco, 2010; Johnson & Summers, 2015).

Findings from these studies have demonstrated that offenders select target locations that are close to their homes (Johnson & Summers, 2015) close to prior homes (Bernasco, 2010), in areas with low social cohesion – indicative of low guardianship (Bernasco & Block, 2009; Bernasco & Nieuwbeerta, 2005; Johnson & Summers, 2015). Bernasco and Block (2009) found that offenders, in this case street robbers, were more likely to select locations that were close to either legitimate small businesses or illegal trading, suggesting the availability of potential targets generated by cash-intensive activities. Illegal markets were also important as a source of drugs, another attractive target for robbers. Bernasco and Block (2009) found that street robbers were more likely to offend in densely populated neighbourhoods that were a similar racial/ethnic mix to the offender's home area or in areas where the ethnicity was very heterogeneous.

Johnson and Summers (2015) examined the target location of theft from vehicle offenders and found that in addition to proximity to the offender's home, locations selected were significantly more likely to be located near to activity nodes (e.g. schools for juvenile offenders and public transport hubs for adult offenders) and in areas that were well connected by major road networks. Clare et al. (2009) found that offenders were more likely to favour areas connected to their own areas by a train line. Johnson and Summers also found that although the availability of targets was significantly correlated with location choice, the effect smaller than anticipated, in other words offenders do not necessarily select the area with the highest availability of targets – rather they appear to select an area where there is at least a suitable number of targets available.

Studies utilising conditional logit models demonstrate a clear relationship between the routine activities of offenders, the awareness space generated by these activities and their selection of target locations. However, it should be remembered that offenders' selection of location may not necessarily follow an explicit decision to make a journey for the purpose of offending. There is much research evidence to support the fact that offenders also act upon crime opportunities that they identify during the course of routine activities when they may not have been explicitly planning to offend (Jacobs, 2010).

The above discussion considered factors which shape offenders' choice of offence location. The research literature also provides insight into how identification of targets takes place. Although identification of a target is essential, active searches for targets are not always conducted. Ekblom and Sidebottom (2008), Felson (2009) and Bernasco (2009) draw upon understandings of foraging to

explain the behavioural tactics used by offenders to exploit the benefits in the environment while managing the risks of harm and wasted effort. Bernasco explores the extent that Optimal Foraging Theory (OFT), used to explain the foraging patterns of animals in the wild, elucidates the ecological restrictions influencing the foraging strategies of offenders, with key restrictions being the availability and distribution of targets and the presence of competitors and natural enemies. The theory demonstrates how the need for predators to be vigilant of other risks reduces the optimal efficiency of foraging. Thus OFT could be used to consider how offenders identify targets while competing with other offenders and handle the challenge of other individuals with conflicting goals (crime preventers, including the victim and bystanders who want to prevent the crime from happening). OFT raises important questions including what additional level of rewards an offender would expect in return for travelling further, to how long an offender is prepared to wait in a hunting patch in order to reap rewards. However, Bernasco recognises a number of limitations to OFT as applied to offending behaviour, most notably the fact that in the wild finding food is an essential behaviour for which there is no alternative, whereas criminals do have other alternatives (although they may perceived these choices as restricted).

Studies into other types of offending provide some useful insights which can be transferred to a consideration of target search scenes for robbery/theft from the person. One important finding is the distinction between the hunting ground and the attack site. Separation of these two situations is a common sexual offence track but is rare for personal robbery and theft from the person (Rossmo 1997 in Beauregard, Proulx, Rossmo, Leclerc, & Allaire, 2007). Another common thread is the distinction between offenders who actively search for their target victims and those who encounter their victims during the course of other routine activities. Copes and Cherbonneau (2006) explained how search strategies are shaped by the motivations of the offender. They distinguished between 'alert opportunist,' offenders, who are not specifically aiming to commit an offence but are conducting their routine activities with alertness to available opportunities, and 'Active Searchers', namely, offenders who set out with the intent to steal, using their experience to identify places and situations in which they are most likely to succeed. Active searchers were more determined to commit an offence and were prepared, if necessary, to wait for these situations to emerge. Copes and Cherbonneau argued that the latter group were likely to be more experienced, applying their knowledge to locate suitable targets, and were more inclined to weigh up the risks and rewards of each opportunity spotted. A third group used force to steal cars and were characterised by greater desperation; their lack of resources for offending was identified as restricting their choices for locating targets and conducting the crime. This group are presented as if they force themselves on

any target regardless of suitability because they do not have the skills to select or the patience to wait.

The available research evidence suggests that the strategies employed in theft from the person/robbery to find and identify victims are most analogous to the 'alert opportunist' (Jacobs, 2010) with a smaller number adopting strategies similar to the 'active searcher'. The third group outlined by Copes and Cherbonneau, using force born of desperation, is a less convincing category when applied to robbery/theft from the person. Later research by the same authors (Copes et al., 2012) has shown that force can be used by both alert opportunist and active searchers, hence the use of force does not automatically indicate desperation, and does not preclude careful target selection (more violent offenders do not necessarily force themselves upon any target indiscriminately). An important insight into robbery target selection comes from Jacobs (2010) who argues that classifications of search strategies are over-simplistic, particularly where they cross-tabulate experience and search approach. Jacobs' research shows that it is not the case that experienced offenders wait until their motivation reaches a given threshold and then embark on a search for appropriate opportunities. To the contrary, Jacobs found that once they had recognised a suitable opportunity, experienced robbers were disinclined to let it pass even when they were not highly motivated at the time. Jacobs' ethnographic study of street robbers and carjackers showcases the ability of motivation to flex based on situational factors (i.e. readiness in CCO terms) and highlights that the transition from indifference to motivation can be extremely rapid and prompted by the presence of a suitable target. Jacobs argues that attributes of the situation can implant the idea of robbery (p.518) meaning that, in a given situation, the presence of an opportunity can spark readiness to offend in someone who is predisposed to offend and possesses the specific moral prerequisites. Jacobs uses the concept of serendipity, the occurrence of beneficial but unforeseen circumstances, to show how offenders frequently take advantage of opportunities that emerge when they were doing something else. In these situations, target search is effectively bypassed as a conscious/standalone stage and decisions are made quickly with target selection, approach and attack occurring almost contemporaneously.

In a discussion, which again highlights the offender as both caused and causing, Jacobs further argues that offenders can manufacture serendipity: i.e. they deliberately manipulate conditions to increase the likelihood that a fortuitous combination of circumstances will emerge. Manufacturing serendipity can involve being at places where opportunities are likely to emerge, a likelihood that is increased by participation in street culture which regularly places offenders in situations conducive to theft (Shover & Honaker, 1992). Jacobs depicts offenders as artfully mixing luck with

preparation<sup>27</sup>. Although luck is important, 'chance is implicated' but it is not the cause (p.516), offenders need the skills to recognise and capitalise on fortuitous circumstances, a capacity coined by Walpole (in Jacobs, 2010) as 'sagacity'. Jacobs does not utilise script frameworks to explain how offenders organise and retain the requisite knowledge and experience required to manufacture and act upon opportunities, but many analogous concepts are referred to. Jacobs refers to 'templates' (p.523) that provide perceptual shorthand and perceptual triggers enabling offenders to map emerging situations to their 'experiential knowledge' (p.520) utilising cues relating to time, place and circumstances. Jacobs continues to explain that templates are updated through learning, and that negative experiences are particularly important in shaping these refinements.

### *Scene: Target selection*

A robbery cannot progress without the recognition and selection of an attractive target. Understanding target selection requires more than listing the features which increase the probability of an individual becoming a victim as the vulnerability of the victim results in a dynamic interplay between the victim, the offender and the situation (Topalli et al., 2015). The process of target selection requires the suitable target to be available to the offender in a conducive wider environment and for the offender to identify and recognise the opportunity. Target selection may be a long drawn out scene, but more commonly, in theft from the person and personal robbery, it is a rapidly made decision. Section 2.2.3 above identified different levels of target; in theft from the person and personal robbery offences, the victim can be regarded as a secondary target who provides additional primary or convertible goods. However, the victim may provide a mixture of rewards to the offender if violence, confrontation or dominance of the victim are ends in themselves, the victim can also be regarded as a primary target.

Research consistently shows that offenders discriminate between crime targets, a finding that has been demonstrated for crime types as diverse as burglary, vehicle crime, shoplifting, robbery, vandalism and sexual offences. In robbery, factors influencing victim selection include the value, to the offender, of property carried by the victim and the value of the victim, as a target in their own right, combined with the attributes of the victim that the offender perceives will facilitate the accomplishment of a successful attack. Valuable theft from the person/robbery victims include those who carry valuable, CRAVED goods (Clarke, 1999) including cash, jewellery and mobile phones (Harrington & Mayhew, 2001) and other high value, portable electronics. Valuable property stolen

---

<sup>27</sup> Jacobs convincingly provides the example of the employment of archaeologists at prior to commencing new building developments as example of steps taken to prepare for a potential, but low probability, fortuitous circumstance.

from robberies can also include drugs (Jacobs, 2010). A form of theft that bridges across personal crime and vehicle crime is carjacking, technically categorised as a robbery in the UK (Home Office, 2015a).

The attractiveness of targets is not solely determined by the nature of the property they are carrying: if the value of committing a robbery includes the thrill (Wright et al., 2006) of conducting the act, then the thrill can be heightened by the type of victim being targeted. Relative deprivation has been identified as a factor motivating and justifying robberies against individuals deemed to be privileged (Deakin et al., 2007; Jacobs et al., 2000), thus students and those displaying signs of affluence have been reported as valuable, justifiable and even deserving targets. De Haan and Vos (2003) found that one of the risks of robbery, expressed by offenders, was the risk of committing a robbery against a target deemed 'low status' by their peers. For offenders, victims often make themselves attractive by their own actions as in the case of 'moralistic robbery' (Jacobs & Wright, 2008) in which the offender has a clear perception that the victims played an active part in their selection. In these cases the formation of readiness and target selection is often simultaneous (Felson, Baumer, & Messner, 2000). In the context of the CCO the victims act as crime promoters as in the eyes of the offender the actions have defined them as 'deserving victims.' Thus, the causes of readiness, predisposition and the role of crime promoters need to be understood to fully understand the choice of targets.

The value and attractiveness of targets and victims are not the only factors taken into consideration by offenders; their vulnerability is also important. This includes victims who appeared nervous or out of place (Wright & Decker, 1997), victims who were distracted by other things, "I prefer to rob somebody that really ain't got they mind on what they doing. You can kind of feel people [and just know] that they ain't thinking about it" (Wright & Decker, 1997, p.97). Jacobs (2010) found that robbers selected victims they perceived would be compliant and less prone to resistance (see also Wright & Decker, 1997). Vulnerability is situationally defined and results from an interaction between attributes of the victim and attributes of the situation, and the wider environment they have entered. Vulnerability can also be generated where victims are unable to call for help: the pressure on school age children not to report the crimes of their acquaintances can increase their likelihood of victimisation (Felson et al., 2000).

Another persistent theme in the literature is the frequency with which robbers elect to rob other criminals, commonly drug dealers (Deakin et al., 2007; Felson et al., 2000; Topalli, Wright, & Fornango, 2002; Wright & Decker, 1997) Drug dealers are vulnerable because they are unlikely to report the offence, and attractive because the robber can steal both drugs and money. Drug dealers

also represent a frequently available target on the street, and bystanders are unlikely to intervene. Topalli and Wright (2004) found that robbers sometimes considered high risk criminal targets when their own level of desperation was high. The disadvantages of this type of attack are that they may be known to the victim, and there is a risk of later recriminations, especially with those higher up the hierarchy. Wright and Decker's robbers stated that in order to attack high level dealers they had to 'catch them napping.' These studies also found that a minority of offenders committed offences against men looking for prostitutes, as these men carried quantities of cash and were less likely to report the offence to police due to embarrassment and shame.

Interestingly, although there are common criteria for vulnerable victims, the types of people that offenders feel meet these criteria appear to vary across offender accounts. Thus in Jacobs' (2010) study, offenders consistently reported that compliance was important but variously suggested that whites, women, older people and drunks met this criteria. This was largely based on their range of experiences with different victim types, thus an offender who successfully robs a female victim may store this experience as a potential future victim type, whereas, another who encounters resistance from a female may be less keen on attacking women in future. Thus criteria for victim selection is often based on experienced based knowledge rather than an assessment of all options.

The factors shaping target selection have been discussed above but really the 'scene' is completed via an active selection on the part of the offender. This step is characterised by speed: victim selection is quick because it needs to be. Selection is frequently a matter of any suitable victim being in the right place at the right time from the point of view of the offender. Feeney (1972) showed that many decisions were quite matter of fact, the victim was in a suitable location and looked to have money. This finding was supported by Jacobs:

"One of the most striking findings to emerge from the interviews is the extent to which targets of predatory robbery materialized almost magically in time and space. Victims 'became' victims after crossing paths with offenders at the right—or wrong (depending on one's perspective) time. In a number of cases, serendipity made the target selection process appear to be almost effortless." (2010, p.518)

Scripts can supply information, or cues, that aid the identification of suitable victims. A routine script performed by an unsuspecting individual could serve to flag their suitability as a victim within someone else's crime commission script. The activities of potential victims can leak signs of vulnerability, such as the manner in which someone crosses a busy city centre street providing an indication that they are not streetwise and/or are unfamiliar with the city (Rogerson, Smithson, &

Hirschfield, 2008). In similar vein Harper (2006) argues that tourists display a range of behaviours that can single them out as being isolated and having little knowledge of their surroundings. Thus interactions and potential clashes between scripts are central to their performance and adaptation. Recognition of criminal opportunities appears to be a skill that can improve with experience (a finding that is also supported by script research and the concept of 'sagacity' discussed above). Topalli et al. (2015) distinguish between procedural (how to) expertise and perceptual expertise, the latter relates to how the offender sees opportunities, both of which can be enhanced with experience. This benefit of experience was directly supported in an experiment conducted with students (Garwood, 2009) which found that the ability to recognise criminal opportunity was directly related to participants (self-reported) offending experience.

From the perspective of Wikström's situational action theory it would be argued that the discovery and recognition of a target has to be linked to something the offender wants in order for the offender to be motivated to convert it into action. Although not explicitly discussed by Jacobs and Wright, their analysis of moralistic robberies, in particular, shows the central place of motivation in shaping target selection. Occasions when the unforeseen actions of another resulted in insult or debasement to the offender, provided a serendipitous opportunity to take revenge or regain the upper hand. Thus the action which provokes readiness to offend, simultaneously provides an attractive and vulnerable target. A suitably resourced offender is able to draw skills, knowledge and sagacity organised with a template (or script) to capitalise on the opportunity.

It is clear that target suitability is co-produced by the attributes and actions of the victim, the offender and the nature of the situation, referred to as "situated vulnerability" (Jacobs, Topalli & Wright, 2006) and while offenders have preferences these are not precise or set criteria (Wright & Decker, 1997). Offender actions can operate to make targets more or less vulnerable. It has also been demonstrated that the nature of the victim selected can influence later stages of crime commission, including how they will be approached and controlled, notably when the victim was another criminal.

#### **5.2.4 Act: Transfer**

Theft from the person and robbery of personal property require item(s) to be transferred from the possession of one person into the possession of another. This act of transfer requires a number of necessary conditions. Firstly it is necessary for the offender to achieve proximity to the victim. For this reason theft from the person scripts require an 'approach' scene. Property must transfer from the possession of the victim to the offender. This transfer is conducted without the permission of

the victim, the transfer scene offers alternative paths through which the offender either circumnavigates permission by committing the offence without the victim's knowledge or gains compliance from the victim; dependent on the track taken transfer can involve engagement with the victim but this is not always necessary. In most cases of theft from the person and robbery, the approach, engagement and transfer scenes are rapid with no clear boundaries between them, the three acts usually occur at the same time and within the same environment. As the location and the actors most often remain unchanged between scenes, applying the dramaturgical analogy suggested in Section 2.3 these will be described as scenes within the same act.

### *Scene: Approach*

An essential accomplishment, in theft from the person, is to gain proximity to the victim (Jacobs, 2012). It is through this scene that the offender achieves offender presence in the situation and more specifically for these crime types, 'co-presence' with the victim (Jacobs, 2012). These thefts are committed in the presence of the victim. However, the victim does not want to be co-present with someone who has the intent to rob them. Given the opportunity, victims would take steps to protect themselves and their property from the undesirable outcome of theft. The goals of the two actors are diametrically opposed, requiring the offender to manage and control the behaviour of the victim; this control begins with the nature of approach (Copes et al., 2012, Jacobs, 2012). For the offender, the imperative that defines the nature of approach is to deny victims awareness of their intentions and consequently deny victims the opportunity to initiate their own counter strategies.

Offenders generally accomplish co-presence through combinations of speed, surprise or secrecy. "Co-presence in carjacking requires that offenders ensnare targets without betraying their predatory intentions" (p.5). This finding is equally relevant to other forms of robbery and theft from the person. Robbers interviewed by Wright and Decker (1997) and by Copes, Hochstetler and Cherbonneau (2011) stressed that the essential ingredient of approach was surprise. Approach is the key stage of these offence scripts where distinct tracks begin to emerge, with choices made here shaping the rest of the script. One available option is for offenders to conceal themselves, aiming to stay 'invisible' to the victim until immediately prior to their contact with the victim. A frequently reported method for doing this was to approach the victim from behind.

An alternative method of approach, identified in many studies including Wright and Decker (1997), Smith (2003) and Copes, Hochstetler and Cherbonneau (2011) is for the offender to conceal their purpose rather than conceal their person. Offenders would aim to deceive victims by 'managing a normal appearance' (Luckenbill, 1981, p.29), giving the impression of a legitimate purpose, such as

asking for the time or for directions. This was the most common method used amongst Jacobs' (2012) carjacking participants, who would approach the vehicle in an inquisitive mode, requesting the time, or assistance, a method which Jacobs notes has long been associated with robbery initiation (Katz, 1988). This notion of functional secrecy can be found in scripts for other crime types including burglary (Cromwell & Olson, 2004; Wright & Decker, 1994) vehicle crime (Michael et al., 2001) and fraud (Tremblay et al., 1998).

An additional means to accomplish surprise, which may be used on its own or in combination with these two tracks, is to take advantage at a time when the victim is busy or distracted, such as when they are at a cash machine (Deakin et al., 2007, Wright & Decker, 1997). Again, this demonstrates how offenders take advantage of the victims' routine scripts. It also highlights how attractive target persons can become more or less vulnerable based on the situations in which they find themselves. Attacks near a cash machine also provide greater probability that the victim will have cash on their person, allowing target selection to blur into approach. Speed is essential to prevent the victim becoming forewarned: consequently, robbers frequently recount a very short time lapse between selecting a victim and acting upon the opportunity (Jacobs, 2012; Topalli & Wright, 2004). The option of identifying a target and then taking time to weigh up the risks and rewards, as identified by Michael, Hull and Zahm (2001) in their study of thefts from vehicles in car parks, is not available in theft from the person scripts. Therefore, in thefts from the person, target selection, approach and transfer are often conflated, occurring rapidly in the course of one or two actions.

As with other scenes, that nature of approach will be shaped by factors in the situation. For example, the track taken for approach will be dependent on whether the environment allows concealment to enable surprise or on whether or not collaborators are available. Male robbers have reported that the presence of a female acquaintance can help to present a non-threatening image (Wright & Decker, 1997). While Alarid, Burton and Hochstetler (2009) found that co-offending amongst robbery offenders led to increased sense of control during the offence, which could encourage the use of a bolder approach strategy.

### *Scene: Engagement*

Engagement has been taken here to start at the point at which the offender makes the victim aware of his presence. It should be noted, however, that approach, engagement and transfer are frequently accomplished with the same actions. Engagement is a scene at which different script tracks begin to diverge. Interaction with the victim, however brief, is necessary for robbery but an

optional scene for theft from the person, where it is possible to accomplish a theft without the knowledge of the victim.

In the case of personal robbery, offenders must engage with the victim in order to communicate a threat sufficient for the transfer of property from victim to offender and for managing the risk of victim resistance. Thus the key aim of engagement is to gain compliance. The offender's aim is to ensure that the victim has an understanding of the risk being posed (Copes, Hochstetler, & Cherbonneau, 2011). Thus contact with the victim is made bold and threatening to ensure that the victim immediately understands the threat and is coerced. From the perspective of the offender, victim resistance is perceived as the greatest obstacle to a successful robbery. Robbers in Copes et al.'s (2011) study were both cognisant of this risk and confident in their ability to manage it. Consequently, an 'understanding of the situation' is frequently communicated by immediate, extreme threats to ensure that the victim understands the consequences of any resistance. The offender must make the victim feel that the threat is 'proximate and real, not distal and theoretical' (Copes et al., 2012, p.12). Threats can be made in the form of verbal commands but also actual force (D'Alessio & Stolzenberg, 1990). By their definition, robberies involve confrontation, although, offenders tend to claim that they avoid using excessive violence, where possible, and aim to use only sufficient violence to induce shock and fear in the victim and to ensure their compliance.

Weapons are clearly an important resource for transmitting a threat. Those offenders who used them, claimed that this was more to communicate a message about the consequences of resistance than out of an intent to use them (Wright & Decker, 1997). The timing of the use of violence marked a distinction between different groups of robberies identified by Smith (2003), with the majority of robbers reserving the use of weapons only to counter victim resistance: only a minority of robbers began the encounter with violent force. Similarly in Copes et al.'s (2012) study, carjackers who initiated victim engagement with assault were rare. Again, the nature of the victim targeted, influences the offender's preparedness to use weapons, and those that decide to target other criminals, did so with the expectation that a higher level of violence, including the use of weapons, would be required (Topalli et al., 2002).

In theft from the person, engagement may occur as a distraction. This can include diverting the target's attention via conversation, a tactic frequently used against tourists (Tarlow, 2006). Katz (1988) identified 'asking for the time' as a tactic of connecting with the potential victim while holding the interaction sufficiently long enough to scan the target for further information. Thus while this activity has been listed here as 'engagement' it also serves to further inform target selection. In their study of bag thefts from licensed premises Smith, Bowers and Johnson (2006)

identified distraction MOs, including making conversation with the victim or showing them a leaflet. The importance of engagement is that it establishes the manner in which an offender will attempt to transfer property, be that by force or stealth, for example. Engagement with the victim is not always necessary in theft from the person. In tracks where property is taken without the victim's knowledge, the offender may take action in order to deliberately avoid interaction with their target.

### *Scene: Search*

In some cases, mainly in relation to robbery of personal property, it is necessary for an offender to locate the target property on the victim. This introduces a short scene of property search. Smith (2003) found that just under a third of robbery victims were forced under threat to turn out their pockets or otherwise reveal their property. In a similar number of offences the offender would conduct the search, 'patting down' the victim to locate items. Wright and Decker (1997) found that searching for property was more common when offenders were working together (allowing one offender to conduct the search while another kept look out). In Smith's study the most common approach (46%) of robberies was for the offender to snatch the property away from the victim. This requires that the property is on view at the time of the offence, a factor which may form part of initial target selection.

### *Scene: Transfer*

To successfully complete a theft from the person or robbery of personal property, the offender must 'permanently deprive' the victim of their property. This requires that the offender removes the property from any target enclosure. For thefts from the person, the victim can be seen as forming the target enclosure, as the property will be in their hands, pockets, bags etc. presenting a physical boundary which the offender must overcome. This target enclosure may present itself to the offender in multiple layers which the offender must breach: e.g. first the victim boundary then a bag, or the target person may, themselves, be contained in an enclosure, such as a car (Jacobs, 2012).

Cohen, Vila and Machalek (1995) examine strategies for the successful misappropriation of property and identify factors that contribute their success. Felson (2009) produced a similar typology of theft 'pathways'; however, Cohen et al.'s classification provides the advantage of considering success, not only in terms of the successful implementation at any given instance, but also, of the successful spread of a strategy relative to alternatives. For Cohen et al. expropriative strategies are more likely to succeed when they are:

- Cryptic, the theft is not detected at time of the offence;
- Deceptive, the activity might be detected but the criminal motive is disguised;

- Bold, the actions are aggressive;
- Surprising, the preparation of a counter strategy is prevented by taking victims off guard;
- Evasive, the strategy avoids or adapts to counter strategies;
- Contagious, the strategy is easily spread through transmission and adoption;
- Stimulating, for the actors performing it, the strategy presents an end as well as a means;
- Mutable, the strategy provides the option of different routes through which it can be performed;
- Mobile, the strategy is easy to pick up, easy to learn or imitate;
- Resistant - resistant to counter strategies, this resistance could be something that is deliberately built in through, for example, fear and intimidation.

These strategies are discussed under transfer but really they have already started to emerge during the discussion of approach and engagement. For example, the description of approach included discussion of strategies that were cryptic, deceptive and bold. Earlier script scenes can be decisive; it is not possible to employ a secretive strategy if the approach to the offender was bold and violent.

Robbery research provides considerable support for Cohen et al.'s characterisation. Based on an examination of 2000 police crime reports and witness statements, Smith (2003) classified robbery offences into five MO categories, Blitz, Confrontation, Con and Snatch, and victim initiated, categories for which there is a significant overlap with strategies outlined by Cohen et al. Blitz is a strategy that is dependent on 'surprising' and bold aggression in which violence is 'the first point of contact between the victim and the suspect,' if there is any verbal communication it follows the initial assault rather than acting as a request or demand for property. Confrontation robberies were bold but, in contrast to Blitz robberies, the contact began with aggressive demands. Threats were, on occasion, followed through with force, but assault was not the first contact. Snatch robberies relied on the element of surprise, the first contact between offender and victim was the physical force used to grab the property away from the victim (in these robberies the property was nearly always on display). In Con robberies, offenders engaged the victims in an interaction with the intention to deceive and/or distract them, thus the initial approach was usually in the form of a 'spurious conversation.' Finally, in victim initiated robberies, it is the victim who initiates contact with the offender and later becomes a victim of robbery. The contact made by the victim frequently related to a drug deal or the procurement of sex, the victims are, therefore, less likely to report the offence due to the embarrassment.

Theft from the person, without threats of violence, can also be understood using Cohen et al.'s strategies: these methods utilise secrecy to ensure that victims are unaware of the offence at the time of commission. This strategy is somewhat challenging given the proximity of the offender to the victim. Johnson, Bowers, Gamman, Mamerow and Warne (2010) identified non-violent methods of theft of bags and theft from bags including:

- Lifting: Picking up unattended items (note that this would not be classified as theft from the person).
- Dipping: Reaching into a pocket or bag. This is a cryptic strategy relying on secrecy.
- Snatching: Pulling the property away from the victim and then escaping. This strategy is based on speed and surprise.
- Slashing: cutting the straps of a bag or bag itself to release the property. This strategy is based on speed and surprise.
- Distraction: creating a diversion to distract the victim's attention away from their property – this is a deceptive strategy, although the victim is aware of the actions of the offender they are not aware of their motivations.

### *Scene: Manage resistance*

As noted above offenders perceive victim resistance as the most important obstacle to theft from the person. Offenders actively aim to minimise the risk of resistance, by way of approach and engagement strategies discussed above. Other tactics include spending as short a time as possible in the presence of the victim (Copes et al., 2012). However, resistance still occurs. Smith (2003) found that a quarter of robberies were likely to involve resistance. In their analysis of US National Crime Survey Data, Kleck and DeLone (1993) identified that offenders are correct to be concerned about victim resistance in their finding that victim resistance reduces the probability of a successful robbery. Kleck and DeLone identified that armed resistance by victims was most likely to be successful; unarmed resistance reduced the offenders' probability of success to a lesser degree but it was also more likely to lead to victim injuries. Kleck and DeLone found that armed offenders were more likely to be able to overcome resistance. As noted above, several studies have shown that offenders who carry weapons frequently do so to communicate a threat, but this threat can escalate to actual violence in response to victim resistance (Luckenbill, 1980 in Kleck & DeLone, 1993).

### *Scene: Escape*

The literature reviewed revealed fewer findings regarding the escape phase of theft from the person and robbery of personal property, although it is acknowledged as an essential scene. The themes of secrecy, deception and aggression, that were threaded through earlier scenes, arguably continue through to the end of the script. Thus offenders utilising cryptic strategies will likely aim to exit the scene secretly. In bold and aggressive script tracks, to prevent the victim retaliating or calling for help, the need to ensure victim compliance continues until the offender has escaped. Offenders interviewed by Wright and Decker (1997) reported two key strategies for achieving this, either forcing the victim to remain stationary through threats or violence or by making the victim leave the scene first. Conklin (1972 in Kleck & DeLone, 1993) demonstrated how offenders used weapons to reduce the threat of pursuit by the victim.

Research has shown that for thefts from the person, as with other crime types, the availability of and the offenders' awareness of, escape routes can influence the choice of location for crime commission. This finding is supported for robbery (Deakin et al., 2007) and theft of bags from licensed premises (Johnson et al., 2010).

### **5.2.5 Act: Post conditions**

As with escape, little has been written regarding the post conditions of theft from the person and robbery. However, findings from research into other crime types highlight that there are later stages of import for the offender in obtaining their goal. In their analysis of vehicle crimes, Michael et al. (2001) identified several stages that followed the theft. These were the examination of property stolen, the safe caching of property for later retrieval, the discarding of unwanted property, the conversion of property into cash or other goods and then finally the consumption of the rewards from the theft. Each of these stages is shaped by wider situational factors such as the offenders' awareness of the availability of suitable cache sites and the availability of networks to convert stolen goods into money or drugs.

After the theft offenders may face retaliation from their victims, particularly where these were other criminals and/or were people known to them. Methods for minimising resistance included adopting a particularly fierce and intimidating persona and using excessive violence during the event to after-the-event tactics such as being cautious not to brag about the event, avoiding the scene of the robbery and maintaining hyper-vigilance.

### *5.3 Summary*

This chapter has used script and CCO frameworks to organise existing research findings about theft from the person and robbery of personal property. This has helped to identify different sets of methods that offenders use to commit these crimes and to highlight the differences between them. It also aids the understanding of the reasons behind those differences. Applying the frameworks has helped to identify and collate available research evidence relating to the key scenes in theft from the person commission. This has included the identification of the sub-goals which must be accomplished to consider a scene 'complete'. These scenes and goals are summarised in Table 5.1. The table shows that some scenes are always necessary to successfully complete a theft from the person (i.e. goal formation, target selection, approach, transfer and escape). The performance of other scenes is dependent on the specific script track taken (preparation, search for target, engagement, manage resistance, convert goods, manage retaliation).

Importantly, the frameworks have encouraged consideration of the interactions between different actors in the scripts. Wikström (2009) noted that intended actions can be subject to interferences, including human and inanimate ones. The script framework has shown how experienced actors are able to anticipate these interferences and drawing on scripted alternatives to avoid, manage or respond to interferences. The literature also demonstrates how offenders incorporate experience into script revisions, including lessons for unsuccessful attempts (Jacobs, 2012).

The reviewed research lends strong support to the applicability of situational crime theories. They have demonstrated that offenders display a decision making calculus, they discriminate between targets, anticipate risks to achieving their goals and apply strategies to minimise these risks including the instrumental (not indiscriminate) use of violence. Use of the CCO framework has shown how the proximal causes within the crime situation, not only determine whether or not a crime will occur but also continue to shape the nature of the offence as it unfolds. All of the CCO elements operate globally across all script scenes, but there are certain scenes where particular CCO dimensions have extra relevance. For example, the nature of target enclosure for transfer, resources (weapons) for engagement with the victim. Crucially, the review of literature has demonstrated the importance of understanding the nature of criminal motivation, rather than taking it as given, understanding how different goals, experience, emotions and morals shape offenders' perceptions of the crime situation, and in turn, shape whether and how they will act upon crime opportunities. Criminal motivation is not simply a case of a decision to offend or not, criminal motivation affects the type of offences that is committed

Findings about some scenes are more abundant about some scenes than others. However paucity of findings is not the same as proving that a scene or act never happens, some acts may be considered as plausible even though they have not been explicitly discussed in the literature, but such propositions should serve as pointers for future research.

While script frameworks have been useful, it cannot be denied that the division of theft from the person and robbery of personal property into discrete acts and scenes does at times feel artificial. While the research lends support to the existence of scripted procedural knowledge, in reality this operates more flexibly than it is possible to present in written form, particularly when the interaction of different actors' scripts are considered.

Table 5-1 Acts, Scenes and Sub-Goals in Theft from the Person Scripts

<b>Act/Precondition</b>	<b>Scene</b>	<b>Sub-goals for scene completion</b>
<b>Readiness%</b>	Goal formation%	Predisposed offender is readied to offend with lack of skills to avoid crime. Anticipation of risk, effort and rewards
<b>Preparation≠</b>	Preparation/≠	As above plus Offender is/perceives is resourced
<b>Target Selection%</b>	Search for target#	Offender situated in area with a likely supply of targets% May include accessing conducive wider environment Potential target recognised
	Target Selection%	Recognition of and decision to target an attractive and vulnerable person%
<b>Transfer of Property%</b>	Approach%	Offender is co-present with victim%
	Engagement#	Victim misunderstands the situation# Offender and victim have common understanding of the situation# Offender gains compliance from the victim%
	Search for property#	Locate property on victim
	Transfer%	Any target enclosures breached# Prevented/Managed victim resistance% Permanently deprive victim of property%
	Manage resistance#	Maintain victim compliance#
<b>Escape</b>	Escape%	Leave scene with property%
<b>Post Conditions</b>	Convert goods#	Convert goods into consumables#
	Manage retaliation#	Minimise the risk of retaliation post robbery#

% Necessary condition - Script dependent on this stage

# Track dependent - Script tracks are available that do not include this scene

## Chapter 6 Understanding the Limitations of Police MO

### Descriptions

The overall aim of this thesis is to explore the potential of police recorded MO descriptions as a source of empirical data for conducting analysis of the nature of offence commission for volume crimes. The following chapter reports on a detailed inspection of MO descriptions conducted on raw data from Force A. The review identifies the key characteristics of these data and moreover highlights those features of MO descriptions, as they are currently recorded, which present challenges for their analysis.

The chapter begins with a description of the processes through which crimes and their related MO descriptions are recorded. The chapter then presents the results of an analysis of a sub-sample of MO descriptions which demonstrates how a combination of the recording process, the nature of natural language and the nature of crime influence the accuracy, clarity and detail of the information captured.

The analysis in this chapter is based on a sub-sample of MOs from Force A, it also draws on the author's past experiences of analysing MO descriptions from approximately 15 different UK police forces. Based on this experience the MO descriptions analysed in this thesis appear broadly representative of police databases throughout the UK.

#### *6.1 Background: The Process of Police Crime Recording.*

The processes and standards for recording both crimes and incidents is stipulated by the NCRS which is reproduced in the Home Office Counting Rules (HOCR, introduced in 2002 and most recently updated August 2015). The NCRS and HOCR aim to ensure consistent data is collected across all 43 England and Wales forces. Her Majesty's Inspectorate of Constabulary (HMIC 2012) outline the stages of crime recording as:

1. A member of the public calls for police assistance, or a police officer observes or discovers a crime. The police create an incident record.
2. If the police decide a crime was committed, they create a crime record (usually straight away).
3. If the police decide a crime was not committed, the incident is 'no crimed', the record remains within the incident file, but a crime record is not created.

4. Where a crime is committed investigations begin as soon as possible, usually with initial enquiries which look for possible leads and gather evidence (a 'primary investigation'). A more detailed, 'secondary investigation' then takes place to consider the evidence gathered in the initial stages."

All police forces in England and Wales maintain crime and incident databases which may be held within the same system, two linked systems or two independent systems. In addition to the core crime databases, detailed investigation files may be kept in other systems, including paper files. The fields that are examined in this chapter and analysed in Chapter 7 to 9 are taken from Force A's crime recording system (incidents were recorded in a separate but linked database). Detailed investigation files were not consulted. This follows from the intention to focus the analysis on types of data that are readily accessible to all forces.

A series of Audit Commission studies has reported overall improvements in the quality of police crime and incident data since the introduction of the NCRS and HOCRs. For example, in 2006/7 the commission rated crime recording as 'good' in 79% of forces compared to 27% in 2003 (Audit Commission, 2007). A National HMIC 2012 review (HMIC, 2012) inspected the quality of crime and incident data and includes data on Force A's procedures for data recording, audit and quality control. The review reported that the quality of incident and crime records was high; and fully compliant with NCRS. The HMIC expressed some concern that in this force a higher than expected number of incidents had been incorrectly closed without creating a crime record. The review noted that Force A's IT systems were outdated and had limited capabilities, which increased the risk of recording crimes and incidents incorrectly. In fact, the report noted that the multiplicity of IT systems in use in the force, including both electronic and paper-based records, overcomplicated the process of auditing data quality.

More recently, the UK Statistics Authority has raised grave doubts over the quality and integrity of police recorded crime data to the extent that they have removed the 'national statistics' status from this data source citing problems including non-recording of crimes, downgrading of crimes to less serious offences and only recording detected crimes (HMIC, 2014).

### 6.1.1 Recording MO

Several fields within police crime recording systems, including date, location, and offender and victim characteristics, are required to piece together a picture of 'what happened' during a criminal offence. However, police crime recording databases commonly contain a field (occasionally multiple fields) labelled 'MO' which provides details of the methods used to commit the crime. The analysis

of data limitations in Section 6.2 and the analysis of police data in Chapters 7 to 9 relates to this MO field. Although there is no national standard format for the reporting and recording of MO details operandi, common practices persist across police forces, these are shaped by commonalities in recording systems and standard aide-mémoires for MO recording that are introduced in training. The MO field is most often a free text field containing a short, usually between 40 and 150 words, narrative describing the crime event, including the central actions and actors (offenders and victims). Some forces use feature codes or keywords to highlight pertinent aspects of an offence MO. These codes are used either instead of or in conjunction with free text fields; however, in Force A such keywords were only recorded for offences of burglary and not for the offences of theft from and the person and robbery.

From personal experience and anecdotal evidence, it appears that MO fields are completed more comprehensively than in the past, although there is no research evidence to confirm this. Until recently, reviews of police data quality have paid little attention to the quality of the MO field. Their enhanced importance, which has arisen from their function in the determination of crime classification, is reflected in their inclusion as an indicator of quality in HMIC reviews (2012). The 2012 review included an assessment of whether “the modus operandi (MO) [is] sufficient to confirm the classification of crime?” It might be assumed that a field lacking in such details would also be of little use, or at best have serious limitations, for other purposes. The review found that the performance of police forces against this standard varied considerably. In the best performing forces, 100% of cases were completed to this standard, in the worst performing force only 59% of cases met this standard, with an average of 95%. In Force A, 100% of cases met this standard.<sup>28</sup> In the data received, MO fields were completed for the majority (over 99%) of crimes and were generally completed comprehensively. It is, of course, difficult to judge the accuracy and completeness of MO records without being in possession of the full facts of the case.

The remainder of this chapter will report on a review of a sample of raw MO descriptions drawn from police data. The review sought to identify issues with the nature and quality of MO descriptions which may affect their analysis. The review identified that these data are shaped by a mixture of factors stemming from the nature of the events being recorded, the systems into which data are being entered, the skills and mind-set of the recorder which is, in part, shaped by the broader organisational culture and the nature of the language used to describe them and the nature of natural language accounts. Descriptions and examples of these characteristics of the data are given

---

<sup>28</sup> Freedom Of Information request submitted to HMIC

in the sections that follow. The examples are faithfully reproduced from the crime information system, only confidential or sensitive information has been redacted and this is indicated by square brackets [...]. Note that the MO descriptions frequently refer to IP (Injured Party), or AP (Aggrieved Party) to refer to the victim. Abbreviations and terminology commonly used in the MO descriptions are included in the list of abbreviations on page 15. A mixture of crime types are used to provide an illustration of the full nature of MO fields. It should be noted that this section aims to highlight common issues with MO descriptions in police data, it does not constitute a full audit or attempt to quantify the prevalence of each of these problems in the data.

## *6.2 The Quality of MO Descriptions*

Throughout this thesis, MO data are treated as representations of actual events in the real world. As 'official data' (Irvine, Miles, & Evans, 1979) crime data are shaped by organisational procedures and, therefore, represent police processes as much as reflect the crimes themselves. MO data should not, therefore, be treated as a completely accurate record of everything that happened during an offence and the discussion that follows highlights some of the processes that lead to variation between MO descriptions and any 'true course of events'. Nevertheless, despite their limitations, this research regards the MO descriptions as a reasonable reflection of the behaviour of offenders. MO descriptions are thus taken at provisional face value unless there are signs indicating to the contrary.

### **6.2.1 MO fields: Issues that reflect the nature of crime**

Foremost of the social factors shaping crime data, is the process of crime reporting. Before dissecting the nature of MO fields, it should be remembered that, according to crime surveys, the police forces in England and Wales come to know about an estimated 38 per cent of all crimes. Of the crimes of interest to this thesis, an estimated 47 per cent of robberies and 29 per cent of thefts from the person are reported to the police (Chaplin et al., 2011). For the purposes of this research, underreporting presents a problem if the crimes, about which the police are aware, have a distinctly different pattern of MO from those that go unreported. It may be the case that attempted crimes are less likely to be reported than successful crimes, and these could execute distinctly different methods. Failed attempts at offending are important as they can highlight stages in offender learning and/or factors that are currently successful in blocking crimes. Also, less likely to be reported are MO that leave victims feeling foolish or humiliated, where the victims themselves fear arrest or where the theft takes advantage of victims in vulnerable circumstances (Felson et al.,

2000). For example, robberies both of and by sex workers are likely to be under-reported while also displaying distinct MO.

In addition to providing an incomplete list of crimes, police recorded crime data only provide a limited subset of information about these crimes. As argued in Chapter 2, crimes are complex processes, which encompass a sequence of actions, at least some of which will take place either before and/or after the commission of the legally defined criminal action. These wider behaviours are of potential interest for the present purpose and for preventative practice. However, MO data, as recorded within police recorded crime information systems, tend to capture only the point in the sequence that is proximate to the commission of the legally defined criminal act. Occasionally, there is information about the period prior to an event, for example, describing the victim's activities immediately prior to the crime, but the focus is largely on the crime *event* itself. This can leave much of the full crime script 'unknown', although, the information may be recorded elsewhere, or stored in the minds of officers investigating the case. For example, Rogerson, Smithson and Hirschfield (2008) interviewed police officers about patterns of robbery MO and compared these to crime data. The officers had identified an MO that involved groups of offenders, with one offender travelling on a bus and his accomplices following behind in another vehicle. The robbery would be committed as the bus approached a stop, with the offender jumping out of the bus and into the car to make a quick escape. Within the police data, no cases could be identified that described the full extent of this crime, as the data within crime records tended to be restricted to what happened on the bus, the 'scene of the crime'. The victims would be unaware of the full MO at the time of reporting. This information was only discovered later by officers conducting full investigations. With possession of this more in-depth knowledge, it is possible to identify cases that hypothetically could be an instance of this MO as the partial information available matches part of the fuller crime script.

The nature of crimes means that much of the activity around an offence remains unknown; not least because offenders have a conscious aim to keep it that way. MO fields are often, therefore, an imperfect trace of real world events. MO fields can, at best, only reflect what is known about an offence and are, therefore, influenced by information of which victims, witness and police personnel are aware and are willing to report. Moreover, with crime there is always the possibility that there will be multiple versions of events. Witness statements can be combined and validated with the interpretation of the traces that are left behind after the offence. For some crimes, this available information will be extremely limited: for example, there tends to be better detail for crimes such as theft from the person and personal robbery where the victim is present, and for burglary, where evidence is generally left at the scene. In contrast, when a phone is pickpocketed at some unnoticed

time, there may be little information to record. Crimes such as theft of a vehicle can pose a particular problem for recoding MO as the 'crime scene' is effectively removed following a successful event.

Sometimes there is insufficient detail to provide a clear picture of what happened, this may be the result of a lack of evidence, the inability of victims to recall details and or the reluctance of victims or witnesses to provide information to the police. This is demonstrated in the examples below. In the first description below, the report that the victim has given is unclear and the police state there is insufficient evidence to classify the crime as a robbery. In the second example, the details of what occurred are unclear due to the victim suffering from concussion.

**\*\*ORIGINALLY CRIMED ON [date]\*\***IP [INJURED PARTY] STATES THAT TWO UK OFFENDERS STOLE #400 FROM HIM WHILST AT HIS HOME ADDRESS AND HAD BEATEN HIM UP RESULTING IN TWO BLACK EYES**\*\*PLSE VIEW LATEST [REFERENCE NUMBER]** UNABLE AT THIS STAGE TO ESTABLISH IF THIS IS A ROBBERY IP ALLEGES OFFENDERS BEAT HIM UP IN HIS OWN HOME IN ORDER TO STEAL. IF FURTHER EVIDENCE PRESENTS TO NEGATE THIS WE WILL REQUEST RECLASSIFICATION AT THAT STAGE.

B.M.T APPROX 4 MALES HAVE APPROACHED IP FROM BEHIND- IP HAS BEEN KICKED TO FACE ONCE CAUSING HIM TO FALL AND SUSTAIN CUTS TO CHIN AND MILD CONCUSSION. OFFENDERS MADE OFF- NOTHING STOLEN. (RE-CLASSIFIED TO ASSAULT W/I TO ROB - WITNESS HEARD OFFENDER SHOUTING `GIVE US YOUR CIGARETTES` AT THE TIME OF THE ASSAULT- IP DOES NOT RECALL THIS AS CONCUSSED. ON BALANCE OF PROBABILITIES ASSAULT W/I TO ROB -)

In the examples below, the details are sketchy, partly as a result of the victims' reluctance to report them.

"ON [DATE] IP WAS APPROACHED BY 3 IC1 MALES AND ONE SAID PAKIS ARE NOT ALLOWED IN THIS AREA"". ONE MALE GRABBED IP AND ANOTHER STOLE HIS MOBILE PHONE. OFFENDERS MADE OFF. RACIAL INCIDENT. IP REFUSED TO TELL OFFICER ANY FUTHER DETAILS. "

ON DATE AND BETWEEN STATED TIMES UNKNOWN OFFENDERS APPROACHED THE IP AND HER FRIEND WHILST ON A NUMBER [NO] BUS. ONE OF THE GROUP GRABBED THE IP'S MP3~PLAYER OUT OF HER HAND AND WHEN SHE TRIED TO GET IT BACK PUSHED HER AWAY. OFFENDERS MAKING GOOD THEIR ESCAPE ON [TOWN] HIGH STREET. (IP DECLINED STATEMENT AS SHE WAS ON HER WAY OUT ON A NIGHT OUT WITH FRIENDS**\*\*BEEN UNABLE TO CONTACT SINCE\*\***

The uncertain nature of the information contained within an MO record is frequently represented in the language used to record them. Hence we find frequent use of phrases such as, "The victim

stated” or “the victim believes” or, as in the example below “From his account”, reflecting that the data is a version of events, rather than a definitive statement of fact. In some cases the police may have reason to doubt the validity of information provided to them by a victim or witness. As shown in the examples below this can include suspicions of exaggerated and false crime reports.

BMT OFFENDER HAS REPORTED TOMV AND PROVIDED STATEMENT TO POLICE. OFFICERS BELIEVE FALSE REPORT DUE TO VEHICLE HAVING BEEN INVOLVED IN NON STOP RTC 40 MINUTES BEFORE TOMV REPORT. IP STATES THAT VEHICLE WAS TAKEN FROM HIM AS A RESULT OF ROBBERY

BMT IP AND TWO FRIENDS WERE APPROACHED BY TWO UNKNOWN MALES AT OFFENCE LOCATION. ONE MALE HAS SEARCHED THE POCKETS OF A 2ND IP ASKING WHAT THE IP HAD FOR HIM - HE HAS THEN ASKED THE SAME OF THE IP - SEARCHED HIS POCKETS TAKING A MOBILE WHICH THE IP SNATCHED BACK. BOTH IPS FEARED VIOLENCE - OFFENDER DETAINED BY SECURITY AND ARRESTED. CROSS REF [CRIME NUMBER] CONSIDERATION GIVEN TO RE-CLASSIFYING AS ROBBERY - HOWEVER – [PC NAME] SPOKE WITH AND WITNESS OVERHEARD IP'S EMBELISHING THEIR ACCOUNT AND THE OFFENDER NEVER THREATENED

In other cases, it is unclear which reported facts are relevant to the crime and which may be incidental. As an illustration, in the example below it is not known whether the victim's visits to the cash machine and branch counter are in any way connected to the offender's crime script. The inclusion of these facts in the crime data suggests that either the victim or the police felt that there may be a connection but there is no solid link<sup>29</sup>.

BMT- IP ATTEMPTED TO USE CASH POINT THE [BANK NAME] ON THE [ROAD NAME]- THE PIN WAS INPUT TWICE- BUT ATM WAS OUT OF ORDER. THE IP WITHDREW #100.00 FROM OVER THE COUNTER INSTEAD. THE IP THEN HAS WALKED DOWN TO [SUPERMARKET] TO SHOP- WHILST SHOPPING. THE IP WAS APPROACHED BY AN EASTERN EUROPEAN FEMALE WHO ASKED HER QUESTIONS ABOUT CHOCOLATE AND NUTS. THE IP FELT A BUMP TO THE BACK BUT CONTINUED WITH HER SHOPPING- ON RETURNING HOME SHE FOUND HER DEBIT CARD WAS MISSING FROM HER HANDBAG ALONG WITH CASH.

## 6.2.2 MO fields: Issues that reflect the nature of crime recording

Once a crime has been reported, a key administrative process that further shapes the recording of the event is the allocation of official crime classifications. In England and Wales these are stipulated by the HOCs. However, it should be noted that offenders do not think or act within the bounds of

---

<sup>29</sup> Note that this is a rare example of an MO which provides information for a longer timeframe than the immediacy of the crime event.

Home Office crime categorisations. Offenders do not set out to commit an 'attempted burglary'; they set out to attempt to commit a successful one. The category against which a crime is eventually recorded will not necessarily reflect the planned script from the point of view of the offender, with the eventual classification reflecting one of many possible narratives. Therefore, two offences with different classifications may be the consequence of the same initial crime script. This is demonstrated in the example below in which it is not possible to ascertain from the MO field alone whether or not the offenders' encounter with a victim was anticipated in advance and whether the offender was operating within an existing script pathway that included violence towards occupants

ON SAID DATE BETWEEN MATERIAL TIMES STATED U/K OFFENDER GAINED ENTRY TO PREMISES VIA FLAT ROOF AT REAR OF BUILDING BY CUTTING A HOLE IN CEILING THE OFFENDER HAS ENCOUNTERED THE VICTIM INSIDE THE BUILDING WHERE HE HAS STRUCK HER WITH A HAMMER TO HER HEAD CAUSING INJURY LISTED OFFENDER MADE GOOD HIS ESCAPE VIA THE FRONT OF THE BUILDING OUT THE FRONT DOOR

Conversely, specific crime scripts that are of interest to analysts are not necessarily reflected in Home Office crime categories. This problem was faced by Donkin and Wellsmith (2006) in their identification of burglaries where cars were stolen. In a similar vein, incidences of 'carjacking' may be of interest to an analyst, but these do not represent a distinct crime category. In the case below, the data includes a helpful flag to aid searches for 'carjacking.'

\*CARJACKING\* BETWEEN MATERIAL TIMES STATED- OFFENDER(S) UNKNOWN HAVE ATTACKED IP IN HIS BLUE SUBARN IMPREZA [REGISTRATION] WHILST HIS IS PARKED OUTSIDE SHOP WAITING FOR HIS GIRLFRIEND. TWO UNKNOWN IC3 OFFENDERS HAVE OPENED DRIVER'S DOOR AND STRUCK IP TO HEAD WITH BAT- DRAGGING HIM FROM VEHICLE. OFFENDERS HAVE MADE OFF IN VEHICLE IN COMPANY WITH 2 FURTHER IC3 OFFENDERS AND SECOND VEHICLE.

This tag is not quite as helpful as it might seem, as the use of such flags is inconsistent both within and between forces; there were plenty of similar cases in this dataset that do not carry this 'tag', as in the example below. Further, flags will only exist for crime types once an MO is recognised as a problem.

UNKNOWN OFFENDER HAS APPROACHED VICTIM IN VEHICLE BLUE AUDI [REGISTRATION] AND SAID TO VICTIM GIVE ME THE KEYS VICTIM HAS THEN BEEN DRAGGED OUT OF THE VEHICLE BY HER HAIR AND CLOTHING OFFENDER HAS THEN MADE OFF IN VICTIMS VEHICLE

These observations serve to highlight the risks of conducting an analysis that narrowly focuses on specific official crime classifications. Consequently, for this thesis, the analysis was based on a range of crime categories relating to theft from the person.

MO fields are brief, in the force in question the descriptions averaged 50 words with longer descriptions reaching around 200 words. The need to record information concisely and rapidly limits the amount of detail that can be provided and can present a risk to quality. Recording must be limited to details that are deemed important for police purposes. In determining which information to record, the police have traditionally prioritised that which will assist with the identification of offenders and the detection of crime. The introduction of the NCRS in 2002 brought an additional priority, that is, information that can assist in the classification of crime but there was still no emphasis on preventative intelligence. However, the priorities for detection and classification may not be the same as what is needed to inform crime prevention.

From all of the available information, those responsible for formulating the record need to select the pertinent details of the offence and actively attempt to construct or synthesise a mini-story. Many descriptions display a degree of skill in being able to communicate detailed information concisely, as shown in the examples below:

OFFENDER(S) HAVE GAINED ACCESS TO REAR GARDEN VIA INSECURE GATE. RUBBER BEADING HAS BEEN REMOVED FROM 2 SIDES OF LARGE WINDOW PANE IN DOUBLE GLAZED DOOR. WINDOW APPROXIMATELY 1/2 M X 1M. NO DAMAGE CAUSED TO WINDOW. NO ENTRY GAINED. OFFENDER(S) MADE GOOD THEIR ESCAPE. DOOR IS UPVC

UNKNOWN OFFENDER(S) HAVE GAINED ENTRY TO REAR GARDEN BY CLIMBING OVER LOW PERIMETRE FENCE- OFFENDER USING A (SCREWDRIVER) HAS ATTEMPTED TO GAIN ENTRY TO REAR WOODEN DOOR BY INSERTING SCREWDRIVER BETWEEN WOOD AND LOCK HINGE CAUSING WOOD TO SPLINTER. NO ENTRY GAINED

Although there is a reasonable consistency about the information provided in descriptions, the details provided can vary. For instance, while both of the examples, below, relate to burglary dwelling offences, the first provides details of the type and location of the dwelling, while the second focuses on information relating to the point of entry. This may be a result of either the available information or the focus of the recorder.

PERSON/S UNKNOWN APPROACH DETACHED PRIVATELY OWNED PROPERTY ON QUIET RESIDENTIAL STREET VIA THE REAR GARDEN TO LARGE WOODEN FRAMED WINDOW INTO STUDY, CONDUCT UNTIDY SEARCH OF THE ROOM AND REMOVE VARIOUS ELECTRICAL GOODS.

AT STATED TIME UNKNOWN PERSON APPROACHES THE SIDE WOODEN FRAMED KITCHEN DOOR VIA OPEN DRIVEWAY AND USING A HALF INCH IMPLEMENT PRISES DOOR 5 TIMES AROUND THE LOCK AREA. OFFENDER/S ALSO GOUGE WOOD REMOVING OUTER HANDLE. ENTRY IS GAINED INTO THE KITCHEN.

Given the brevity of fields, to maintain clarity it is essential that material is organised to aid comprehension. While most MO descriptions are comprehensible, some require a degree of interpretation, such as the example below where the reader must assume that the “and was hit” refers to the victim being hit by the offender.

THE IP WHO IS A 14 YEAR OLD BOY WAS STANDING WITHIN THE GROUND FLOOR AREA OF [RESIDENTIAL FLATS] [TOWN NAME] WHEN A UNKNOWN MALE APPROACHED HIM AND WAS HIT IN THE CHEST AND GRABBED IP' S MOBILE PHONE FROM HIS HAND.

Some descriptions are so disorganised that significant effort must be invested in order to understand the events being described:

BETWEEN MATERIAL TIMES IP HAS LEFT KEYS TO VEHICLE WHICH HE HAS BORROWED FROM HIS FATHER A [MAKE AND REGISTRATION] WHICH IS BLACK IN COLOUR WHILE AT A PARTY ON TABLE OUTSIDE ADDRESS IP NOTICED KEYS WERE GONE AS WELL AS VEHICLE A MEMBER OF THIS PARTY ALSO LEFT THE PARTY OFFENDER MADE GOOD HIS ESCAPE (VEHICLE RECOVERED WITH ACCIDENT DAMAGE CAUSED TO IT).

OFFENDER UNKNOWN HAS FOLLOWED VICTIM TO THE AISLE WITH CAKES -WHILST VICTIM HAS BEEN SELECTING ITEMS OFFENDER HAS MADE 2 ATTEMPTS TO TAKE AN ORANGE CARRIER BAG CONTAINING ITEM FROM ANOTHER BAG - VICTIM WAS HOLDING. OFFENDER HAS BEEN SUCCESSFUL ON THE 2ND ATTEMPT AND ONCE IN POSSESSION OF VICTIMS BELONGINGS HAS MADE A CASUAL EXIT FROM THE STORE WITHOUT PURCHASING ITEMS OF HER OWN AND THEN MAKING GOOD THEIR ESCAPE.

A further risk is that MO fields are left incomplete as below, with no clear reason for their truncation.

BETWEEN TIMES ON DATE STATED IP WAS WALKING HOME ALONE- SEES 3 MALES. ONE IS HEARD TO SAY HE IS ON HIS OWN ROAD

Despite the need for brevity, the MO fields still contain a lot of information which appears redundant. In some cases, the fields contain information that may be useful for other purposes but not to this research; of greater practical concern is information that represents an unnecessary duplication of other fields. As will have been noted in the previous data extracts, the MO fields frequently contain phrases such as ‘at offence location’, ‘between material times’. These phrases are representative of the localised grammar but do not add anything to our understanding of the crime.

The regularity with which these phrases are used demonstrates the influence of training and culture. The use of these phrases also demonstrates the recorders' attempts to construct a self-contained story with time, place and event – even though parts of this story are covered by other fields in the data. This is an issue that could be rectified beneficially with training/guidance saving time in the recording process and freeing space for additional details to be recorded. These phrases create additional problems because they are applied inconsistently and are, therefore, difficult to remove via automated processing. For example, we can see the use of Between Material Times but also BMT, B.M.T and B/M/T. Similar variations exist for phrases such as Between Times Stated and Make Good Escape. When a human analyst reads the data, these terms can simply be ignored as irrelevant, however, when automated approaches are applied, there is a need to apply rule sets which define which stopwords to be ignored.

In the Force A data it was observed that, in addition to the describing the offence, MO descriptions often contain information pertaining to the recording process. For example, the extracts below include comments about a data audit, classification and decisions relating to the crimeing of reports.

BETWEEN TIMES STATED- UNKNOWN OFFENDERS HAVE APPROACHED VICTIM AND DEMANDED PROPERTY. WHEN IP HAS REFUSED- ONE OFFENDER HAS PROCEEDED TO PUNCH IP TO REAR OF HEAD. OFFENDER HAS ATTEMPTED TO GRAB IP AROUND HEAD. VICTIM HAS FORCED OFFENDER OFF AND MADE HIS ESCAPE. AUDIT FAIL WRONG CLASSIFICATION RECTIFIED BY ASSESSOR

B/M/T U/K OFFENDER/S HAVE UNZIPPED TROUSER POCKET AND STOLE MOBILE PHONE THEN MADE GOOD ESCAPE. IP DID NOT FEEL ANYTHING OR SEE ANYONE. NO SUSP- WITS OR CCTV. IP STATES COULD HAVE LOST PHONE- TO DCI FOR NO-CRIMING DECISION

DISCUSSED WITH DATA QUALITY AT [HQ]. THIS [CRIME REFERENCE] HAS BEEN TAKEN OUT AND CRIMED AS PRINCIPAL OFFENCE IN ACCORDANCE WITH HOCR. (ATTEMPT ROBBERY) RE IP'S STATEMENT AND IP LED. X REF WITH [CRIME REFERENCE] (CRIMINAL DAMAGE).THIS HAD BEEN INCORRECTLY RECLASSIFIED FROM ATTEMPT ROBBERY BY F1. AND DETECTED WITH LESSER OFFENCE OF CRIMINAL DAMAGE. IN VIEW OF THIS BEING DONE WE HAVE BEEN INFORMED TO TAKE OUT THIS FURTHER NUMBER AS THE PRINCIPAL OFFENCE NEEDS TO BE CRIMED AND WILL REMAIN UNDETECTED. ALL DETAILS

As seen above, this administrative information can end up dominating the whole field leaving little or no description of what occurred. While this type of information is important for the audit trail required to justify the process of 'crime finalisation' (crimeing events and classifying them according to Home Office guidelines), it adds nothing to the understanding of the crime occurrence and would

probably warrant its own field. However, the design of current data recording systems may preclude the addition of new fields and even if such a modification were technically possible, there may be other administrative/time saving pressures requiring the number of fields to be kept to a minimum. Where it is possible to identify these terms in advance they can be added to a stoplist and ignored within textual analysis.

The tense used to describe an MO can vary both between and within individual MO descriptions. Descriptions are frequently reported in perfect tense rather than preterite which is the norm for story-telling. The example, below, shows how multiple tenses have been used within the same description. This shows that the analysis of text must be able to recognise the same words when they are reported in a different tense. As explained in Chapter 4 this is handled via a process known as stemming.

"OFFENDERS UNKNOWN (X2) HAVE SEPARATED IP AND FRIEND- []- 15 YEARS WITH ONE OFFENDER APPROACHING IP FROM BEHIND AND PLACING COLD SHARP OBJECT TO NECK/THROAT. THE IP HEARD A VOICE SAY GIVE ME YOUR GOLD"". OFFENDER MADE OFF WITH PHONE ONLY. IP RAN FOR HELP. OFFENDERS MADE GOOD ESCAPE IN UNKNOWN DIRECTION"

Moreover, the recording officer may sometimes use the phraseology of the original reporter, while at other times, they will cast the description in occupational terminology. While understandable, the use of this language can cause problems for some methods of pattern extraction, as such phrases increase the variety of terms used to describe.

For concision, abbreviations are commonly applied. The author is aware of some variations in these abbreviations between police forces e.g. the use of AP (Aggrieved Party) rather than IP to refer to the victim. However, within force inconsistency can also be identified and some abbreviations appear quite ad hoc for example the use of ST (Street), SUS (Suspect), UK (Unknown) etc. At times those recording the MOs use their own abbreviated words for brevity, the example below contains AMBO as an abbreviation for ambulance. There are parallels here with trends in 'text speak' and the use of social media networks. These characteristics of MOs all reflect the pressures of time and concision required when recording information but present challenges when attempting to extract key terms from the data via automated processing.

BETWEEN TIMES STATED - U/K OFFENDERS HAVE MET THE PIZZA DELIVERY DRIVER O/S THE PREMISES WHERE HE HAD A DELIVERY - ONE OFFENDER PUNCHED THE DRIVER TO THE FACE AND TOOK THE PIZZA OFF HIM - WHILST THE OTHER OFFENDERS PUNCHED HIM AROUND THE BACK OF THE NECK AND HEAD - 3 X OFFENDER/S IN TOTAL - IP SUSTAINED A SWOLLEN BOTTOM LIP - 2 LOOSE TEETH - AND A CUT GUM -

AMBO CALLED FOR TREATMENT - OFFENDER/S MADE GOOD ESCAPE WITH THE PIZZA  
ALONG [XXX] ROAD - PIZZA WAS STILL IN THE HEAT BAG - ANS SUNDRIES TO THE  
VALUEOF #48.2

In addition to the problems already discussed, the MO descriptions are beset with common errors, including spelling mistakes, missed words, both missing and surplus spacing between words, punctuation and the inconsistent application of non-text characters such as ~,# and \*. Cocx and Kusters (2006) refer to police MO fields as 'polluted.' They explain that despite protocols for completion, the reports 'suffer' from the inclusion of abundance of police terminology and spelling mistakes. Although spellcheckers can be used to identify errors, it should be noted that at times the recorded information is far removed from the intended spelling, meaning errors will be missed by automated checking. Thus spelling errors can dramatically increase the variety of terms occurring in the data and requires the development of data dictionaries to identify the terms used and convert them into a smaller vocabulary. The preparation of data dictionaries for this thesis are described in Section 7.2.

### 6.2.3 MO fields: Issues that reflect the nature of natural language

The problems faced when analysing natural language, particularly with a large volume of cases, were outlined in Chapter 4. Inspection of the police data revealed repeated examples of all of these problems confirming that MO descriptions can be regarded as natural language.

A central issue identified in Chapter 4 was the need to interpret words in their context. Extracting words from their context may often introduce ambiguity. Bache, Crestani, Canter, & Youngs (2010) provide the example of four words occurring in burglary MO fields: electricity, electric, electronic and electrical – these appear to be synonyms or near synonyms, but the first two were used in relation to the coin-operated electricity meters that were targeted for their contents, the latter two refer to the nature of goods that were stolen and removed from dwellings. Understanding this distinction is important for distinguishing two distinct script pathways, however, this distinction could only be uncovered with a fair degree of familiarity with the free text.

Of key importance to accurately interpreting MO descriptions is the identification of negating phrases. Failing to identify these phrases can reverse the meaning of terms within the data. The examples below contains two such phrases each: *without* being observed and *did not* discover, felt nothing and saw *no one*. Other examples include nothing stolen, no verbal.

BETWEEN MATERIAL TIMES PERSON(S) AT PRESENT UNKNOWN UNZIPPED IPS HAND  
BAG AND REMOVED FROM THE TOP OF THE BAG A SMALL PURSE CONTAINING JUST

CASH (CARDS ETC IN OTHER PURSE) WITHOUT BEING OBSERVED OR FELT IP DID NOT DISCOVER THEFT UNTIL GOING TO PURCHASE CIGARETTES AFTER CLEARING TILLS WHERE SHOPPING WAS PAID FOR ON CARDS SHOP

PROPERTY STOLEN IN ZIPPED UP POCKET OF RUCKSACK CARRIED ON IP' S BACK. PERSON UNKNOWN UNZIPPED POCKET AND STOLE PROPERTY. POCKET ZIPPED BACK\_UP. IP FELT NOTHING AND SAW NO ONE. AIRPORT

In fact, police data represent what has been termed a 'local grammar' a subset of natural language which represents the communication of a specific community with a common purpose (Crowston et al., 2012). The specialised vocabulary utilised within a local grammar can help with disambiguation, for example homonyms will exhibit a more restricted set of possible meanings that are relevant to that community, in this case to crime and policing. While ambiguity is bounded by the local grammar, understanding and analysing this text requires a comprehension of the shared set of meanings and concepts applied within the domain. Further, this comprehension must be programmed into any automated methods designed to assist the human analysis. While it is possible to tailor systems to local grammar it should be borne in mind that too close attention to local terminology could restrict the broader applications of an automated system to other contexts. System comprehension also needs to keep abreast of evolving language, both within local grammar and language more broadly.

#### **6.2.4 MO fields: Issues that reflect the nature of narrative accounts**

The information in MO fields is not restricted to a description of the activities of the offender(s). Reflecting their status as brief narratives, the fields can be used to describe the actions of a range of actors on the scene. This includes actions of victims prior to an attack and victims' acts of resistance during offences. The connections between action and actor are important to fully comprehend the narrative. However, some forms of analysis lose this connection. For the first examples below, an analysis that quantifies the presence of particular words or phrases (see Section 4.3.3) would identify the presence of hand but not whose hand, there is a risk of assuming that all acts of biting are committed by the offender against the victim and the ingredient of victim resistance could be lost from this and other cases.

OFFENDER APPROACHED IP WHO WAS ABOUT TO USE HER MOBILE PHONE TO CALL THE POLICE TO REPORT A FIGHT PUSHED HER AGAINST A WALL AND THREATENED TO ASSAULT IP IF SHE DIDNT HAND OVER HER PHONE THE OFFENDER THEN KNOCKED THE PHONE OUT OF HER HAND AND STOLE SAME

BMT THE IP WHO IS A TAXI DRIVER DROVE HIS VEHICLE A TOYOTA AVENSIS [VRM NO]WAS DIRECTED BY HIS BASE TO ATTEND [ROAD NAME]TO COLLECT A FARE- HE

COLLECTED X3 IC3 FEMALE OFFENDER(S) ONE OF THE FEMALES HAS THEN PLACED HER HAND\_OVER IP' S MOUTH AND POINTED WHAT WAS PROCLAIMED TO BE A FIREARM TO THE IP' S HEAD. THE IP HAS THEN BIT THE OFFENDER' S HAND- AT WHICH POINT ALL THREE OFFENDERS HAVE MADE GOOD THIER ESCAPE. NO PROPERTY STOLEN

Crime narratives are co-generated by actors, at least one of whom is an offender, in interaction and reaction to the actions of those who share the scene.

### *6.3 Summary*

As outlined in Chapter 2, the aim of the data review was to examine MO descriptions and to identify any limitations which could make analysis difficult. The review also aimed to consider whether techniques available from the fields of natural language processing and content analysis can be used to address these limitations in order to facilitate a meaningful analysis of MO patterns. This chapter has reviewed the nature of MO descriptions from a sample of Force A data. As noted above, in the author's experience the Force A MO descriptions are broadly representative of fields recorded across the UK. It is clear that police free text reports of offence MOs contain data that provide a valuable insight into the methods that have been used to execute an offence. However, the information provided in the descriptions may not be complete and the characteristics of the data mean that the task of extracting this intelligence is far from straightforward. Comprehension of MO descriptions may be straightforward to the human reader, particularly when they are familiar with police terminology and local knowledge (but not always, a number of MO fields require multiple reads through to understand what is being described). This section has demonstrated that the challenges stem from a mixture of factors relating to the nature of crimes themselves, the nature of crime recording and the nature of natural language. Some of these issues present a problem mainly for automated or computer-assisted analysis; others present a problem for both human and automated problems. However, the challenges are more significant when using automated methods of text searching and pattern extraction and therefore, for the analysis of large volumes of cases on a frequent basis. For this reason, it is necessary to consider whether the data is suitable for computer-assisted approaches to analysis, and what types of pre-processing might be necessary.

Comparing the data problems identified in the current chapter with the methods described in Chapter 4 it appears that some of the issues relating to the nature of recording and most of those relating to the nature of natural language can be addressed, or at least reduced, through the application of the natural language processing and content analysis techniques. The development of data processing tools that are tailored to MO descriptions has the potential to reduce many of the

barriers to the effective analysis of MO descriptions. The more technical issues relating to spelling errors can be reduced through the modification and application of dictionaries; removing all punctuation removes the problem of inconsistent application. Terminology associated with superfluous information, including audit processes, can be identified and added to stop lists allowing it to be ignored by subsequent analysis. Multi-word lists can be used to ensure that important phrases are identified, where individual words would not have conveyed the true meaning. These lists can also be used to capture negating phrases again ensuring that descriptions are correctly interpreted. Standardisation of the data to reduce the range of vocabularies used includes stemming words of different tenses. Standardisation can be extended to identify synonyms and replace them with standardised terms (known as lemmas), this process can incorporate variations in technical jargon and local vernacular. Although the application of coding dictionaries is automated, their preparation and modification is not. Given the time and effort that is required to inspect data and developed tailored coding dictionaries it will be important for this thesis to explore whether these methods really reduce the burden of analysis or simply shift it from analysis to data processing.

It is apparent that the tools offered by natural language processing and content analysis hold promise for getting the most from the information provided within MO descriptions. However, these tools do not offer assistance in assessing or mitigating threats to the validity of the data. These tools do not help with data that may be missing, incorrect, biased or filtered by organisational processes. What is more challenging is dealing with information that might be missing, or factors which shape the information as it is provided. The review has shown that MO descriptions are shaped by social and institutional as well as technical factors. As Bache et al. explain (2010, p.306) "The mechanism by which the actual events surrounding a crime are transmitted to the crime description is a noisy channel." Although Bache and colleagues identify that crime records are 'authored,' processed and communicated by police employees, they conclude that MO descriptions can be treated as descriptions that were ultimately generated by the behaviour of the offender. This thesis has also adopted this assumption but also aims to place the findings from data analysis in the context of existing literature and theoretical frameworks. Interpreting the analysing findings in this wider context will help to identify gaps in the data, but also helps to assess whether any patterns identified are compatible with current evidence and theory.

While this chapter has identified limitations to MO recording, any recommendations for improvement should be realistic and should consider the current pressure on police resources, the drive to reduce police bureaucracy and the real time pressures on those undertaking the recording task in what may be a noisy, chaotic environment. While improvements to MO recording would be

welcome, this thesis assumes that there will be no radical changes in the near future and will focus predominantly on suggesting ways of making best use of the systems and data as they currently exist.

# Chapter 7 Identifying Crime Scripts Through Computer-aided Content Analysis.

## 7.1 Introduction

This chapter reports the process of preparing a text based database for computer-aided quantitative content analysis. As described above the data comprised police recorded text fields describing the MOs of 22,193 cases of robbery of personal property and theft from the person committed in one police force over a period of three years (January 2007 to December 2009).

The cleaning, standardisation and later analysis of data was conducting using a number of “off-the-shelf” tools and software packages. The data, which were described in detail in Chapter 6, were prepared and analysed using a basic text editor (Notepad++)<sup>30</sup> and various analytical packages within the statistical software R. R is a free software environment for statistical computing and graphics. The R environment can be extended through the addition of packages. Packages are collections of R functions and compiled code. Packages can be developed, shared and evaluated by the R community<sup>31</sup>. The R task view for natural language processing provides a detailed list of the extensive range of validated packages that include commands relevant to conduct natural language processing and text mining<sup>32</sup>. The R packages used in data preparation and analysis were:

- TM (Hornik, Feinerer, Hornik, & Meyer, 2008);
- Snowball (Bouchet-Valet, 2015);
- Aspell (Lang, 2005);
- Wordnet (Feinerer, Hornik, & Wallace, 2015);
- Cluster (Maechler et al., 2015).

Other popular open source software for textual analysis includes the Natural Language Toolkit (NLTK) written in the Python programming language (Bird, Klein, & Loper, 2009) and the Text Mine module in the Perl programming language (Francis & Flynn, 2010). R was selected because of the ability to conduct additional statistical analysis and visualisation from within the same platform.

---

<sup>30</sup> [notepad-plus-plus.org/](http://notepad-plus-plus.org/) (last checked 06/08/2015)

<sup>31</sup> <http://www.cran.r-project.org/> (last checked 06/08/2015)

<sup>32</sup> <https://cran.r-project.org/web/views/NaturalLanguageProcessing.html> (last checked 06/08/2015)

## 7.2 Data Preparation

The starting point for the analysis was a database of 22,193 MO descriptions and related variables (crime type, date, time etc.). As discussed in Chapter 4, free text data require data cleaning and standardisation to convert them to a form suitable for statistical analysis. The R Text Mining (TM) package was used to carry out several data preparation functions including the removal of superfluous white space between words, the conversion of all text into one case (lowercase) and the removal of punctuation. Following this first stage of processing, an initial inspection of the database revealed that across the 22,193 MO descriptions, 22,250 different words were used at least once. The occurrence of more words than cases suggests a high level of heterogeneity across the database, with each of the MO descriptions seeming to be very different from the others. However, this impression is exaggerated by the characteristics of natural language which increase the number of words found in the data through instances of:

- misspelled words (every misspelled variation of a word will appear unique);
- singular and plural forms of the same word;
- the use of different tenses (e.g. walks, walked, and walking will all be counted as different words);
- stopwords;
- words with closely related meanings including (synonyms, hyponyms and hypernyms);
- the use of proper nouns to describe features such as streets, licensed premises and shops (which may also be misspelled or subject to variations).

In addition to reducing the variety of words, prior analysis was also necessary to reduce the ambiguity of terms in the data to increase the potential for words to be interpreted correctly. This included identifying multi-word phrases and negating phrases and disambiguating homonyms.

### 7.2.1 Dictionary development

To accomplish the tasks of reducing flexibility and ambiguity, a set of dictionaries was prepared and applied to the data. The starting point for building these dictionaries was the use of freely available natural language tools including the R packages *Aspell* for spell checking, *Snowball* for stemming and the English stopword list available with the TM Package. The *TerMine* service, provided by NatCTem, was used to generate a list of the most frequently occurring multi-word phrases in the data (Frantzi et al., 2000). The use of the *WordNet* lexical database to identify related words was explored but this

provided difficult to implement in the R environment (other than searching for one synonym at a time). While all of the tools used were revised and tailored to the data after a series of iterations; the identification of synonyms and related words was the task for which the most manual effort was required. To ensure anonymity, personal names and proper nouns relating to streets and other local locations were identified and removed.

As already suggested, the 'off-the-shelf' tools provided a starting point for data standardisation but using them in their raw form did not resolve all of the issues in the data. For example, the spellchecker only identified and corrected a small proportion of errors in the MO descriptions. Spelling errors were identified as a problem in Chapter 6, but the data preparation phase highlighted the true extent of this problem. Not only were there particularly high levels of errors, the incorrect spellings were in many instances far removed from the intended term meaning they were not picked up by spell checkers. Therefore initial spelling errors had to be identified manually and included in the dictionary. Table 7.1 highlights just one example of this issue, showing that there were 36 different spellings of *escape* in the data, in addition, the alternative words and phrases for *escape* and their misspellings provided a further 36 terms. These 74 terms were all recoded to the token *escape*. Similarly, there were 127 variations of *offender* including spelling variations of *offender*, *suspect*, *aggressor* and *assailant* (but not including descriptions where terms such as *youth* or *gang* were used in place of offender). There were also 48 terms used for victim including *IP*, *injured party*, *victim*, *complainant* and spelling variations).

Table 7-1 Table showing the alternative words and incorrect spellings of 'escape' found in the data

Spellings of Escape in the Data	Alternative Words for Escape in the Data
andescaped	absconded
eacape	egress
eacpe	egressed
eascpe	exited
ecape	exitted
ecapes	exitting
ecsape	fled
eescape	flee
erscape	fleeing
esacape	flees
esacepe	getting_away
esacpe	get_away
esape	goddescape
escae	goodescape
escaoe	madeoff
escaope	made_off
escap	make_off
escapebus	makinggood
escaped	makingoff
escapeon	making_off
escaper	ran_off
escaperoad	running_away
escapes	running_off
escaping	run_away
escaple	run_off
escapoe	egress
escapre	egressed
esccape	exited
escpae	exitted
escpe	exitting
escpwe	fled
escqape	flee
esdcapae	fleeing
rscapae	flees
secapae	getting_away
theirnescape	get_away

The TerMine service generated a list of suggested multi-words phrases, each of which would be recoded as one token, but these needed checking by hand and refining by adding and removing terms. Phrases that were added included: *under\_the\_influence*; *made\_good\_their\_escape*;

*stood\_in\_her\_way* and *racist\_insults*. It was particularly important to identify multi-word phrases containing negating words which reverse the meaning of a term or phrase. Phrases where the negative term was directly collocated with the connected phrase were effectively picked up by TerMine, i.e. phrases such as *no\_injury* and *without\_warning*. Other instances were more complex and needed to be identified manually e.g. *without\_any\_warning; not\_put\_in\_any\_fear\_of\_violence; no\_threats\_or\_weapons* (in the latter examples *no\_threats* would be picked up automatically but *no\_weapons* would not). These multi-word phrases were recoded as a token which was treated as one word in which the negative element was clear, some examples of recoding phrases relating to no injury and without provocation are provided in Table 7.2. Negating terms may not always be as obvious as *no*, *didn't*, *did not* etc. For example in the MO description, below, the inclusion of the words *slow* and *speed* completely change the meaning of the word *speed*:

<<BETWEEN TIMES STATED IP WALKING IN A RESIDENTIAL AREA WHEN A VEHICLE STARTED FOLLOWING HIM AT A SLOW SPEED. THE VEHICLE PULLED OVER AND OFFENDER GOT OUT AND APPROACHED IP AND SAID WHAT HAVE YOU GOT FOR ME"". HE PUSHED IP TO GROUND BEFORE TAKING HIS WALLET FROM HIS JEANS POCKET AND THEN MADE OFF IN VEHICLE DRIVEN BY OFFENDER 2 IN THE DIRECTION OF XXXX ROAD ROAD CAR>>

Table 7-2 Examples of multi-word phrases (including negating phrases) from the data

Raw Text	Changed to Token
No injury No visible injuries Not hurt No visible injury Not injured Did not suffer any injuries	no_injury
Without provocation Without any provocation No provocation Un provoked	no_provocation

As noted in Chapter 4, the meaning of many words can only be understood in relation to their context. The methods used to analyse data in this thesis largely examines words in isolation from their context and can therefore introduce ambiguity. This is a particular problem with homographs, distinct words that share the same spelling but have different meanings ((e.g. saw (verb) and saw (noun)). To reduce the uncertainty of the meaning of terms, potentially ambiguous terms had to be identified and their meaning clarified. Again existing lists of known homographs in the English

language can provide a starting point, but much of this work had to be conducted manually by inspecting the database. KWIC lists, also known as a concordance view, aided the process of identifying the meaning of ambiguous words. These lists display every occurrence of a word surrounded by their context. This facilitates the inspection of the different meanings that words may have within the data. Figure 7-1 below shows a KWIC list for the word '*escort*.' The list highlights the two different meanings for the similarly spelled words *escort* and *escorted*. If stemming had occurred prior to disambiguation all instances of *escorted* may have been recoded to *escort* and then *car*, a misinterpretation of the original word. Disambiguation can be made more difficult by the use of slang and local variations, for example the names of pubs or local landmarks taken out of contexts could lead to misinterpretation (imagine the impact that many crimes occurring at a Punch and Judy Inn would have on the perceived level of violence in the data if the issues was not recognised and corrected). In the current data, the location code (park, licensed premise, shop etc.) was consistently recorded allowing a quick check for the names of licensed premises and the removal of related proper nouns.

Some advantages in data preparation stemmed from the data's status as local grammar. This restricted the available meanings of many words that otherwise could have been ambiguous. Thus *stalk* always referred to the verb, not the noun, while both *implement* and *object* always referred to the noun. However, *phone* could relate to the noun as in mobile phone, but often included the verb as in *phoned the police*, this term also appeared as *phone box*. This issues was resolved through the identification of multi-word phrases relating to calling the police or calling 999 and also the identification of words and phrases relating to phone boxes.

RACE NOT USED OFFENDERS MADE OFF IN R REGISTERED FORD ESCORT WITH THIRD OFFENDER AN IC3 ROAD BETWEEN MATERIAL T  
 Y STOLEN. OFFENDERS MADE GOOD ESCAPE IN AN M REG FORD ESCORT GHIA DWELLING KNOWN OFFENDERS AWAITED OUTSIDE IPS  
 FFENDERS THEN MADE OFF TOWARDS ##### RD IN BLACK FORD ESCORT OLD STYLE VEHICLE VRM UNKNOWN. PUBLIC FOOTPATH DES  
 MANAGED TO ESCAPE.OFFENDERS WERE IN A LIGHT BLUE FORD ESCORT. ROAD BMT UNKNOWN OFFENDERS HAVE DRIVEN PAST THE I  
 HAS BEEN APPROACHED BY 3 U K OFFENDERS IN A BLUE FORD ESCORT MOTOR VEHICLE. OFFENDERS BELIEVED IP TO BE A PROST  
 ON LAMP POST OFFENDER MADE GOOD ESCAPE IN WHTIE FORD ESCORT NO VRM IN U K DIRECTIOPN NO PROPERTY STOLEN ROAD I  
 HEN GOT INTO THE FRONT PASSENGER SEAT OF A BLACK FORD ESCORT AND MADE OFF FROM THE SCENE. NO CCTV AND NO KNOWN  
 KET. OFFENDERS HAVE MADE GOOD ESCAPE IN A SILVER FORD ESCORT. NO VRM TAKEN. MADE OFF ON ##### WAY TOWARDS #####  
 G WITH SECOND OFFENDERS VEHICLE POSSIBLY A LIGHT FORD ESCORT NO VRM. OFFENDERS 2 WHITE MALES HAVE APPROACHED IF  
 VEN BY THE SECOND MALE THE VEHICLE INVOLVED BLUE FORD ESCORT ROAD REC BETWEEN MATERIAL TIMES ##### HAS ATTENDE  
 EAR. 2 UNKNOWN OFFENDERS HAVE GOT OUT OF A WHITE FORD ESCORT AND TOLD IP TO HAND\_OVER HIS MONEY THEN PUNCHED IF  
 CHANGING HIS TYRE ON HIS VEHICLE NAMED A WHITE FORD ESCORT VAN HAS GOT INTO THE DRIVERS SEAT AND STARTED TI C  
 E CROSS HE WAS TALKING ON HIS MOBILE PHONE WHEN A RED ESCORT NO TRACE STOPPED TWO BLACK MALES ALIGHTED AND DEMA  
 IS CAR ON PARK CAR PARK.OFFENDERS HAVE ALIGHTED A RED ESCORT AND SURROUNDED VEHICLE.ONE OFFENDER HAS SAID ' LET  
 N SAFETY OFFENDERS HAVE MADE GOOD THIER ESCAPE IN THE ESCORT and WITH THE IPS CAR X1 IC1 MALE AND X1 IC3 MALE C  
 D PARTY' S BAG AND MADE HIS ESCAPE INTO A DAIRY WHITE ESCORT TYPE VAN VEHICLE TOWARDS ##### ROAD. ROAD BETWEEN  
 IS POCKETS OFFENDERS PRODUCED A HANDGUN AND KNIFE AND ESCORTED IP TO OTHER LOCATIONS DURING WHICH OFFENDERS MAC  
 OFFENDER HAS ENGAGED IP IN CONVERSATION ABOUT RTC AND ESCORTED HER TO HER CAR WHERE HE HAS GRABBED AND SQUEEZED  
 D MONEY AND PHONES. OFFENDERS HAVE TAKEN PROPERTY AND ESCORTED IPS FROM BUS THREATENING TO ' STAB THEM' IF THEY  
 ED MONEY AND PHONES OFFENDERS HAVE TAKEN PROPERTY AND ESCORTED IPS FROM BUS ROUTE THREATENING TO STAB THEM IF T  
 HAS ENTERED IP' S HOUSE PUT A HANDGUN TO HER HEAD AND ESCORTED HER ROUND THE HOUSE DEMANDING PROPERTY. WHILST C  
 AILED ON THE WC202 HAS HELD THE IP AT KNIFE POINT AND ESCORTED HER UNDER DURESS TO THE CASHPOINT ONCE THERE HE  
 NE ADULT OFFENDER HAS APPROACHED 2 JUVINILE MALES AND ESCORTED THEM TOWARDS PARKLAND BEFORE PRODUCING LOCK KNIF  
 WEEN MATERIAL TIMES ##### HAS ATTENDED IP' S FLAT AND ESCORTED HIM TO POST OFFICE WHERE HE HAS WAITED FOR IP TC  
 D OFFENDER AND THEN OFFENDER TOOK HOLD OF IPS ARM AND ESCORTED HIM TO PARK. WHILST IN THE PARK OFFENDER TOOK IF  
 ES AND ARE DETAINED BY STAFF DUE TO THIS. AS THEY ARE ESCORTED TO THE STAFF ROOM CCTV SHOWS ONE OFFENDER DROP A  
 Y' AND PLACED IT IN HIS RIGHT COAT POCKET HAVING BEEN ESCORTED BY ' KELLY' TO THE POST OFFICE HOLDING HIS RIGHT  
 ND SEARCHED IP FINDING IPS BANK CARD.IP HAS THEN BEEN ESCORTED TO THE ##### CASHPOINT IN ##### WHERE IP HAS WI  
 K BY OFFENDERS. OFFENDER 1 HAS CAUGHT IP AND FORCIBLY ESCORTED IP TO OFFENCE LOCATION. OFFENDERS 1 AND 2 HAVE T  
 DIDN' T HAND\_OVER HIS PHONE IP DID THIS OFFENDER HAS ESCORTED THE IP TO ##### ROAD WHERE OFFENDER HAS THEN MAC  
 THE PHONE AND TOLD IP TO GET OFF THE BUS OFFENDER HAS ESCORTED IP OFF BUS. OFFENDER REMAINED ON BUS MAKING GOOD  
 OPED TO EREND FROM WALKING WHILST OTHER 2 OFFENDERS ESCORTED TO FURTHER ALONG ROAD. A GOLD COLOURED KNUCKLER

Figure 7-1 A screenshot of a KWIC list showing all instances of *escort* in the data

One issue which arose in the standardisation of the text was the stemming of plural words, when the distinction between singular and plural words was important for the analysis. This issue was relevant when MO descriptions included either the number of offenders or the number of other actors in the situation. This problem is heightened by the fact that the standardisation process removes punctuation, so that 'offender's' would become 'offenders' and then 'offender'. Further complications are added by the fact that this punctuation was often missing or misplaced in the raw data. Even if it had been possible to accurately retain the singular or plural in these cases, the number of offenders, or other actors was often indicated by shorthand such as X2. Looking at the examples in Figure 7.1 above once more, we can see details of the number of offenders given in various ways including:

- 3 U K OFFENDERS
- THIRD OFFENDER AN IC3
- OFFENDERS 2 WHITE MALES
- 2 UNKNOWN OFFENDERS
- X1 IC1 MALE AND X1 IC3 MALE
- OFFENDERS 1 AND 2

Similar problems emerge when attempting to uncover the number of victims in a description, or the number of people that were in the company of a victim at the time of the offence. It was decided that the data processing would not attempt to ascertain the number of victims / offenders in each description. This represents a limitation within the analysis as co-offenders are a key resource in scripts (the token group later provided significant in the analysis) and other actors have the potential to be crime preventers or crime promoters. This problem could have been addressed with the use of additional variables relating to the numbers offenders and victims and their characteristics, but this data was not part of the original request made to Force A. Providing a way of ascertaining the number of actors in each scene would certainly provide an improvement to any future research using similar techniques.

### 7.2.2 Coding iterations

The development of data dictionaries was an iterative process, beginning with the standard tools and then refining the coding frames based on inspection of the data. Initial analysis of the data also began to reveal anomalies that remained in the data. In response to this, the coding frames applied to the data were revised i.e. stoplists and multi-word lists were extended and code lists revised. For example, in an early run of the analysis, the term 'Belcher' emerged as a significant term, dominating one of the clusters. This term was not known to me or my supervisors. On further investigation it became clear that a Belcher is a form of gold chain worn as jewellery and the word was recoded accordingly. Other refinements included the revision of the stoplist to add words and phrases that related to administrative processes rather than the commission of crime, examples of these terms include *between material times*, *at offence location*, *re-classified*, *xref* and *cross reference*. A further issue was the frequency with which the names of particular police officers and staff members occurred in the data; particularly those individuals involved with auditing the data. These proper nouns were identified and added to the stopword list.

One of the most challenging and time consuming (because it was the least automated) element of data preparation was the extension of tokenisation to identify and group together synonyms, hypernyms and hyponyms and other words and phrases that indicate the same or similar class of object or action.

The identification of words with similar meaning is not an objective process. The inspection of data and coding frames will be guided by the understanding of the analysis task, and potentially by previous experience with the data or the real world problems that the data represents.

In this case the analyst (myself) had previous experience in analysing crime data, including some basic experience of inspecting crime MOs and a background knowledge of theory and research evidence relating to theft from the person. It is important to consider the extent to which this prior knowledge may have informed and facilitated the processes of data preparation and analysis, or may have biased them. If the data is prepared with prior knowledge there may be a risk of missing crimes that do not fit expectations. This is avoided to a degree by conducting exploratory analysis to data that has been minimally cleaned, and ensuring that anomalies in the early iterations are not ignored. It is, therefore, important to acknowledge that this process is by no means totally automated and free of subjectivity.

In this research, insights gained from the literature on theft from the person and robbery of personal property were used to consider key elements of offence commission that might be described in the MO and then explore the vocabulary being used to describe those elements. Data preparation was being conducted in parallel with the development of conceptual frameworks, therefore the emerging script frameworks provided a structure for considering elements of the description. This prompted the consideration of different scenes, such as target selection, approach, engagement and transfer, what is known about the accomplishments the offender must complete within each scene and then potential terminology that could be used to describe this. For example, the literature demonstrated that approach is a critical element of robbery of personal property and that one of the different routes for making an approach is to create an element of surprise. This prompted a search of the terms in the data that might describe how surprise might be generated e.g. jumped out, hiding etc. Similar exercises were conducted for other script scenes.

Different research questions require greater or lesser attention to the levels of lexical detail in the data. So, for example, in this analysis the coding frame did not make a distinction between severities of injury. Tokens relating to injury such as injury, head\_injury, swelling and bruises were all coded simply as 'injury'. For the purposes of a different research question, these distinctions may be important and would need to be preserved. However, where this type of detail is important, the ability to accurately code the data to the desired level of detail is dependent on the amount of information that is available in each of the MO descriptions; and many of the descriptions in the dataset simply stated 'injury'.

In the current analysis distinctions were made between verbal interactions, verbal aggression and verbal abuse. It was anticipated that these distinctions would be helpful in understanding the engagement scene of offence commission. Verbal interactions were characterised by conversations between actors in a script (normally the offender and the victim), where these verbal interactions

were elevated to the level of shouting, or use of aggressive profanities this was labelled as verbal aggression and where it was clear that the language was directly abusive it was recoded as verbal abuse. As noted above, it was not always possible to accurately code the nature of interaction because descriptions may only have included the word 'verbal.' Consequently, coded data contains imperfections and limitations. If this study was interested in hate crime it would have been necessary to ensure that, wherever the data permitted, distinctions could be made between different types of abuse such as homophobic, racial, disablist.

The cycle of refinement and re-analysis can be repeated many times, although the marginal improvements to the results become smaller with each iteration. Overall, the coding frames underwent twelve iterations of refinements. The results of the analysis with the transformed corpus were generally consistent with those of the early iterations. The initial round of transformations is an essential stage of analysis (cluster analysis on completely raw text did not produce any meaningful groupings); however, smaller refinements, while they may have benefits for the presentation and communication of results, did not fundamentally influence the results of the analysis. In other words, there is a point where refinements prevent something strange from occurring on a presentational wordcloud diagram but the changes do not change the overall interpretation of the cluster. There is, therefore, a careful balance to strike between producing a perfectly clean and unambiguous dataset and the time and effort that is invested to achieve this. However, where time has been invested in the creation of revised dictionaries, these can be used as an improved starting point for future analysis.

Although standardisation means that, for the purpose of analysis, words in the raw text are replaced with a narrower range of tokens, a permanent link is retained between the tokenised data and the raw data. This means that the raw form of the MO descriptions can be recalled at any time in order to assist with the verification and interpretation of the results.

A final stage of data preparation was to remove 'sparse terms'. This excluded from the analysis any tokens that occurred in the whole dataset with a frequency less than six. This helps to reduce the processing time involved in analysing the data and to simplify output tables and graphs while losing little in terms of additional information.

### 7.2.3 Summary

This chapter has described the process of data pre-processing and analysis which prepared the data for subsequent analysis. Chapter 6 described characteristics of MO fields that present challenges for their analysis. The above discussion has shown how techniques available within content analysis and

natural language processing can help to resolve these issues. However some problems remain, including the identification of the number of actors in a scene.

The dictionary refinements reduced the number of distinct tokens from 22,250 to just 1,920. This striking simplification of the data helps to identify MOs that are similar but have been described using different terminology. The process of data preparation and dictionary development is clearly time consuming which may seem a contradiction to the aims of using methods appropriate to volume crime data. However, effort input into this stage produced the legacy of a set of dictionaries – and a developed procedure – which can provide a starting point and crucial time-saver for any future analysis.

## Chapter 8 Identifying Crime Scripts and Tracks

This chapter reports the results of a computer-aided quantitative content analysis of police recorded text fields prepared as described in Chapter 7. The analysis is based on the assumption that offences, committed in the same way, will be described using similar vocabulary. Analysis of the terms used in the descriptions was used firstly to test whether it is possible to distinguish between different types of offence and secondly, to identify groups of crimes that were committed in a similar way. In this regard, MO descriptions were treated as (imperfect) traces of events in the real world and the analysis aimed to identify regularities in the manner in which offences were carried out.

This chapter describes the results of the application of cluster analysis to MO descriptions, and applies the analytical framework developed from integrating script and CCO frameworks in Chapter Five to organise, interpret and compare the clusters.

### *8.1 Distinguishing Between Protoscripts*

A preliminary analysis was undertaken to determine whether computer-aided content analysis could reliably distinguish between the two key protoscripts in the data, namely, theft from the person and robbery. To conduct this analysis, the database was split into two groups, based on the Home Office crime classification robbery or theft from the person. A contingency table was created which cross-tabulated crime type with each of the tokens in the database. Further analysis of this table (using the chi-square statistic) enables the identification of tokens that are distinctive to each group of cases. Distinctive tokens are significantly overused in the descriptions for a given crime type compared to the usage that would be expected if all tokens were evenly distributed across the two groups. These tokens can, therefore, be regarded as distinctive. The distinctive tokens for robbery and theft from the person are presented in Table 8.1. The table includes only tokens which were statistically significantly overused for each crime type. Tokens that are distinctive of a group of cases are not necessarily exclusive to those groups, although those tokens which were found exclusively for one crime type are marked with an asterisk.

The table shows that the computer-assisted content analysis and chi-square analysis has been successful in distinguishing between the two crime types. The distinctive tokens serve to highlight characteristic features of the method of crime commission for each crime type. The distinctive tokens do not, on their own, provide details of MO types. In isolation, they require background knowledge (common sense or expert) in order to interpret their role in the overall MO narrative.

Script and CCO elements can be used to help contextualise and interpret the meaning of distinctive tokens in relation to the overall MO story.

It should be noted that some of the tokens require context in order to understand their role in the offence, although roles in some cases can be assumed based on previous research evidence. Cases can be investigated to provide more information on the context of word usage. This will be expanded on in the inspection of specific theft clusters in Section 8.2.

**Table 8-1 Distinguishing between robbery and theft from the person. distinctive and exclusive tokens.**

<b>Distinctive and exclusive* tokens</b>	
<b>Robbery</b>	<b>Theft from the person</b>
threat	shop
fear	entertainment
demand	market
verbal	supermarket
assault	shopping_centre
push	crowd
force	pay
injury	offender_not_seen
face*	nothing_felt
brandish*	missing
knife	no_force
weapon	distract
vehicle	dip
door	undo
knocked_door*	bag
knock_on*	
search	zip
ground	
tie_up*	remove
unconscious*	trolley
hand_over	contain
male	
friend	
jewellery	
twist*	
engine*	
bang*	
discharge*	

Table 8.1 highlights that the key differences between robbery and theft from the person, in terms of the information available within the MO descriptions, relate to the nature of expropriation (the transfer act). Other differences relate to the offence's wider environment and the enclosures that are breached by the offenders. The distinctive tokens do not help to provide any information, or

highlight any differences in terms of the motivations, or provocations that caused the offences or the processes of searching for, and identifying a target, a suitable target location and a victim. Perhaps, with the exception of the token *pay*, which, depending on the context, may be indicative of activity conducted by victims that may make them more vulnerable (this is investigated further in Section 8.2). These omissions relate to the limitations of the MO descriptions, discussed in Chapter 6, that they only provide a snapshot of a restricted part of the full script of offence commission.

The table highlights key differences in the wider environment in which the search for targets and/or the transfer of property takes place. The locations that are distinctive to theft from the person scripts are busy places, populated by other people indicated by tokens including, *entertainment*, *supermarket* and by the token *crowd*. Nothing in the table provides information on the wider environment of robbery offences.

Key differences between the two crime types are highlighted by words that relate to the transfer/expropriation of property. A key distinction is in the nature of approach and engagement between robbery and theft from the person. Theft from the person is characterised by stealth and secrecy and an absence of engagement between victim and offender tokens such as *offender\_not\_seen* and *nothing\_felt*. Target enclosure tokens, in the theft from the person group, relate to *bags and trolleys* with the offenders breaching them through *undoing, dipping and removing*.

In contrast, the table suggests that robbery is characterised by verbal and physical contact between offender and victim (*threat, demand, verbal, assault, force*). Rather than relying on stealth, robberies use fear and force to achieve the transfer of property; weapons are distinctive to robbery and victims are forced to comply with the transfer rather than it occurring without their knowledge (*search, ground, tie\_up\*, unconscious\**). Only the robbery offences group contained a distinctive token relating to the property stolen i.e. jewellery, other items of property are stolen in robberies but are likely in common with theft from the person, whereas theft of jewellery, other than via robbery, was rare. Neither of the columns provides information on distinctive features of escape or post-offence activity.

## *8.2 Identifying Patterns in Methods of Offending: Distinguishing Between Scripts*

Following the analysis at the protoscript level, a more detailed analysis was conducted of patterns in methods of offence commission at the script level. A cluster analysis was conducted to identify regularities in the way in which offences are carried out and to assign cases to clusters based on these regularities. This analysis was conducted using the R “Cluster” (Maechler et al, 2015) package. Commercial software is available to perform the steps for text preparation and analysis described below, see for example Lexmancer<sup>33</sup> and T-lab (Lancia, 2012)<sup>34</sup>. However, open source programmes, which have the advantage of being free of cost, were favoured. Open source software had the added advantage of allowing data and results to be imported into and exported from alternative software if required, rather than being restricted by proprietary file formats.

This section discusses the results of the cluster analysis, inspects each of the clusters in turn and makes comparisons between the clusters. The clusters identified relate to distinct scripts for performing the protoscripts of theft from the person and robbery, although as will be explored below, there was not always a one-to-one matching of clusters to scripts. To use Borrion’s (2013) terminology, these scripts are ‘performed scripts’ derived from empirical data as they describe sequences of actions that have actually occurred. Interpreting the clusters through the script/CCO framework enables the systematic consideration and comparison of scripts and the exploration of potential causal mechanisms that have shaped the perpetrations of scripts performed by offenders.

As detailed in Chapter 4, the method used to identify clusters, uses an N-gram method which operates based on frequently used tokens. Cases using similar combinations of tokens are grouped together. N-gram analysis does not completely ignore the context of word use, as the same word used with different combinations of other tokens (and therefore, potentially different meanings/interpretations) will be grouped into different clusters. However, interpretation of the role of individual tokens in clusters needs to be exercised with caution, and sometimes establishing meaning requires reinvestigation of the data. This is particularly the case with assigning actions to individuals, such as determining which actor was walking, talking or running etc.

---

<sup>33</sup> <http://info.leximancer.com/> (Last checked 10/08/2015)

<sup>34</sup> User manuals for these software packages were extremely informative on exploring potential approaches to the analysis of data.

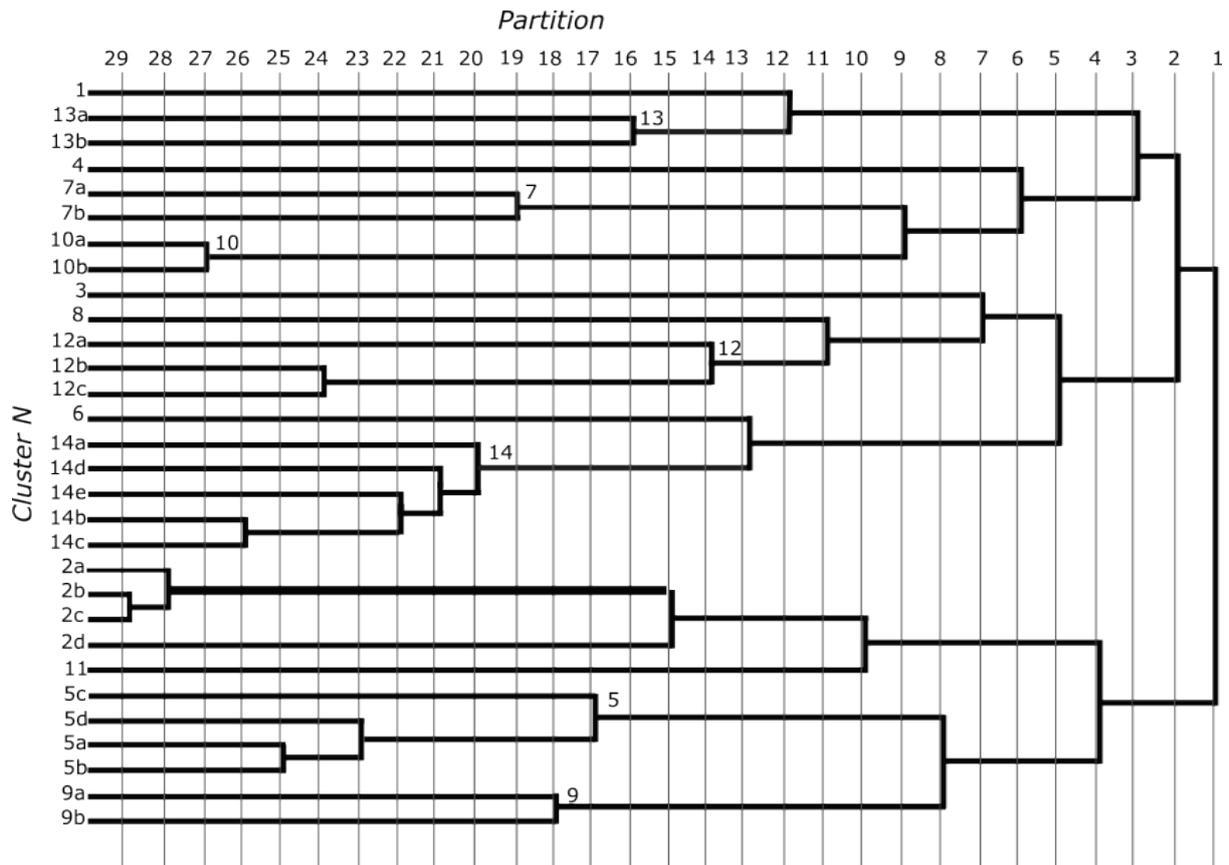
An overview of all of the clusters is provided below, followed by a detailed description of each cluster.

### 8.2.1 The clusters described

The BKM method of cluster analysis requires that the analyst specify the desired number of clusters prior to running the procedure. To compare the effect of requesting different numbers of clusters a range of solutions was produced and their results compared. Iterations were produced ranging from 2 to 30 clusters. Table 8.2 shows the between-cluster variance/ total variance ratio for each of these cluster solutions, and the iterative improvement between each one. The table suggests that a 14 cluster solution offered the best ratio of between-cluster variance and total variance. This coefficient rises steadily as the number of clusters increases (although there are a number of stages that result in a temporary reduction in this metric, for example partitions 8, 11, and 18). With the aim of selecting a parsimonious solution a 14 clusters solution was initially explored. However, to consider the implications on analysis and interpretation, the sub-clusters produced by a 30 cluster solution were also inspected.

**Table 8-2 Comparing the between-cluster variance to total variance ratio of cluster solutions**

<b>No. Clusters</b>	<b>Ratio between-cluster variance/ total variance</b>	<b>Improvement between solutions</b>	<b>No. Clusters</b>	<b>Ratio between-cluster variance/ total variance</b>	<b>Improvement between solutions</b>
<b>7</b>	0.192	0.038	<b>19</b>	0.418	0.018
<b>8</b>	0.173	-0.019	<b>20</b>	0.438	0.02
<b>9</b>	0.219	0.046	<b>21</b>	0.45	0.012
<b>10</b>	0.252	0.033	<b>22</b>	0.471	0.021
<b>11</b>	0.249	-0.003	<b>23</b>	0.491	0.02
<b>12</b>	0.27	0.021	<b>24</b>	0.498	0.007
<b>13</b>	0.293	0.023	<b>25</b>	0.51	0.012
<b>14</b>	0.334	0.041	<b>26</b>	0.53	0.02
<b>15</b>	0.361	0.027	<b>27</b>	0.541	0.011
<b>16</b>	0.386	0.025	<b>28</b>	0.55	0.009
<b>17</b>	0.413	0.027	<b>29</b>	0.545	-0.005
<b>18</b>	0.4	-0.013	<b>30</b>	0.562	0.017



**Figure 8-1 Dendrogram showing the hierarchical structure of clusters**

Figure 8.1 presents a dendrogram of the clusters identified by the analysis. In this diagram, each numbered vertical line represents a partition at which an existing cluster is split into two. It is, therefore, possible to ‘roll’ partitions backwards and forwards to consider the different structures that would be produced for different numbers of partitions. This is an advantage of hierarchical methods over pure k-means clustering. As noted above, the first iteration of the analysis stopped at 14 clusters. A second iteration drilled down further to produce a 30 cluster solution. The clusters identified in the second iteration have been labelled as sub-clusters of their parent cluster. Thus the sub-clusters that are produced by the further division of cluster 14 become 14a, 14b, 14c, 14d and 14e and so on. Overall, 20,608 cases were successfully assigned to one of the 30 clusters, leaving 7% of cases unclassified<sup>35</sup>. Guided by the hierarchical structure that is a characteristic of BKM cluster analysis, the 30 clusters have been presented as 14 ‘parent’ clusters with a variable number of sub-clusters. The clusters and sub-clusters that were identified are listed in Table 8-3 below. The current

<sup>35</sup> Inspection of unclassified cases did not reveal any specific patterns within the group. Many of these MO descriptions were very brief or incomplete and the group included many MOs in which administrative and audit relevant information dominated the crime description.

section will describe the clusters produced by the analysis, Chapter 9 will consider the integrity of the cluster solution as a whole.

**Table 8-3 Clusters identified from theft from the person and robbery of personal property MO descriptions**

	<b>Cluster</b>	<b>No. Cases</b>
1	Pickpocketing at Venues	692
2	Bag snatches	2392
2a	<i>including resistance and injury</i>	110
2b	<i>from pushchairs and including verbal aggression</i>	144
2c	<i>offender on cycle</i>	1583
2d	<i>including following victim home</i>	555
3	Jewellery Snatches	682
4	Theft of Cycles	757
5	Theft from Bags (Shops)	2185
5a	<i>theft from Bags (Shops)</i>	1388
5b	<i>theft from Bags (with distract)</i>	285
5c	<i>including banks, cards and withdrawals</i>	424
5d	<i>thefts from shop staff</i>	88
6	Robberies within Dwellings	1355
7	Demands and Threats	1472
7a	<i>not involving knives</i>	1010
7b	<i>involving knives</i>	462
8	Involving taxis	313
9	Vehicle related crimes	2318
9a	<i>Carjacking and thefts from cars</i>	1645
9b	<i>Car thefts from dwellings</i>	673
10	Robbery involving young people	1262
10a	<i>in parks</i>	378
10b	<i>on footpaths, alleyways and in schools</i>	884
11	Theft from bags at Licensed Premises	689
12	Robbery with Assault	2842
12a	<i>with assault and injury</i>	837
12b	<i>with assault no injury</i>	204
12c	<i>with assault domestic</i>	1801
13	Mobile phone theft	1355
13a	<i>with communication</i>	989
13b	<i>without communication</i>	366
14	Thefts on buses and at bus stops	2294
14a	<i>at the bus station</i>	133
14b	<i>at the bus stop</i>	445
14c	<i>on the bus</i>	1578
14d	<i>leaving the bus</i>	22
14e	<i>misc</i>	116
	Total	20608
	<i>Not assigned to a cluster</i>	1585
	<b>Grand total</b>	<b>22193</b>

Table 8-3 and the descriptions of each cluster below highlight that the analysis did not produce a one-to-one mapping of clusters to scripts or tracks. Some of the clusters can be interpreted as a script track in a straightforward manner. Other clusters contain a small number of tracks relating to the same script or reveal different permutations through which the same track is performed. In some cases the clusters contain different tracks that are not related to each other and have grouped together cases that could not, through logic or experience be considered part of the same script. Thus the identification of different script tracks stems partly from the clusters, but also from additional inspection and interpretation of the data. The relationship between the clusters, sub-clusters, scripts and tracks are depicted in Figure 8.31 at the end of this chapter.

The following sections will describe each of the 14 clusters and their sub-clusters. Script and CCO frameworks were used to organise the description of the distinctive features of each cluster. Script frameworks helped to organise the information, available from the analysis relating to the distinct script tracks available to accomplish different theft from the person scripts and the combinations of script procedures that are required to complete these tracks. The CCO framework helped to identify causal factors, predominantly those in the situation, that potentially shape offenders' actions and the procedures they elected to follow.

Each cluster description will be accompanied by a series of figures, graphs and tables. This includes a graphic depiction of the tokens that are characteristic of each script. This is represented in a wordcloud. The wordclouds include only those tokens that are characteristic of a given cluster i.e. those that occur significantly<sup>36</sup> more than would be expected if all clusters shared a similar proportion of all MO features. The greater the significance of a token for the cluster, the larger the font size in the diagram – the font size is weighted by the chi-square score and not by the raw frequency with which token occurs in the cluster. Each cluster description also includes a table highlighting examples of the cases that best represent the cluster. These cases were selected on the basis of the cluster membership score (see section 4.4.1). The cluster membership scores for each case are proportional across the 14 parent clusters. Therefore a hypothetical case that had equal probability of being assigned to any cluster would have scores of around 7% for each cluster. A case membership score of 100% indicates total confidence that a case is member of a given cluster (and all other cluster membership scores would be zero). Finally, a tree diagram has been produced for

---

<sup>36</sup> Chi Square > 10.83 (p<.001 df=1) see Methodology Section 4.4.1.

each cluster showing the relevant script tracks and CCO elements most indicative to that cluster (see Appendix 2).

Throughout the discussion of clusters, example MO descriptions are faithfully reproduced from the police data, the only changes to raw text being the removal of any identifying information, the conversion of multi-word phrases to one token and the highlighting of significantly distinctive tokens in red.

*Cluster 1: Pickpocketing at venues*

**Cluster 1 Mapped to one theft from the person track: Pickpocketing at Venues**



Figure 8-2 Significant tokens in cluster 1 (font size relative to chi-square score).

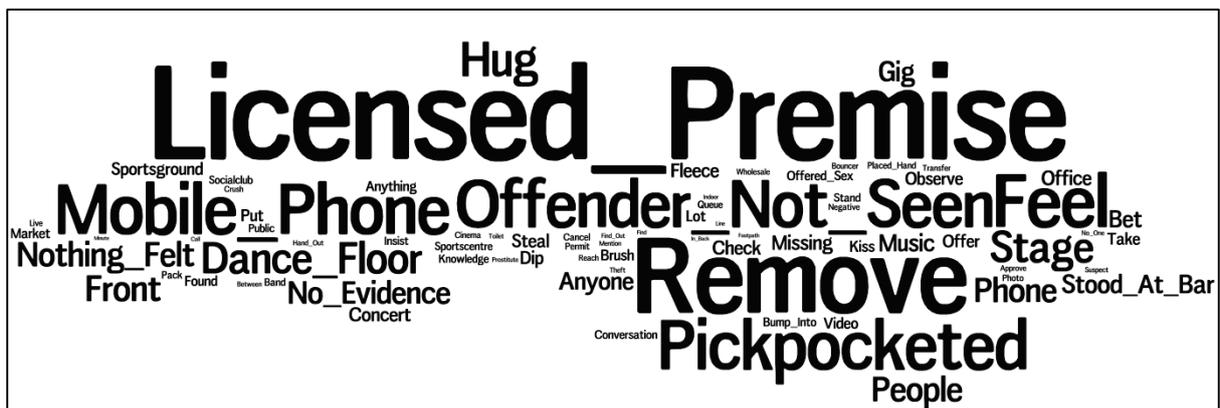


Figure 8-3 Significant tokens in cluster 1 – top 5 tokens removed (font size relative to chi-square score).

**Table 8-4 Examples of MO descriptions assigned to cluster 1: pickpocketing at venues. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

Cluster Membership Score			
	A) High	B) Average	C) Low
i	<p>at_relevant_times &amp; location unknown offenders have <b>stolen</b> a <b>mobile_phone</b> from the ips [victim's] <b>pocket</b>. offenders made good escape.</p> <p><b>Entertainment</b></p> <p>CMS = 75%</p>	<p>between_material_times up [sp]was in <b>busy crowd</b> at <b>music_venue mobile_phone</b> was stolen from his jeans <b>pocket</b> by unknown offender.<b>nightclub</b></p> <p>CMS = 30%</p>	<p>the ip had_been <b>standing</b> outside [name of] <b>bookmakers</b> having a <b>cigarette</b> whilst the offender has approached the ip from_behind <b>reached</b> into the ips left <b>jogging bottom</b> and <b>stole</b> cash <b>from_his_pocket</b> estimated value of #200 offender made good his escape towards the [name of] hotel <b>no_force</b> was used and no_words were used road</p> <p>CMS = 7%</p>
ii	<p>bmt at offence_location ip was in main floor of a large <b>concert</b> when ip had <b>mobile_telephone stolen</b> from <b>jeans pocket</b> by uk [unknown] offender. <b>offender_not_seen</b> and made good escape. <b>entertainment</b></p> <p>CMS = 72%</p>	<p>between material_times and at offence_location uk offender has knocked into ip and removed ips <b>mobile_phone</b> from her <b>trouserpocket</b>. offender has mge in uk direction. <b>nightclub</b></p> <p>CMS = 30%</p>	<p>"taxi: between stated_times two unknown offenders have approached a black_cab taxi in [street name] offenders asked ip how much into town?". offenders then lent <b>forward</b> and <b>picked</b> ip taxi <b>badge</b> off ip's <b>front sweater</b> "road</p> <p>CMS = 7%</p>
ii	<p>ip walking out_of <b>concert</b> when persons unknown <b>took mobile_phone</b> from out_of <b>pocket</b> of <b>jacket</b> ip was <b>wearing</b>. <b>entertainment</b></p> <p>CMS = 70%</p>	<p><b>bmt</b> uk offender/s have <b>removed</b> ips <b>mobile_phone</b> from his <b>back_pocket</b>. ip has <b>felt touch/pressure</b> on his <b>rear pocket</b> where his <b>phone</b> was <b>being</b> kept. within 2 <b>minutes</b> he has <b>checked</b> his <b>pockets</b> having left the <b>dance_floor</b> and <b>found</b> that his <b>phone</b> was <b>missing</b>. <b>phone</b> was not handed in <b>at</b> offence_location. uk offender/s have made_off in uk direction by uk means. <b>nightclub</b></p> <p>CMS = 30%</p>	<p>offenders unknown to ip approached ip from_behind and <b>stole</b> ips <b>wallet from_pocket</b> before making_off on_foot offenders x 2 located in close proximity</p> <p>CMS = 8%</p>

Words highlighted in red are significantly characteristic of this cluster (Chi-Square  $p < 0.001$ )

Cluster 1 related to instances of theft from the person occurring principally in entertainment/concert venues, but also, more generally in licensed premises. This was apparent in Figure 8-2 and Figure 8-3 which highlight these tokens as by far the most distinctive (i.e. those

significantly overused) in the cluster. Figure 8-2 presents all of the tokens that significantly represent the cluster. However, the significance values for *entertainment*, *pocket*, *clothing* and *crowd* are so high that all other words are obscured. In Figure 8-3 the top five tokens have been removed to allow the inspection of the remaining significant tokens. To further indicate the characteristics of this cluster, Table 8.4 provides a selection of cases, with the column A listing those cases that were most characteristic of this cluster (highest cluster membership scores) and column B listing cases that represent an average case in this cluster (average cluster membership scores). Column C provides examples of cases that were included in the cluster but may have been a poor fit i.e. they did not share many key attributes characteristic to the cluster.

This cluster was very clearly defined by the *wider environment* in which the offences take place. Drawing from the literature it can be presumed that target search, selection and offence commission appear to take place in the same location. *Entertainment* was very distinctive to this cluster with the token used infrequently in the other clusters with the exception of cluster 11 below. Thus, while a significant number of pickpocketing crimes occurred in entertainment venues, it was also the case that the majority of thefts occurring in entertainment venues were pickpocketing; there were very few robberies in this setting.

Other significant tokens highlighted the features of the wider environment that were conducive to committing pickpocketing offences. *Crowd* featured highly amongst the tokens characterising this cluster, suggesting that offenders may have taken advantage of crowded locations to gain proximity to the victim. This is supported by the fact that there were no significant tokens relating to the offender approaching the victim. The very dominant role of specific environments, in this cluster, pointed to crime pattern theory's concepts of crime generator and crime attractor. However, from the MO descriptions, alone it is not known whether offenders made the journey to these venues specifically to offend or travelled to the venues for other reasons and took advantage of opportunities once they were there.

Other than their location in a suitable location, there was little to indicate why specific victims were targeted. Tokens such as *dance\_floor* and *stood\_at\_bar* may suggest victims were distracted by other things at the time of the offence but this is not clear, even from inspection of the raw data. Despite their location in entertainment venues and licensed premises, there was little reference to alcohol or intoxication.

There were few significant tokens relating to engagement between offenders and victims, in fact, the reverse is true, with significant tokens indicating that, in many cases, victims were aware of little

other than someone *brushing* them past and in a significant number of cases, the *offender was not seen*. There was some evidence of limited verbal engagement between victims and offenders in this script. There were 90 references to *conversation*, a statistically significant level for this script. There was little reference to physical contact, although, 17 cases included '*Hug*' and this was a distinctive token for this cluster and not significantly overused in the other clusters. There is no significant mention of physical violence or threats in this cluster.

Other keywords in this cluster related to the *target enclosure* and indicated, that, in the majority of these cases, the property stolen was located within the victim's clothing (including specifically in their *pocket*). These features were also shared with other clusters including 12 and 14. Of note, is the fact that the token '*bag*' did not occur to a significant level in this cluster, it appears that, in this group of cases, bags were not stolen and items were not removed from bags. Amongst the poorer fitting cases in Table 8.4, clothing and pockets appear to be one of the few attributes shared with the cluster overall.

The *targets* or items stolen included *mobile phones* and *wallets/purses*<sup>37</sup>, again features shared with other clusters. With regards to the transfer of property, there was no reference to the snatching or grabbing of property. The transfer of property, from victim to offender, was conducted covertly and was dependent upon stealth. Thus, items appear to have been *removed* and *pickpocketed*.

### ***Summary of cluster characteristics***

In summary, the cases in this cluster all relate to one clear track for the commission of theft from the person; there was little variation in the cluster relating to how the track was performed. The offences in this cluster are those in which the offender has taken advantage of the proximity afforded by crowded locations and the distraction of would be crime preventers, to remove commonly carried valuable items that are relatively accessible (in the context of a crowded situation) by being carried in pockets. The need for a specific tactic of approach is removed by the crowded nature of the situations; similarly engagement with the victim is not required, in fact, it may be avoided as the transfer of property relies on stealth and secrecy. This cluster, as will others, demonstrates the intrinsic link between the situation in which the offence takes place and the methods chosen to complete it. As can be seen in the dendrogram (Figure 8-1), this cluster was not split when further partitions divided the data. Eighty eight per cent of cases in this cluster were classified as theft from the person, this confirms that the cluster related to theft from the person

---

<sup>37</sup> Purses were coded as wallets for the purpose of tokenisation

scripts but suggests from the remaining 18% that there were a sizable minority of cases that were a poor fit for the cluster.

### Cluster 2: Bag Snatches

Cluster 2 Mapped to five bag snatch robbery tracks, four of which were robbery scripts and one theft from the person.

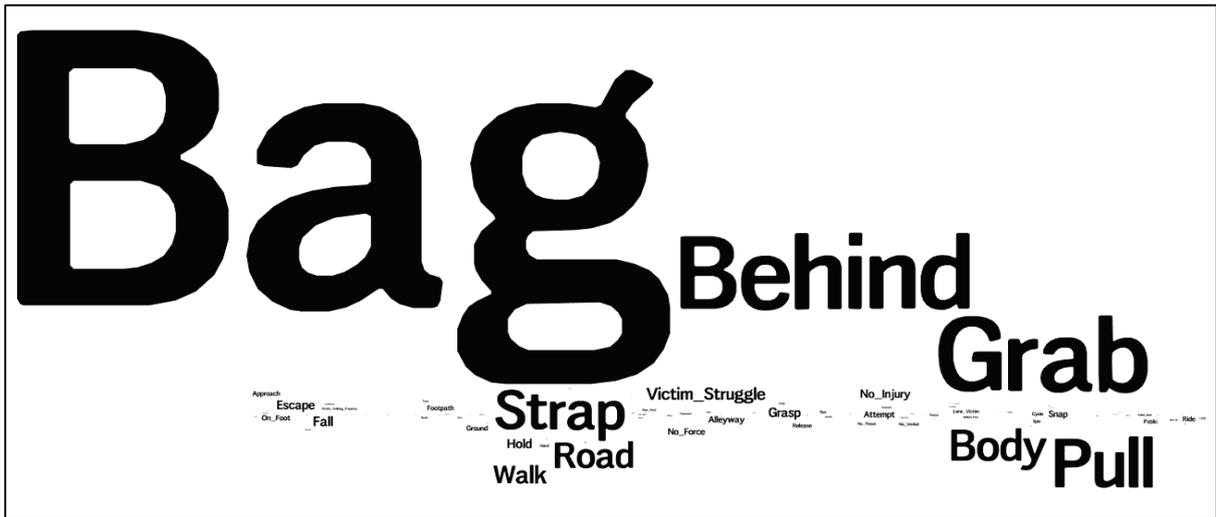


Figure 8-4 Significant tokens in cluster 2 (font size relative to chi-square score).

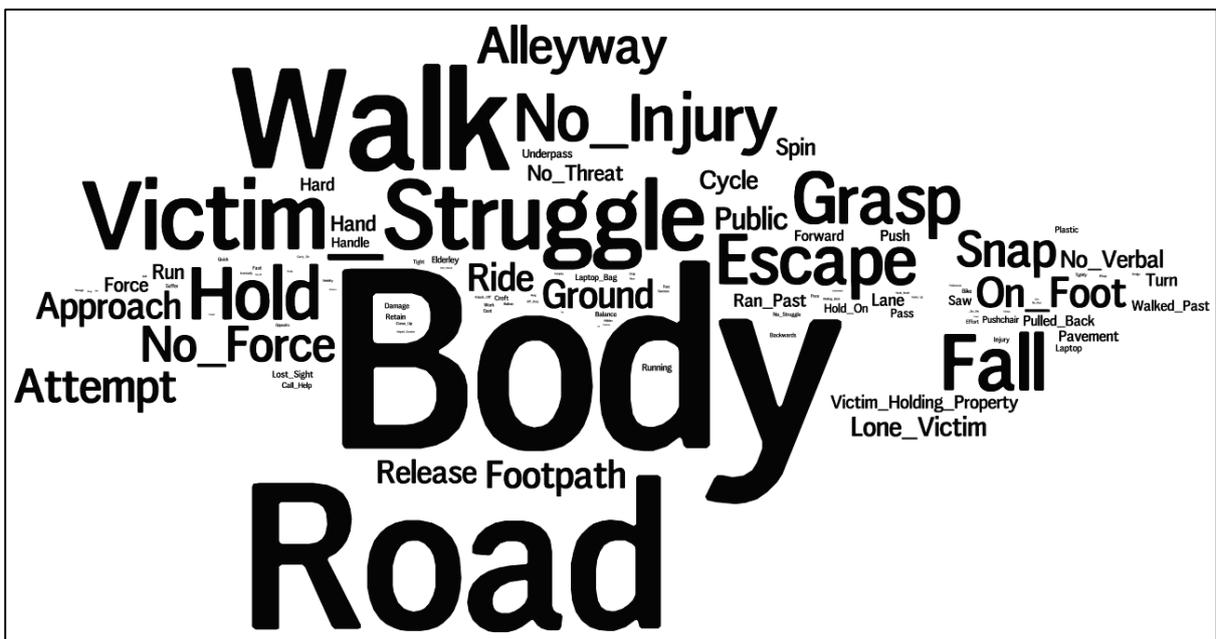


Figure 8-5 Significant tokens in cluster 2. Top 5 tokens removed (font size relative to chi-square score).

Cluster 2 is a relatively large cluster with 2392 cases, around 12% of the cases successfully assigned to clusters, however a fairly varied range of offences and related script tracks are contained within

the cluster. It is apparent from Figure 8-4 and Figure 8-5 that cases in this cluster are grouped together, primarily due to a common *target* i.e. the property stolen which, as is overwhelmingly apparent from Figure 8-4, is *bag*. Other tokens significantly overused in the cluster indicated that this cluster is characterised by theft offences in which bags are stolen by grabbing or snatching the entire bag from the victim, rather than dipping into the bag to retrieve other property.<sup>38</sup>

The wider environment in which offences take place (and most likely the identification of victims too) for this group of cases is characterised by *roads, alleyways and footpaths*. *Underpass* was also significantly over-represented here. In contrast to other clusters, tokens indicating an *approach* were significantly over-represented, although, often cases only stated *approach* and nothing more; cases in this cluster can still be distinguished from those where there was no clear approach. Where the vocabulary used provides more information, it appears to indicate approaches that are engineered to ensure surprise, *run and run-past, behind*.

*Walk* is a significantly over-represented token in this group; however, it is necessary to inspect the context in which the token is used to determine whether this relates to the victim, the offender or both. Inspection of cases revealed that, in the majority of cases, this token was used to describe the victim walking. This is one of the few clusters where some detail relating to victim characteristics was significantly overused, in this case the token *lone\_victim*. While this was significant, it was not present in every case<sup>39</sup>.

The transfer of property, in this cluster, is predominantly achieved through *grab, pull and grasp*. However both *force* and *no force* are significant within this cluster. This shows that there are different script tracks within the cluster<sup>40</sup>. The use of physical force is commonly directed towards the target property rather than the victim (as shown in example Ai in Table 8-5 below) and again, points to the role of surprise and speed to expropriate property. Similarly, both *no\_injury* and *injury* were significantly over-represented in this cluster. Injury was frequently related to falls to the ground caused by the force of the grab (as in examples Bi and Bii below).

---

<sup>38</sup> Note the absence of other target words such as mobile phone or wallet which would be present in a dipping script.

<sup>39</sup> In the frequent MO descriptions that do not refer to either a lone victim, or a victim with friends/others it is not possible to determine whether or not the victim was alone, or whether this information has simply not been recorded.

<sup>40</sup> There may also be different interpretations of 'force' within the data with some officers referring to physical force generally and other using the term to refer to forceful assault

No tokens relating to verbal communication were significantly overused in this cluster. Although verbal communication occurs in some cases (Table 8-5 Example Aiii below provides an example), the lack of speech shows that it is not necessary, and at times, counter to a script track hinged on the element of surprise.

**Table 8-5 Examples of MO descriptions assigned to cluster 2: bag snatches. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

				Cluster Membership Score		
				A) High	B) Average	C) Low
i	between_stated_times unknown offender has <b>approached</b> ip on [name of] <b>road</b> and by using <b>force</b> has <b>snatched</b> ip's <b>handbag</b> unknown offender has then <b>made_off</b> towards [name of] <b>road</b> with <b>handbag. road</b>  CMS = 73%	offender a single white male has <b>approached</b> the ip whilst she was stood at the front door. offender has <b>grabbed</b> the ip's <b>handbag</b> and after a <b>brief struggle</b> where the ip has <b>fallen</b> and <b>injury</b> her head the offender has <b>pulled</b> the <b>bag</b> from her <b>grasp</b> and has then <b>made_off</b> towards the [name of] <b>road. road</b>  CMS = 31%	unknown offender has <b>walked</b> <b>passed</b> disabled ip knocking his <b>arm</b> and <b>pulling</b> his ring <b>from_his_hand</b> . robbery considered. ip did not notice theft element until afterwards. under guidelines from stats this is <b>not</b> robbery. Shoppingcomplex  CMS = 6%			
ii	between_material_times at offence_location 6 asian males have <b>approached</b> the ip <b>from_behind</b> and <b>grabbed</b> her <b>handbag. handbag strap</b> has <b>broken</b> and the <b>bag fallen</b> to the <b>floor</b> . one of the offenders <b>grabbed</b> the money from the <b>bag</b> and the <b>bag</b> and <b>ran_off</b> in an unknown <b>direction. road</b>  CMS = 73%	as ip was <b>walking</b> along [name of] <b>street</b> with her sister as they have got to the <b>junction</b> of [name of] <b>crecident</b> a group of ic1 males aged 16 to 18 years have <b>approached</b> them one of the males has <b>pushed</b> the ip in her <b>left_arm &amp; grabbed hold</b> of her <b>shoulder_bag</b> the ip has <b>hung</b> onto her <b>bag</b> until <b>falling_on</b> the <b>ground</b> the offender & other two males have <b>run_off</b> towards the <b>direction</b> of [place name] <b>road</b>  CMS = 31%	offenders unknown have bumped into ip <b>spinning</b> him round and at the same time taking his wallet and <b>contents</b> offenders <b>walked_off</b> with same <b>no_injuries</b> to ip <b>road</b>  CMS = 6%			
iii	"unknown offender has <b>approached</b> the ip <b>from_behind</b> and then <b>grabbed</b> her <b>bag</b> . the ip has <b>held</b> onto her <b>handbag</b> . the offender has then said give_me the <b>bag</b> " the ip has refused to let go and the offender made good his <b>escape</b> " <b>road</b>  CMS = 71%	31% offender <b>pushed</b> the ip <b>from_behind</b> her left <b>shoulder</b> causing the ip to <b>fall</b> onto her <b>hands</b> and <b>knees</b> and then onto her left <b>shoulder</b> . using unknown means took the ip's <b>handbag</b> from her left <b>shoulder</b> and made good <b>escape</b> <b>running</b> along [name of] <b>road</b> towards [name of] <b>road</b> . the ip suffered <b>sorness</b> and <b>reddening</b> to both <b>palms</b> and <b>pain</b> to her left <b>shoulder. road</b>  CMS = 31%	7% unknown offender has <b>approach</b> ip and <b>grabbed hold</b> of her swipe card. the ip has <b>held</b> onto the card and the offender has let go and <b>walked_away</b> . the offender has <b>pulled</b> the card with <b>enough</b> <b>force</b> to open the lanyard <b>around</b> the ip's <b>neck</b> reclass att. rob. healthcentre  CMS = 7%			

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

The target enclosure is most commonly the victims themselves i.e. the bag is *held* on their *body*; with *body* amongst the most significant token for this group. This presents a challenge for the offender and also the greater potential for the victim to resist, hence the overuse of the token *victim\_struggle*, although, the term did not pervade all the cases. The additional challenges to the offender, in this cluster, may account for the significant over-representation of the token *attempt*.<sup>41</sup>

Escape is mentioned to significant levels, although, again, there is little detail, but it is pertinent that the escape is witnessed and noted in these cases in contrast to more stealthy offences. The token *on\_foot* was significant in this cluster, again this could relate to both offender and victim but tended to relate the offenders' mode of escape (on foot is more distinctive of the local grammar used to describe offender actions than that of victims).

Note that in Table 8.5 example Aiii includes verbal communication between the offender and victim; with the exception of track 2c, discussed below, verbal communication is not characteristic i.e. not significantly overused in this cluster.

As can be seen from the dendrogram (Figure 8.1), cluster 2 was divided into a total of four sub-clusters 2a, 2b, 2c, 2d. In Cluster 2 overall 82% of cases were robberies and the remaining 18% were theft from the person but these proportions varied across the sub-clusters. The sub-clusters also revealed a range of distinct tracks that provided different methods for achieving a bag snatch.

Of these, the first track, sub-cluster 2a retained all the features that were most prominent in the original parent cluster. This group of offences were bag snatches, involving offenders *grabbing* the property from the victim's body (normally the shoulder). Compared to the other three sub-clusters, offences in this track were significantly more likely to involve the *victim resisting* or struggling to keep their bag, and consequently, to involve the victim *falling* to the ground or sustaining an *injury*. 82% of offences in this cluster were classified in the police data as robbery. Track 2a may not be a distinct set of methods, but may represent the consequence of the nature of interaction with victim actions, which the offender must be prepared to handle. Track 2a offences were most likely to occur on the road.

The second sub-cluster also retains many features of the original cluster but is differentiated by a number of new features. For example, a significant proportion of thefts in this cluster are of bags

---

<sup>41</sup> This also indicates that victims are more likely to be aware of failed attempts in offence types where offenders need to reveal themselves in order to execute the transfer. This contrasts to dipping offences where victims may not be aware of failed or aborted attempts at theft.

placed on the back of *pushchairs* and the token *female* is significantly overused within this group (both of these tokens were significant in this sub-cluster compared to the corpus as a whole and to the other sub-clusters in cluster 2). This was the only sub-cluster of cluster 2 in which *verbal aggression* occurred to a significant level. This sub-cluster was predominantly comprised of robbery offences (69%) although this proportion was smaller than for 2a (82%) and 2c (86%). There is, potentially, two tracks here, namely, 2b.1 thefts from pushchairs/trolleys and 2b.2 thefts with verbal aggression.

Sub-clusters 2c and 2d are more distinctive and depart further from the original parent script. There are key differences in the methods of approach but they can still be regarded as related to the other bag snatch tracks. In track 2c, the tokens *ride* and *bike* are significant; a feature not shared with the other sub-clusters in this family. Inspection of these cases confirms that this cluster includes a significant proportion of offences involving offenders on bikes snatching bags from victims as they rode past. Compared to the other sub-clusters, the tokens *behind*, *walk along*, *footpath*, were significant. 86% of cases in this sub-cluster were classified as robbery. Track 2c is a rare group of offences in which a physical resource i.e. cycle, is required to accomplish the offences.

Track 2d can be distinguished from the sub-groups by the significant occurrence of the tokens *follow*, *door* and *doorbell*; again, on inspection, it is clear that this sub-cluster is distinguishable by cases involving offenders following victims, including following them to their home (again, this does not account for all of the offences in this sub-cluster, but those that involve these actions are grouped here). This sub-group is also distinct from all the other groups across the data in that the location in which the victim is identified, differs from the location of approach and property transfer. There were also a significant number of *elderly* victims in this sub-group. Track 2d is also distinctive in that in the police data the offences were predominantly (65%) classified as thefts from the person.

Across the sub-clusters, differences can also be identified in the environments in which sub-group offences take place. Tracks, 2a and 2c included road and alleyway to a significant level, 2b footpaths and bus stops and 2d dwellings. Tracks 2b and 2d can also be distinguished from the other sub-clusters by the significant inclusion of crime preventers including *witnesses* and of individuals giving *chase*; these do not account for all of the cases in these two clusters, but where such actions have occurred, they are grouped in cluster 2b or 2d.

### ***Summary of cluster characteristics***

Overall, cluster 2 covered general bag theft offences but these included both theft from the person and robbery scripts. Five separate tracks were identified in the cluster. Offences were accomplished via an approach which was reliant on stealth and surprise and minimal engagement with the victim (with engagement largely reserved for instances of victim resistance), the transfer of property involved surprise, speed and physical force, but force was used to transfer the property as well as to assault the victim. The target of the theft was mainly bags. Offences tended to take place in public places, out of doors and particularly streets and roads. There were indications that victims were most likely to be walking alone. These offences did not appear to take place in premises such as licensed premises of venues.

This division into script tracks was useful in identifying a set of distinct variations. These tracks retained central features of the parent scripts i.e. stealth, surprise and minimal engagement but differences included: the actions of the victim and the offenders' subsequent reaction, the enclosure surrounding the property (person, pushchair), the mode with which the offender approached the victim (from behind, following, on a bike) and the location at which offences took place (on a road, bus stops and at the victims' dwelling).



**Table 8-6 Examples of MO descriptions assigned to cluster 3: jewellery snatches. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

				Cluster Membership Score		
				A) High	B) Average	C) Low
i	offender <b>approached</b> ip <b>from_rear</b> and <b>tried</b> to <b>snatch</b> 2 <b>gold chains</b> <b>around</b> ip <b>neck</b> . ip <b>grabbed</b> 1 <b>gold chain</b> but offender <b>wrestled</b> other <b>chain</b> from ips <b>neck</b> . this <b>caused</b> ip to go to <b>ground</b> . offender made good escape <b>on_foot</b> then in car with 1 <b>gold chain</b> . public <b>footpath</b>  CMS=79%	between_material_times unknown offender has <b>approached</b> ip and demanded <b>gold chain</b> . ip_refused and ran_off along [xxxx] <b>road</b> 2 x offenders have given chase when ip <b>fell</b> over. offenders <b>punched</b> ip to his <b>face causing pain</b> and <b>discomfort</b> . ip has got <b>back</b> on his <b>feet</b> when offenders pushed him into park area and further <b>assaulted</b> him. following ip <b>screaming</b> offenders made_off <b>on_foot</b> along [xxxx]. <b>road</b>  CMS=38%	bogus official 2 x offenders unknown have gained access to ip's address claiming to be from the utility made tidy search have then <b>struggled</b> with the ip to obtain money on his person. <b>during</b> this <b>struggle</b> the ip has <b>sustained torn</b> skin to his <b>lower</b> right <b>wrist</b> . ip attended [hospital name] butterfly <b>stitches</b> administered to right <b>arm</b> . dwelling  CMS=7%			
ii	between_times_stated 2x uk offenders <b>approached</b> ip told him to give them a <b>gold</b> in <b>colour chain</b> one of the offenders then <b>snatched</b> the <b>chain</b> from <b>around</b> ip <b>neck</b> offenders then made good escape <b>on_foot</b> taking <b>chain</b> from ip <b>road</b>  CMS=77%	between_stated_times known offender <b>approached</b> ip <b>grabbed</b> his t_shirt and <b>tried</b> to steal his <b>neck_chain</b> ip pushed offender <b>away</b> and offender produced a knife offender then searched ips and <b>removed</b> #12. 50 from his pockets <b>road</b>  CMS=38%	between_material_times on the said date offender has entered east wing of the manor <b>hospital</b> and has <b>approached</b> a bed of a sleeping <b>patient</b> who's mobile_phone is on the bed <b>by</b> her offender has taken the phone off the bed while doing so has knocked a draining tube that was attached to the females <b>throat area</b> offender has made good his escape <b>on_foot</b> out_of the <b>hospital ward hospital</b>  CMS=7%			
iii	u/k offenders x2 have <b>approached</b> ip and <b>pulled</b> at <b>gold chain</b> on <b>neck</b> with t_shirt <b>ripping</b> t_shirt and <b>causing injury</b> offenders have then <b>tripped</b> and <b>kicked</b> ip and <b>tried</b> to <b>remove bracelet before</b> running_away towards [xxxx] <b>road</b> offenders made_off with <b>gold chain road</b>  CMS=76%	bet_mat_times ip was crossing <b>road</b> texting on his mobile_phone. ip has then noticed 3 x black males walking towards him ip put his phone <b>away</b> . <b>first</b> offender said'what have you got_for_me <b>blood</b> '. ip has then been <b>grabbed</b> he has <b>fell</b> to the <b>floor</b> offenders have then <b>punched</b> ip to the <b>head</b> <b>grabbed</b> his <b>chain</b> from <b>neck</b> and all 3 made_off <b>road</b>  CMS=38%	offender i/c with friend [xxxx xxxx] was <b>ip_shot</b> in the <b>neck</b> with <b>pellet</b> gun outside location offenders made good thier escape. the ip will need surgery due_to the <b>pellet</b> being <b>lodged</b> in his <b>neck</b> near his <b>vertebrae</b> . <b>road</b>  CMS=8%			

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

Cluster 3 contained 682 cases that, as shown in Figure 8.6, were all clearly related to a script of robbery of *jewellery* (which were predominantly necklaces and neck chains). Therefore, the most predominant tokens to characterise this cluster relate to the *target* stolen, that is, *jewellery* and *gold*. Two contrasting wider environments, were significantly over-represented in the cluster, namely, *road* and *licensed\_premise*. Although *road* was significant, *alleyway*, *footpath* were not.

There did not appear to be any distinctive features of victims in this cluster, other than presumably that they were wearing valuable and visible jewellery. However, the tokens *acquaintance* and *domestic* were both significant, both tokens relate chiefly to partners and ex-partners showing that, in a small but significant number of cases (at least 31), the offender and victim were known to each other – a feature which was distinctive to this cluster.

There were no significant tokens describing the nature of approach, this may indicate that offences occurred very quickly with offenders benefiting from surprising the victim. In common with the bag snatch script in cluster 2, the transfer of the target property was through physical force, as illustrated with the predominance of tokens (many of which are shared with cluster 2) including *grab*, *pull* and *rip*. There appeared to be more instances where force is directed at the victim as well as towards the target property. As with the tracks in cluster 2 the enclosure was the victim's body, however, in this Jewellery Snatch script the enclosure tended to be the victim's neck and breaching this enclosure appears to have required greater proximity and greater force requiring the more frequent use of physical force against the victim and resulting in the significant inclusion of the terms *assault* and *injury*; while *no\_injury* was not significant.<sup>42</sup> The more frequent use of force is reflected in the police classification of offences in this cluster, 94% of which were coded as robberies of personal property.

Although there was limited evidence of engagement of victims by offenders in this script, the token *argue* was significant. Examination of these cases shows that some offences begin with an argument and property is taken following the altercation. This was supported by word sequence analysis that showed that transfer-related tokens such as *snatch\_grab* were more likely to be successors to, rather than precursors to, the token *argue* (*snatch\_grab* is a successor to *argue* in 118 and precedes it in 40 cases).

---

<sup>42</sup> Note that *scratch* is also a significantly overused token in this cluster. *Scratch* was not coded to *injury* because *scratch* is an ambiguous term, not all scratches relate to human injury. Accurate coding of this token would have required an inspection and disambiguation of all instances of the words

Although the number of cases was small (4 from a total of 9 in the database) the token *victim\_shot* was significant in this cluster, indicating a (rarely used) script permutation with a high level of violence.

There was evidence of victim resistance in this script, indicated by the significance of tokens such as *victim\_defence* and *victim\_struggle*. Inspecting the cases revealed that victim defence tended to relate to descriptions where victims are assaulted prior to transfer and they aim to defend themselves. In comparison, victim struggle tended to be used where victims aim to resist the force applied to the target. In a large proportion of cases (25/45) this struggle was met by an escalation of violence from the offender towards the victim. Mention of and details about methods of escape were not included in this to a significant extent although, *on\_foot* was significant which can relate to both approach and escape.

As seen in the dendrogram (Figure 8.1 above), cluster 3 did not separate into further sub-scripts or tracks.

#### ***Summary of cluster characteristics***

This cluster was interpreted as a distinct script track of jewellery theft taken from the victim's body. The offences took place on public roads and also in licensed premises. There was little information regarding victims, although, there appeared to be an available permutation in which property was taken from close acquaintances.

The script was built on the element of surprise, with minimal engagement with the victim and transfer taking place through snatching and grabbing. There is a possible track option in which engagement with the victim began with argumentative engagement. Although, in most cases, force was initially applied to the target property, there was an apparent preparedness to use greater force, and force against the victim, where necessary. This escalation of force commonly followed victim resistance. Here, there was an overlap with scripts using bold controlling methods of transfer. There was an identifiable track using weapons, specifically firearms, but this related to a small number of cases (4). There is little data relating to escape or disposal of property, although, escape was characteristically on foot.



**Table 8-7 Examples of MO descriptions assigned to cluster 4: theft of cycles. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

		Cluster Membership Score		
		A) High	B) Average	C) Low
i	<p>offenders have <b>approached</b> ip have <b>stopped</b> him on his <b>bike</b>. have <b>told</b> him to <b>hand_over bike</b>. <b>ip_refused</b>. he has then been <b>threatened</b> and removed from his <b>bike</b> by two offenders. <b>bike</b> has_been taken from him and <b>rode</b> off by the offender. other offenders have then made good thier escape. <b>road</b></p> <p>CMS – 87%</p>	<p>at time stated ip was <b>approached</b> by <b>male</b> as described in the <b>skateboard</b> area at location. <b>male asked</b> if he <b>could</b> have a <b>go_on</b> his <b>bike</b> and <b>ip_refused</b>. <b>male</b> then <b>said</b> he had a similar <b>bike</b> that had_been stolen and he would take it <b>home</b> to see if the <b>serial</b> numbers<b>matched</b>. <b>male</b> took the <b>bike</b> and made good escape. he was <b>in_company</b> with another 2 ic1 <b>males</b>. <b>sportsclub</b></p> <p>CMS – 43%</p>	<p>between stated_times u/k offender/s have <b>approached</b> the ip and her <b>friend</b> [name of] <b>lane</b>. u/k offenders then physically <b>pushed</b> the ip on her blue and silver <b>scooter</b>. u/k offenders have then stolen the <b>scooter</b> and made_off to return a short time later and <b>chased</b> ip and <b>friend</b> into [xxxx] <b>park</b>. <b>road</b></p> <p>CMS – 12%</p>	
ii	<p>known offender has <b>approached</b> ip whilst he was on his <b>bike</b> offender has <b>told</b> ip to <b>give</b> him his <b>bike ip_refused</b> offender has then <b>told</b> ip that he would <b>batter</b> him if he <b>did</b> not <b>get_off</b> his <b>bike</b> offender has then <b>pushed</b> the ip off his <b>bike</b> causing <b>no_injury</b> offender has <b>ridden</b> off with ips <b>bike road</b></p> <p>CMS – 86%</p>	<p><b>known</b> offender has <b>approached</b> ip whilst he <b>was</b> on his <b>bike</b> offender has <b>told</b> ip to <b>give</b> him his <b>bike ip_refused</b> offender has then <b>told</b> ip that he would <b>batter</b> him if he <b>did</b> not <b>get_off</b> his <b>bike</b> offender has then <b>pushed</b> the ip off his <b>bike</b> causing <b>no_injury</b> offender has <b>ridden</b> with ips <b>bike road</b></p> <p>CMS – 43%</p>	<p>offender u/k has <b>approached</b> ip who was <b>pushing</b> a <b>scrambler</b> &amp; offender has <b>said im</b> having this offender has then taken the <b>scrambler</b> from ip without force being used however ip <b>fear_for_his_safety</b> on being <b>approached</b> by the offender <b>wasteground</b></p> <p>CMS – 10%</p>	
iii	<p>bmt u/k offender <b>approached</b> ip who was on his <b>bm</b>x. got on <b>bike pushing</b> ip off and then made good escape on the <b>bike</b>. <b>road</b></p> <p>CMS – 85%</p>	<p>ip has_been <b>approached</b> by a <b>group</b> of 5 7 <b>males</b> on <b>bikes</b>. offender 1 asked_the_time. the <b>group surrounded</b> the ip making her feel <b>intimidated</b>. offender 2 pulled bag from ip's shoulder <b>group</b> made_off towards [place name] on thier [sp] <b>bikes</b>. <b>road</b></p> <p>CMS – 43%</p>	<p>offenders unknown <b>approached</b> ip on yellow and black honda <b>motor_bike</b>. offender ( 2 ) searched ip looking for items of value making <b>threats</b> if ip <b>did</b> not <b>comply</b>. offender found nothing on ip. offenders then got on <b>motor_bike</b> and <b>rode</b> off having stolen items from ips <b>friends</b> as_per cross_reference above. <b>farm</b></p> <p>CMS – 6%</p>	

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

Cluster 4 contained 757 cases which describe a distinct script involving thefts of cycles. The fact that these thefts were found within a database of theft and robbery from the person offences, suggests that these cycles were taken from the person, when the victim was sitting on, riding or pushing<sup>43</sup> the bike. Cases predominantly involve pedal cycles, but also include motor bikes and scooters.<sup>44</sup> This cluster represents a script for theft of cycles rather than scripts where offenders used bikes as a resource. This was evidenced from the inspection of characteristic cases (such as those in Table 8-7) and from the absence of any other significant tokens relating to types of property stolen. Further support was drawn from analysis of word sequences, illustrated in Figure 8-10. Tokens that were used within close proximity to *bike* in Cluster 4. This diagram shows the tokens that most frequently preceded the token *bike* in cluster 4. Words connected to *bike*, by shorter lines, were more frequent predecessors. In the diagram, the predecessors are generally indicative of bikes being stolen. However 'ride->bike' and 'approach-> bike' do not help to clarify the role of the bike in the description. It was not always clear, even from the original text descriptions, who was riding the bike as demonstrated in example Ciii in Table 8.7.

Significant tokens that indicated the track's wider environment, included *park* and *road*. *Road* occurs more frequently but *park* received a higher relevance chi-square score because *park* occurred less frequently in other clusters. There was little to indicate the nature of victims in this cluster, although, the tokens *friend* and *youth* were significant. Inspection of the data shows that, as a characteristic of the local grammar, *friend* was predominantly used in relation to the victim and his or her friends whereas *youth* almost exclusively referred to the offender or offenders.

It was clear that offenders approached victims by *stopping* them. This highlighted the mobility, or potential mobility, of victims who are on bikes and the offender(s) need to take conscious steps to contain victims for the duration of the offence. How victims were stopped was often not clear from the descriptions, although, *cornered* was an overused token. There was a higher level of engagement between victim and offender than seen in the clusters so far. It is not possible to steal a bike merely through speed, surprise or secrecy when the victim is effectively the target enclosure and needs to be removed from the bike. It was, therefore, necessary for offenders to communicate their intention to steal by making demands. This was reflected in the significance of tokens *demand*, *verbal*, *get-off* and *hand over*. This communication was frequently aggressive, with the tokens *verbal\_aggression*

---

<sup>43</sup> Note that "*push bike*" is recognised by the coding dictionary as *bike* therefore the token *push* relates to the verb

<sup>44</sup> Note that as the cluster analysis does not determine the meaning of words, the location of pedal cycles and motor cycles within the same cluster is not a consequence of their semantic similarity, rather it is a result of the tendency for MO descriptions including pedal cycles and those containing motor cycles to share common vocabulary.

and *threat* significantly over-represented here. Perhaps as a result of the explicit requests for the property, instances of *victim\_refuse* were significant. The descriptions recount that victims of crimes reported being in fear during the offence, this feature is shared with clusters 7, 10 and 14<sup>45</sup>. This is a probable result of the increased engagement, communication and use of verbal demands. Tokens associated with assault did not occur to a significant degree, although, the significance of *push* indicated that victims are removed from the bike by force and *knife* was also significant. This did not occur in all cases but it was more likely to be used here than other clusters. The absence of tokens relating to injury suggested that knives were used in threats rather than to directly injure. There was evidence of victim refusal but this does not appear to have resulted in escalation of violence as seen in other scripts. There were no significant tokens relating to escape, although, inspection of cases suggested that offenders frequently used the bike itself as means of escape. The close association of bike and escape are also observed in Figure 8-10

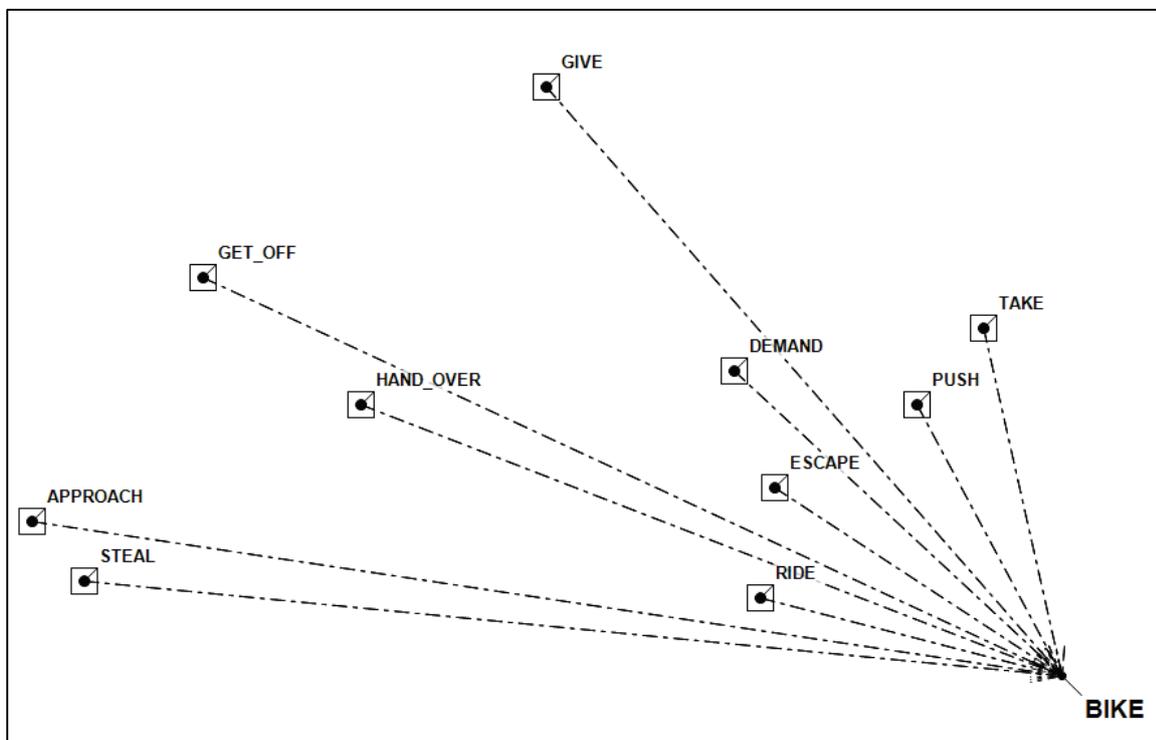


Figure 8-10 Tokens that were used within close proximity to *bike* in Cluster 4

<sup>45</sup> It should be noted that officers specifically seek to clarify whether victims felt fear as this element is crucial in determining the distinction between theft from the person and robbery in charging and prosecution.

### ***Summary of cluster characteristics***

The track represented by this cluster related to a robbery of personal property track for thefts of bicycles and motorcycles. The track was based on control of the victim, but control was gained through verbal aggression and threats rather than direct physical assault. The offence environments included public spaces, particularly roads and parks. The level and extent to which offenders engaged with victims was greater than for other scripts, this included taking steps to 'stop' the victim, the request for the property was also made explicit through verbal requests and the demeanour of offenders appears to be more threatening to victims. However violence, assault and injury were rare in the cases in this script.

### ***Cluster 5: Theft from Bags (Shops)***

**Cluster 5 Mapped to four tracks, two theft from the person theft from bag tracks and one thefts from shop staff track.**

Cluster 5 contained 2185 cases, the majority of which were thefts from bags, sometimes referred to as 'dipping'. A clear theft from bags in shops script was identified. Within the script three separate tracks were identified and additional robbery track which was distinct from the other three was also identified.

Figure 8.11 and Figure 8-12 highlight the words that were significant to the whole of cluster 5. The distinction between the theft from bags script here and the theft of bag script in cluster 2 above, is apparent from alternative target-related tokens visible in the wordclouds below, most notably *wallet*<sup>46</sup> but also *money* and *bank\_cards*. It is interesting that other items, such as mobile phones, were not significantly included in this cluster. This cluster was characterised by the wider environment in which these offences took place, the items stolen, the target enclosure and the secrecy under which thefts took place. Typical environments were *shop* and *supermarket*. *Bank* was also significantly mentioned, although, this was not always the location of the theft; in some cases victims had visited the bank or *ATM* to withdraw money shortly before the offence. The MO descriptions do not make clear (probably because it was unknown) whether these visits were causally linked to the crime or a mere coincidence. The regularity of this pattern may suggest a causal connection. Payment and transaction points (e.g. *pay*, *till*) were also significantly over

---

<sup>46</sup> The token wallet includes both purses and wallets.

reported in this cluster. Victims either had items stolen shortly after completing a transaction or noticed that their property is missing at the time/place when they came to pay for something



Figure 8-11 Significant tokens in cluster 5. (font size is relative to chi-square score).



Figure 8-12 Significant tokens in cluster 5 – top tokens removed (font size is relative to chi-square score).

In addition to bag, there were significant tokens relating to other levels of enclosure. *Undo* and *zip*, were both significantly overused in this cluster, showing that offenders need to breach the enclosure of the bag. In addition, the significant overuse of *shopping\_trolley* and *trolley* indicate that bags were placed inside or were hanging on trolleys at the time of the thefts.

In this cluster, the theft is accomplished by disguising the act from the victim. This secrecy is reflected in the significance of the tokens *notice* (related to not noticing), *missing*, *offender\_not\_seen* and *nothing\_felt*.

There was little contact or engagement between offender and victim in this cluster. One exception are the tokens *brush\_past* and *bump-into*; with victims often not realising until later that this contact was likely the result of an offence being committed. There was also little verbal communication or physical contact in this cluster. However, *distract* and *display* (usually looking at a display) were both highly significant in this script, suggesting that perpetrators create or take advantage of the victims' inattention in order to transfer the possession of property. It is notable that many of the MO descriptions for this script are very brief (see examples Ai to iiii in Table 8.8), this is indicative of the fact that the nature of these offences leaves many of the details unknown. Tokens that were notable, by their absence, include those that relate to communication, force, injury and escape.

Cluster 5 was divided to produce a total of four sub-clusters, two of which (sub-clusters 5a and 5b) clearly provide permutations of the parent theft from bags in shops track. The remaining two sub-clusters appear to relate to distinct tracks. The permutations identified in 5a and 5b retained the characteristics of the parent theft from bags in shops track, with the significant overuse of tokens including *shop*, *bag*, *undo*. There is little to distinguish these two clusters apart from the fact that in track 5b, the tokens *display*, *browse*, *look* and *distract* were significantly overused. These tokens provide more detail on the actions of victims' at the time of the offence; these activities may have provided an opportunity for offenders (see example Biii Table 8.8). However, this is not to say that the victims of offences, described in track 5a, were not similarly distracted; the difference between these two tracks may only be the result of the decision to preserve different details about the offence.

The third sub-cluster, was more distinctive from the parent cluster in that it was dominated by tokens relating to *banks*, *bank cards* and *cash withdrawals* (an illustration can be seen in example Bi in Table 8.8). The sub-cluster contained a mixture of offences occurring at or near banks, offences that occurred shortly after a victim visited a bank or cash machine, and cases where victims realised that the offence had taken place when they later came to use their bank card. The first of these groups is straightforward to identify because offence locations have been comprehensively labelled and can be used to identify this track, this might not be the case in data from other forces. This could be regarded as a separate track as the offences occur in proximity to banks, rather than shops. The group of cases where victims notice that property is missing when they come to use their card could also be regarded as a separate track but it is more accurately an element of victim activity which cuts

across several script tracks. Following a victim from a bank to a place that is conducive to bag theft, could be regarded as a distinct track.

**Table 8-8 Examples of MO descriptions assigned to cluster 5: theft from bags (shops). Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

	Cluster Membership Score		
	A) High	B)	C) Low
i	unknown offenders stole purse from ips handbag made good his/her/thier escape shop CMS=71%	between material_times offender ip has attended bank and withdrawn #2000 money placed in handbag ip left bank and made her way to shop in [xxxx] town_centre upon leaving store ip discovered #1000 missing from handbag shop CMS=71%	between 2200 31. 12. 06 and 0200 01. 01. 07 ip was at [xxxx] whilst on the dancefloor he felt somebody touch him around his right pocket area. on attending the bar approx 15 mins later the ip reached inside his pocket to find that his wallet had been stolen. inside was approx #30 driving licence and army id_card. offender ( s ) was not_seen entertainment CMS=71%
ii	offenders unknown stole purse from ip's handbag by means unknown made good escape shop CMS=71%	unknown offender ( s ) have taken ips purse which she states was either in her coat pocket or small bag. both of which were on her person whilst shopping in [xxxx]. offender ( s ) made good thier escape. shop CMS=33%	"the ip befriended a male by the name of [xxxx]. he withdrew some money from a cashpoint with [xxxx] in attendance. later that evening his card was fraudulently used on six occasions. this was after [xxxx] and the offender had been dancing together. the offender disappeared for no reason. location and offence changed in_line_with balance_of_probabilities. nightclub CMS=5%
iii	uk offender/s have approached ip & by uk means stolen her purse from out_of her shoulder_bag & made good thier escape shop CMS=71%	on above date and between times person ( s ) unknown opened ips hand_bag and stole her purse while she was looking at display in location hand_bag was on ips left_arm but ip felt no pushing or pulling and was only aware of the theft when she came to pay for an item made good thier escape shop CMS=33%	"between times stated unknown offender has entered premises via public access and walked behind the service counter and started pressing buttons on the till. the pharmacist tried to stop him by words alone where the offender just shouted open the till""the offender then grabbed the supervisor and shouted""open the till"". the supervisor believed she would be attacked if she did not so she opened the till where he grabbed notes and pound coins and made good his escape by foot towards [xxxx] road.""shop CMS=4%

Words highlighted in red are significantly characteristics of this cluster (Chi-Square  $p < 0.001$ )

The fourth sub-group of this cluster, 5d, was very distinct from the other sub-clusters. As with the others, this grouping is characterised by tokens that relate to shops. However, additional tokens were identified which were not significant in the other sub-clusters in this group. These included *till*, *counter*, and *staff*. Also significant were tokens such as *enter*, *take away* and *customer*. Inspection of cases from this cluster revealed that group 5d are not theft from bags in shops. They are, in fact,

robberies committed in shops (and take away restaurants), in which the person from whom property is transferred is a member of staff, as shown in the example Ciii in Table 8.8<sup>47</sup>. Sub-cluster 5d is a very distinct grouping of 88 cases which were hidden away within the parent cluster. These cases are predominantly robberies of personal property (78%) in contrast to 5a-5c each of which comprised 91% thefts from the person. Sub-cluster 5d represents an additional unrelated track, rather than an alternative track for the theft from bags (shops). The significance of tokens such as *behind*, *open*, *bag* and *shop* resulted in these crimes being grouped with the theft from bag offences, even though, in reality, the offences have little else in common. Identifying this cluster would be important to counter these thefts. These crimes could not be identified from the crime classification alone, but their identification required careful investigation of cluster solutions. This finding highlights the influence of the number of clusters in a solution and the difficult balance between simplicity and detail.

### ***Summary of cluster characteristics***

Overall, the theft from bags in shops track was accomplished by secrecy and ensuring that the victim was unaware that an offence was taking place. Consequently, engagement and contact with the victim is negligible. The primary targets in this track were wallets, money and bank cards, which were enclosed within bags, and at times, in or on shopping trolleys. Drilling down into the cluster revealed two additional permutations and two additional distinct tracks. One permutation of the theft from bags in shops track took advantage of (and/or created) victim distractions to create the opportunity to act without the victims' knowledge. Another permutation appeared to identify victims' making transactions as possible targets. Two tracks were identified that were not related to thefts of bags from shops, the first of these was a script for thefts from bags in the proximity of banks and finally, a track was identified that related to robberies of staff in shops.

---

<sup>47</sup> Example 5i is a poor fit to the original cluster indicating its low relevance to the overall theft of bag from shops script.



**Table 8-9 Examples of MO descriptions assigned to cluster 6: robberies with dwellings. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

				Cluster Membership Score		
				A) High	B) Average	C) Low
i	between stated material_times 6 unknown offenders have <b>forced entry</b> into <b>property</b> via <b>front door</b> . <b>occupants asleep upstairs</b> . offenders in_possession_of <b>knives untidy search</b> of <b>rear bedroom items</b> selected and the offenders made good escape through <b>front door</b> and made_off on_foot. <b>dwelling</b>  CMS= 62%	between times 3 x unknown white males have <b>forced entry</b> to offence_location through <b>front door</b> whilst ip was watching <b>television</b> in the <b>lounge</b> . one male has <b>pushed</b> ip and reached into her handbag taking #150 <b>cash lying loose inside</b> . other two males have <b>searched</b> rest of <b>downstairs</b> . nothing else taken. offenders made good thier escape. ip not_injured. ip <b>pushed</b> immediatly prior_to <b>cash</b> being <b>stolen</b> ). <b>dwelling</b>  CMS=25%	known offender [name]dob: 21/1/86 <b>attended</b> the location & <b>accused</b> the ip of having something to do with some <b>charges</b> on his account relating to a <b>mobile_phone</b> which the ip states he has no_knowledge of. the offender has then ripped the ip's <b>mobile_phone</b> from around his neck & then left. <b>dwelling</b>  CMS=6%			
ii	60% between times stated unknown offenders have gained access via <b>insecure ground_floor window</b> and conducted an <b>untidy search</b> of all <b>rooms</b> . when entering upstairs <b>bedroom</b> they have found 1 x occupant <b>in bed</b> occupant has_been told to <b>shut_up dwelling</b>  CMS=60%	25% between stated_times 3 offenders have <b>entered</b> the premises <b>armed</b> with <b>weapons knives bats</b> and a <b>baton</b> . <b>threatened</b> 2 of the <b>occupants</b> to hand_over thier stuff. offenders then began a <b>search</b> of the premises and <b>stole property</b> . offenders made good escape. <b>dwelling</b>  CMS=25%	5% "offenders whilst <b>legitimately</b> at offence_location have argued over equal <b>allocation</b> of <b>coke</b> " for purposes of <b>supply</b> " <b>dwelling</b>  CMS=5%			
iii	uk offenders x 2 <b>entered terrace dwelling</b> via <b>insecure kitchen door ( rear )</b> and <b>entered living_room</b> where ip was offenders then <b>pushed</b> ip on chest causing ip to fall <b>back_into armchair</b> . offenders <b>demanded money</b> from ip who then took offenders <b>upstairs</b> to <b>bedroom</b> and <b>stole money</b> from <b>bedroom untidy search</b> offenders left via <b>front dwelling</b>  CMS=57%	"between stated_times and date two unknown offender <b>forced entry</b> via <b>rear lounge ground_floor window</b> . offenders produced 18 inch <b>knife</b> and a <b>iron rod</b> . offenders <b>demanded</b> ip's <b>property</b> as ip fear for his saefty he handed_over the <b>property</b> ." " <b>dwelling</b>  CMS=25%	the complainant was <b>asleep</b> at location when offender removed ip's <b>bank_card</b> from his wallet which was located in his <b>rear</b> trouser pocket. offender made_off with <b>bank_card</b> direction u/k. ip's <b>ex_girlfriend</b> was at location during the time_of_offence and had not <b>returned</b> despite requests from ip. ip suspects <b>ex_girlfriend</b> to be responsible. <b>dwelling</b>  CMS=6%			

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

Cluster 6 contained 1355 cases that represented a track for robberies of personal property that were committed within dwellings (96% of the cases in this cluster had been classified as robbery of personal property). *Dwelling* was significantly over-represented, as were related tokens including

*door, bedroom, room, flat, kitchen*. Note that, although the location variable (value *dwelling*) was influential in defining this cluster, other tokens related to domestic dwellings were identified and drawn directly from the free text field that also point to domestic settings.

In this track, dwellings provided the wider environment, but could also be considered a layer of *enclosure* that must be breached to commit the offence. As the offences were robbery of *personal property*, a second layer of enclosure was the victim themselves. To accomplish this track, offenders must gain entry to the dwelling and gain the compliance of the victim.

Analysis of the cases in this cluster revealed a number of different permutations available to complete the tracks, these were used in varying combinations. These include the use of physical force, verbal aggression and some degree of secrecy. These show flexibility in the way in which the tracks were performed but it is difficult to describe them as distinct tracks because the options were not mutually exclusive. For example the track could be performed using physical force *and* verbal aggression. Another occurrence might begin with secrecy but then move to the use of physical force.

In these offences the target enclosure (i.e. the dwelling) was frequently breached using physical force, as seen with the overused token *forced\_entry*. *Force* is also a significant token but, in isolation, it is unclear whether this force is exerted to gain entry to the property, to control the victim or both. In another permutation of the track, entry into the dwelling was gained more conventionally by knocking on the door (*knocked\_door* and *doorbell*) and engaging with the occupant. Only 9 offence descriptions in the whole dataset included the token *bogus*, six of these are located in this cluster, relating to the use of deception to gain entry, making the token statistically significant for this cluster. *Distract* was not statistically significant for this cluster. There was a small opportunity for secrecy within this script, in that dwellings were entered through *insecure* doors and windows to a significant degree. However, transfer of property from the victim was not conducted through stealth. The track included permutations providing an element of surprise in that in a significant number of cases the occupants were asleep at the time of the break in (tokens *sleep* and *wake*).

There was a high degree of engagement between offenders and victims in the track. Offenders in these cases did not attempt to conceal their actions from their victims, although, in a significant number of cases, offenders attempted to conceal their identity (*face\_covered*). The level of verbal communication was high, aggressive and threatening. *Verbal* communication and *verbal aggression* were both significant, along with *threats*, including *threats to kill* and *threats to shoot*. These threats were reinforced with *weapons*, including *knives* and *firearms*. The success of these threats was, perhaps, demonstrated by the absence of victim resistance in this cluster of cases.

Despite the level of threats in this track, the actual use of physical force against victims appeared to be limited. *Push* was significant but *violence*, *assault* and *injury* did not occur within this track to significant levels. Victim compliance and transfer of property was, therefore, gained via bold threats and aggression rather than direct physical violence. This was the first of all of the tracks discussed so far to have included a significant mention of a *search* (for items to be stolen) and in the case of these crimes, the majority of searches were *untidy*.

For the majority of those cases where the information is available, the object of theft appears to be *money* although *TV*, *games console*, *computer* and *laptop* were also significant. Also significant was *drugs*, this included theft of drugs (including cannabis plants) but also demands made by offenders for drugs, it is unclear in the latter cases, whether any drugs were actually available in the dwelling.

As the MO descriptions all focus on the activities occurring within the dwelling, looking at MO descriptions in isolation provides little information to suggest why these dwellings or these victims were selected. Although the combination of MO descriptions with other variables would help to investigate this.

Due to the problems with identifying the number of offenders involved in each case (detailed in Chapter 6) the analysis provides limited information on the role of co-offenders for each script. The token *group* was not used significantly to describe this script, however, inspection of cases reveals there are certainly permutations available that exploit group offending. Benefits of co-offending included increasing the threat and control over the victim and allowing a division of labour between tasks such as distraction, use of force and searching.

Although there are a number of permutations through which this script can be accomplished, drilling down did not identify any sub-clusters or specific tracks. This is likely to be because different permutations tended to be used in variant combinations rather than providing distinct tracks.

### ***Summary of cluster characteristics***

This cluster could be interpreted as one track relating to robberies occurring in dwellings. The wider environment, or outer enclosure of the dwelling, provided an additional challenge to the offender in the need to overcome both access to the dwelling and control over the victim. These challenges were addressed through a wide range of permutations covering secrecy, surprise, threats and force. The level of engagement, between victim and offender was high, with a high level of aggression. Aggression was expressed verbally and through brandishing weapons, the use of physical force was restricted to entering the dwelling rather than controlling the victim.



**Table 8-10 Examples of MO descriptions assigned to cluster 7: demands with threats. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

Cluster Membership Score			
	A) High	B) Average	C) Low
i	<p>"ip was walking with offender when offender demanded money then said i'm going to cut you""ip in_fear_of_his_safety handed_over his property offender then made good escape""road</p> <p>CMS=66%</p>	<p>"bmt u/k offender has approached ip and whilst clearly holding a 6 bladed kitchen_knife has said""give_me your phone "road</p> <p>CMS=31%</p>	<p>offender attended [xxxx] childrens_home. whilst at this location he has threatened a staff_member with a small knife stating yo carehome</p> <p>CMS=10%</p>
ii	<p>"bmt u/k male approached ip produced a knife and said to ip give_me your wallet"". ip handed it over fear the knife. offender made good his escape in direction of [name of] road.""road</p> <p>CMS=64%</p>	<p>ip has_been approached by 4 u/k males who have surrounded him and demanded money before implying that the ip would die if he did not adhere to demands ip has screamed and offender/s have made_good_escape. fear_for_his_personal_safety road</p> <p>CMS=31%</p>	<p>offenders unknown have approached ip and asked for_a cigarette. ip has pulled_out a packet of cigarettes containing #95 which is on show within the clear film of the packaging. offender has snatched cigarette packet and made good his escape on_foot road</p> <p>CMS=9%</p>
iii	<p>offender unknown approached ip in the street offender has then said to ip'give_me your money or i'll stab_you'ip fear_for_safety has handed_over #40 in cash. offender then said'give_me your mobile_phone '. ip then handed_over mobile_phone. offender then made good escape. public footpath</p> <p>CMS=33%</p>	<p>"3 x unknown offenders approached 3 x ips one held one ip around throat from_behind with hunting knife to his neck while another showed the other two a craft knife and said hand_over your wallets phones everything now""ips handed_over property and offenders made good escape""road</p> <p>CMS=31%</p>	<p>rec between material_times [name] son of [name] has attended ip's flat and escorted him to post_office where he has waited for ip to collect benefits ( #91. 00 ) then demanded monies off ip ie ( #91. 00 ) and given him back #10. 00. 03/08/09. full circs unclear. mo reads possibly as robbery. suggest we give this some time to be done. public footpath</p> <p>CMS=8%</p>

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

Cluster 7 contained 1472 cases that can be mapped to a robbery script characterised by threats, two sub-clusters reflected clear tracks, one involving knives (7b), the other not (7a). Both tracks had been

classified as predominantly robbery of personal property (99% of cases in 7a and 95% of cases in 7b). The track is characterised by a bold strategy, gaining compliance from the victim through aggressive demands and threats creating feelings of fear for victims. The most dominant token in the wordcloud, Figure 8.15 is *hand\_over* representing demands made by the offender that the victims relinquish their property. Other significant tokens related to the strategy of verbal aggression included *threats*, *demand* and *fear*.

Wider environments characteristic of both tracks included *alleyway*, *road* and *underpass* suggesting that these offences took advantage of secluded locations with poor guardianship and surveillance. Other types of location included proximity to *ATMs* showing that offenders selected sites with an increased likelihood that victims will carry money, the most dominant property targeted in the script (mobile phones are also significant as targets).

The analysis of MO descriptions provides little detail on the victims of these offences, neither tokens relating to being alone or those relating to being with company were significant.

As with other tracks, there is little data relating to offenders' approach. *Approach* itself is highlighted as significant, showing that, in contrast to secretive scripts, the offender is seen by the victim. Offenders had their *face\_covered* in a significant number of cases, countering the risk of detection posed by this bold script. The token *behind* was not significant, suggesting that with the use of a bold strategy, the element of surprise may be less central.

The offenders' first contact with victims was either directly aggressive or initially innocuous. The token *verbal* related to non-aggressive conversation, including asking for the time, asking for a cigarette or a light. In other script permutations, contact was aggressive from its initiation (as seen in examples Aii and Aiii of Table 8.10). Another characteristic of this script was the use of weapons to reinforce threats. This predominantly involved *producing* or *pulling\_out a knife*, or other *object* used as a weapon. An extremely small, but significant number of cases (5), involved *syringes*. In other cases, the weapon was not seen, but the offender *indicated* or the victim *believed* that the offender had a weapon. However, interaction between offender and victim was principally verbal and weapons were rarely used. The token *stab* was significant but related to the content of threats issued by the offender e.g. 'The offender threatened that he was going to stab her'.<sup>48</sup> The absence

---

<sup>48</sup> An attempt was made through dictionary building to differentiate between actions of violence and threats of violence. For example the phrases 'Stab you', 'Stab ya', 'Shank ya' were identified and coded as 'threats.' However not all threats were successfully identified and later inspection of the clusters revealed the use of alternative phrases such as 'threatened he was going to stab him.' From this phrase the threat would be identified but whether or not the mention of stab referred to the same threat or an actual stabbing would be unclear. The

of tokens relating to assault or injury further confirms the predominance of threats over physical violence in this track.

Offenders demanded property from victims, which was then relinquished due to fear. In this script, offenders did not search pockets or bags, as demand led to victims handing over property. The sequence of threats, fear and compliance is demonstrated by the examples in Table 8.10. Tokens relating to snatching and grabbing property were not significant as the victim transferred the property to the offender, albeit under duress. *Victim\_refuse* was significantly overused in this cluster, although it did not represent a majority of the cases.

As noted above, the involvement of knives formed a distinction between two distinct script tracks. These tracks were identified via the cluster analysis. In other respects, these two tracks were largely similar. Money was mentioned significantly more in track 7b (with knives) compared to track 7a. Victim struggle was significantly more prevalent in track 7a than track 7b. This may point to the effectiveness of weapons in gaining victim compliance.

### ***Summary of cluster characteristics***

Cluster 7 can be interpreted as two distinct tracks relating to robbery via threats and demands. The tracks are distinguished from others by the bold expropriation strategy that is reliant on verbal threats and demands; in a distinct script track these demands are reinforced by the brandishing of knives. Victims relinquish their own property as a consequence of compliance gained through fear.

Wider environments appear to be secluded and lacking in surveillance/guardianship and/or are locations where victims are likely to be accessing or carrying cash. The property stolen was most often money, although mobile phone theft was also characteristic of these tracks.

---

further away the two verbs are from each other in the description, the harder it becomes to link them through semi-automated processes. Although the dictionary building phase can appear quite pedantic there are clear benefits from successfully distinguishing as many ambiguities as possible. However, even with an imperfect dictionary these issues can be identified with careful interpretation of a narrowed down subset of results.

Cluster 8 Involving Taxis

Cluster 8 was mapped to a robbery track relating to the victimisation of taxi drivers. A potential track involving offences committed by taxis drivers was also identified.

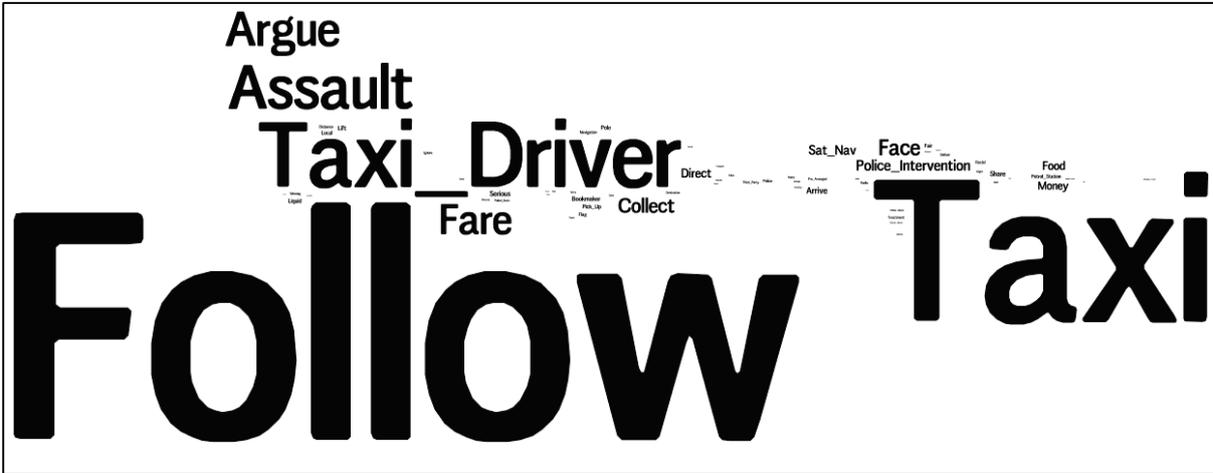


Figure 8-17 Significant tokens in cluster 8. (font size is relative to chi-square score).

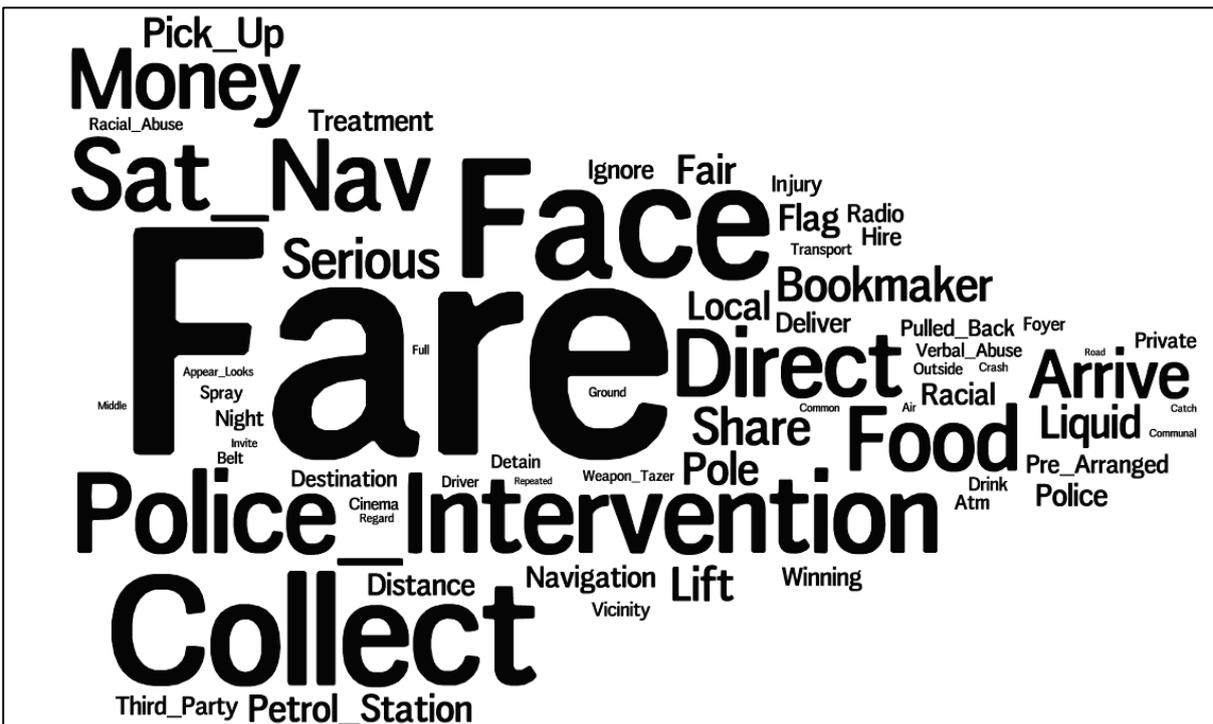


Figure 8-18 Significant tokens in cluster 8. top 5 tokens removed (font size is relative to chi-square score).

**Table 8-11 Examples of MO descriptions assigned to cluster 8: involving taxis. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

Cluster Membership Score			
	A) High	B) Average	C) Low
i	<p>offender <b>taxi_driver punched</b> ip to <b>face causing</b> ip to fall to <b>floor</b> offender then went through ips pockets whilst holding him down with force <b>stole</b> items and made_off in his <b>taxi road</b></p> <p>CMS = 55%</p>	<p>between stated_times and dates unknown offender has apprached ip and held <b>a piece of wood</b> to her <b>face</b> and <b>demanded</b> her <b>money</b> ip has ran_off into the <b>road</b> and <b>flagged a taxi</b> down got in and fled home unknown offender has made_off in unknown direction. <b>road</b></p> <p>CMS = 27%</p>	<p>victim was on the side of [xxxx] <b>resvoir night fishing for</b> carp when he got into <b>a verbal_altercation</b> with <b>a</b> youth. this youth later returned with accomplices who <b>assaulted</b> ip who jumped into <b>resevoir</b>. as this occured his carp <b>fishing</b> equipment was <b>stolenby</b> offender. towpath</p> <p>CMS = 8%</p>
ii	<p>u/k <b>male</b> has called <b>taxi</b> to <b>an</b> address whilst in <b>taxi</b> has <b>directed</b> it down <b>a</b> dark small <b>road</b> then tried to <b>strangle</b> the <b>taxi_driver</b>with his seatbelt. <b>taxi_driver</b> has struggled <b>free</b> u/k has got out and threatened <b>taxi_driver</b> with <b>a</b> blade although it wasn't seen <b>bydriver</b>. u/k <b>male</b> has then <b>punched</b> and <b>kicked</b> <b>taxi_driver</b> into <b>unconsciousness</b> and <b>stolen</b> property from the <b>taxi</b> <b>before</b> making good his escape <b>on_foot</b> towards [name of] <b>road</b>. <b>road</b></p> <p>CMS = 54%</p>	<p>b. m. t uk offender has requested <b>a taxi</b> on <b>arrival</b> ip saw <b>4 males</b> coming out_of 150 [road name] rd and a further <b>2 males</b> standing <b>outside</b>148 [road name] rd. ip drove all <b>6</b> offenders to a location on [road name] st believed <b>pak</b> supermarket. <b>4</b> offenders alight <b>taxi</b> and whilst ip is moving seat to allow <b>for</b> <b>2</b> offenders to get_out of <b>taxi</b> ip is <b>attacked</b>. offenders <b>demanded money</b> and ip has between #110 #120 taken from him including his <b>sat_nav</b>. ip <b>during subsequent attack lands</b> on <b>ground</b> and is <b>kicked</b> to the back and his hea <b>road</b></p> <p>CMS = 27%</p>	<p>offender has approached ip in alleyway near_to offence_location and <b>barged</b> past him. the opinion of the ip on <b>purpose</b> made good escape towards [road name] ip <b>states</b> that he was holding his <b>bus_pass</b> at time_of_offence that the offender <b>barged</b> into him on <b>causing</b> him to <b>drop</b> his <b>bus_pass</b> that the offender <b>picked_up</b> the <b>bus_pass</b> and ran_off it with it. the ip <b>picked_up</b> the empty <b>buspass</b> holder <b>a short distance</b> away. <b>road</b></p> <p>CMS = 6%</p>
iii	<p>ip <b>private_hire</b> <b>taxi_driver</b> has <b>picked_up</b> the three offenders in [xxxx] <b>street</b> [xxxx] they have <b>directed</b> him to <b>an</b> alleyway off [xxxx] <b>street</b> [xxxx] once at this location the offenders have <b>robbed</b> and <b>assaulted</b> ip one offender has <b>restrained</b> bictim whilst other offenders have <b>struck</b> him the the <b>face</b> and forcably searched him <b>stealing</b> #200 in <b>cash</b> and <b>a sat_nav</b> device offenders made_off <b>on_foot</b> <b>road</b></p> <p>CMS = 50%</p>	<p>'offender <b>followed</b> ip out_of <b>a pub</b> ([xxxx]) asked ip <b>for money</b> <b>hit</b> him in the <b>face causing a thick lip</b> and <b>stole money</b>. ip was intoxicated at time_of_offence when later visited has refused to make <b>complaine</b> or give statement regarding incident <b>states</b> offender has since apologised and returned <b>money</b> "<b>road</b></p> <p>CMS = 27%</p>	<p>"unknown offender has approached ip placed his hands_on ip's <b>hips</b> and jumped and down shouting oi oi oi""this laster <b>for ten seconds</b>and was <b>repeated five</b> minutes_later. offender made good escape <b>by</b> unknown means."""</p> <p>CMS = 6%</p>

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

Cluster 8 is a smaller cluster containing only 313 cases that relate to offences involving *taxi\_drivers*. Significant tokens related to taxis were *taxi*, *taxi driver*, *fare and collect*. The majority of offences, covered by this cluster, related to a track which victimised taxi drivers, a smaller number can be mapped to a track related to offending **by** taxi drivers (as appears to be the case in example Ai in Table 8-11 above). Distinguishing the roles and actions of different actors is one of the tasks that is difficult in computer-aided context analysis, and these two tracks were not distinguished through sub-clusters, therefore, determining the number of taxi drivers that were offenders/victims required inspection of the original MO descriptions. Example Bi in Table 8-11 above, shows an additional, but infrequent, use of taxi in this cluster; that is as a mode of escape.

The most significant token in this cluster is *follow*. Analysis of word sequences show that the most common successors of *follow* are *argument* and *verbal*. Therefore, in this script *follow* relates to the order of events rather than a verb (e.g. following an argument).<sup>49</sup>

Given the nature of offences, it is unsurprising that the most significant location/environment for the offences is *road*. It is notable that *direct* is significant, as this relates to the offenders directing the taxi driver to another location. This demonstrates a possibility, for offenders to direct the driver to an environment more conducive for theft. This appears to be the case in examples Aii and Aiii in Table 8.11, although, it is not known whether the offenders were planning to commit a theft at this stage in all cases.

The most significant targets of theft were *money* but also satellite navigation equipment (*sat\_nav*). The transfer of property involved a high level of engagement between offender and victim (perhaps necessitated by the proximity of being in a vehicle); while tokens relating to more innocuous verbal communication were not significant, *arguments* were, and theft often followed *arguments* about a *fare*. Physical force was used to gain victim compliance and transfer property.

This taxi victim track is, therefore, characterised predominantly by the nature of its victims. Transfer of property was bold, and often physical, 95% of cases in this cluster were classified as robbery of personal property. *Assault* and *injury* were significant token for this script. Weapons were not characteristic of this script. Tokens relating to victim resistance were not significant in this cluster.

---

<sup>49</sup> This highlights an area for improved disambiguation in a revised coding dictionary. The results of the current analysis can be used to guide and speed up the disambiguation process.

### Summary of cluster characteristics

This track was characterised, predominantly, by the nature of its victims. Although a potential alternative track was identified that was characterised by the nature of offenders. Transfer of property was bold, physical and violent. *Assault* and *injury* were also a significant token for this script.

### Cluster 9 Vehicle related crimes

Cluster 9 can be mapped to three robbery tracks, carjacking, thefts from persons in cars and thefts of cars from dwellings.

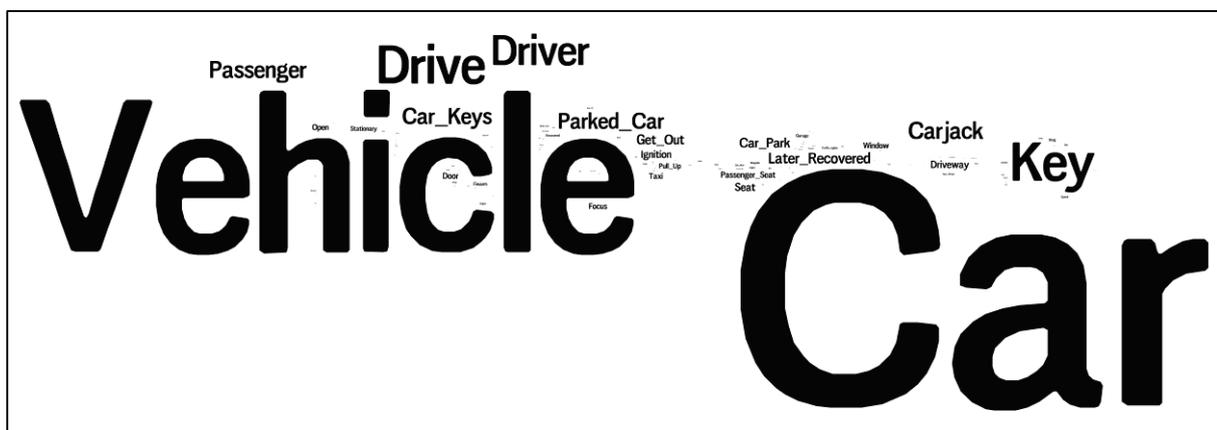


Figure 8-19 Significant tokens in cluster 9. (font size is relative to chi-square score).

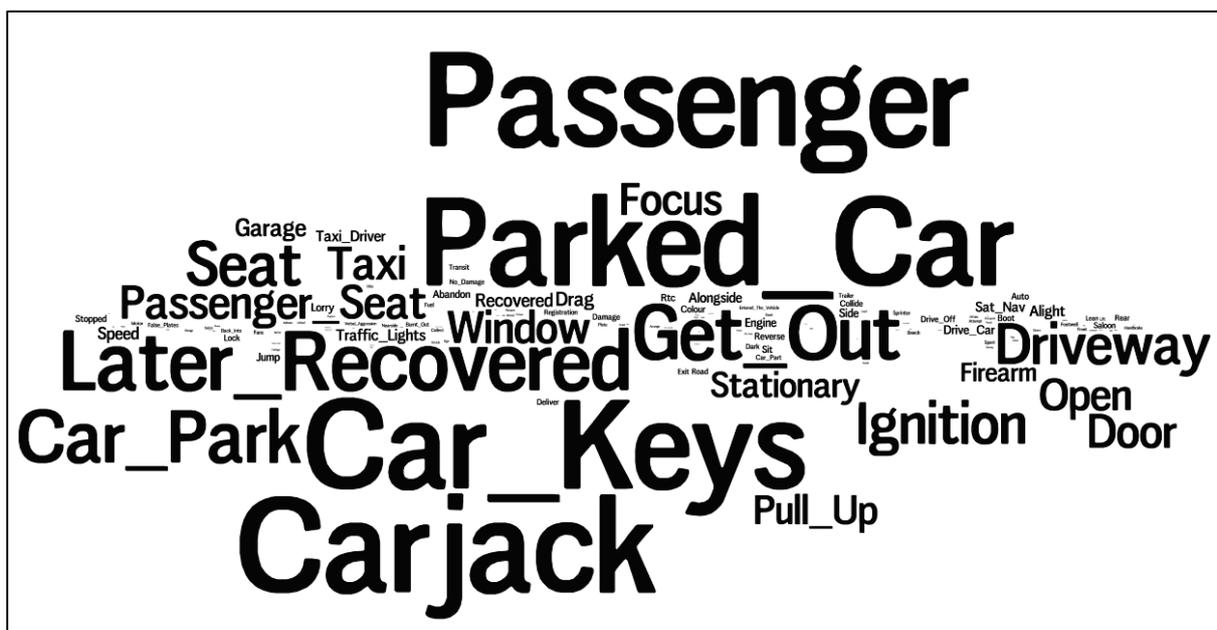


Figure 8-20 Significant tokens in cluster 9. with the top 5 tokens removed.

**Table 8-12 Examples of MO descriptions assigned to cluster 9: vehicle related thefts. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

Cluster Membership Score			
	A) High	B) Average	C) Low
i	<p><b>car_jacking</b> between times stated x2 offenders have approached ip in <b>vehicle</b> silver <b>ford focus</b> [vrm] <b>opened passenger sidedoor</b> and threatened_ip stealing <b>car</b> and ips wallet offenders have made_off in <b>vehicle road</b></p> <p>CMS=64%</p>	<p>offender unknown to ip has got into <b>passenger side</b> of ips <b>vehicle</b> a blue <b>toyota yaris</b> [vrm] whilst <b>stationary</b> at traffic <b>lights</b> offender has produced a knife and instructed ip to give him money offender has <b>directed</b> ip to a cashpoint and fear that the offender will harm the ip the ip has withdrawn money from cashpoint and handed it to offender offender has then made good escape on_foot <b>road</b></p> <p>CMS=27%</p>	<p>unknown offender identifying himself a [name] has approached ip on the <b>car_park</b> to the raer of [name of public house name][name of <b>street</b>] and has taken #20. 00 2 x #10. 00 notes from ip without using <b>force</b> offender has then gone <b>back_into</b> [name of street] and has later been arrested <b>car_park</b></p> <p>CMS=5%</p>
ii	<p><b>car_jacking</b> 2 unknown offenders have approached ip's <b>vehicle</b> opened the <b>driver's door</b> and forced occupant out. both offenders entered <b>vehicle</b> and made good escape in ip's <b>vehicle</b> a black <b>ford focus</b> [name of] <b>road</b></p> <p>CMS=64%</p>	<p>the <b>driver</b> was asleep in the <b>driver's seat</b> of his <b>audi_a4</b> ( green in <b>colour</b> ) [vrmx] and between material_times some persons u/k have <b>opened</b> the <b>door</b> and stolen the <b>keys</b> from the <b>ignition</b> and his mobile_phone taken from his jacket <b>outside</b> pocket. no_force used against ip to facilitate the theft. x_ref with [crime number] theft of <b>motor veh. road</b></p> <p>CMS=27%</p>	<p>ip was standing_by motor_bike when known offender pushed ip out_of the way got onto motor_bike and <b>drove</b> off direction u/k <b>road</b></p> <p>CMS=4%</p>
iii	<p><b>car_jacking</b> bmt offender ( s ) <b>pulled</b> ip from his <b>vehicle</b> silver <b>honda civic</b> type [vrm] whilst he ( ip ) was <b>parking</b> the <b>vehicle</b> on [name of] <b>road</b>. offenders have punched him several times to head causing a graze to his left cheekbone ips <b>vehicle</b> then stolen and it made_off up [name of] <b>road</b> towards [name of]. <b>road</b></p> <p>CMS=60%</p>	<p>"at times &amp; location stated u/k offender ( s ) have approached ips <b>vehicle</b> whilst ip was <b>waiting</b> at traffic <b>lights</b>. u/k offender ( s ) have<b>opened drivers door</b> &amp; <b>attempted</b> to strike ip. u/k offender ( s ) have <b>thrown</b> meat_cleaver blade 8 long at ip. ip <b>drove</b> off in fear_of_being_attacked and in_fear of having <b>vehicle</b> stolen."""<b>road</b></p> <p>CMS=27%</p>	<p>whilst <b>driving</b> to the bank ip has_been verbally abused and threatened by known offenders namely [name] and [name] causing ip harassment alarm and distress bank</p> <p>CMS=3%</p>

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

Cluster 9 contains 2318 cases and is dominated by tokens relating to vehicles, notably: *car, vehicle, driver, keys, ignition*. Wider environments for offences across the whole cluster included *car\_park* and *road* but also dwelling related tokens including *driveway, dwelling, garage*. The cluster included two sub-clusters, the first of which (9a) included carjackings and thefts from persons seated within cars and the second (9b) related to thefts from cars at dwellings. Common across all tracks in this cluster were bold theft strategies, this reflects the crime classifications of the cases in these sub-cluster, 90% of which were robberies of personal property. There were significant levels of *demands* in both sub-clusters including *get\_out*. However, *assault* and *injury* are not significant. *Weapons* were significant, including *firearms* (but not knives), the absence of injury related tokens again suggests that weapons were used to reinforce threats but not to assault victims. Resistance to the theft from victims did not occur to a significant level within this cluster. Again this may be related to the level of threat and the use of weapons.

Although grouped within the same sub-cluster, the offences in cluster 9a could be mapped to two distinct tracks with different primary targets. The first of these related to a carjacking script. The frequency and significance of *Carjack*, in this cluster, is aided by the fact that the police have made attempts to flag these crimes by including this code at the beginning of MO text descriptions<sup>50</sup>. In the *carjacking* track, the car has a dual status as both the primary target of the offence and target enclosure. The car presents a barrier that must be breached to remove the victim and gain the property. The significance of tokens such as *damage* and *drag\_from* highlight the forcefulness with which enclosures are breached in this script.

Not all crimes in cluster 9a were carjackings; crimes also included thefts of items from persons, particularly *passengers* seated within vehicles. For this track, the car represents the enclosure but not the primary target. The victim does not need to be removed from the vehicle to obtain the target property and the enclosure can be breached through an open *window* or an unlocked *door*. Significant tokens in this track indicated that opportunities to breach the target enclosure are made possible when the car is *parked* or *stationary* such as at *traffic\_lights*.

The second sub-cluster, 9b, included offences relating to an additional script. The significant overuse of the token *dwelling* along with *entry, upstairs* and *knock* these tokens appeared to relate to a script

---

<sup>50</sup> Although applied relatively comprehensively not all carjackings are identified with this tag.

which included cases where offenders followed, or waited for the victim to reach home, at which point the victim is robbed outside their home and car keys/cars stolen, as below.

IP DRIVING A BLUE R32 VOLKSWAGON GOLF [VRM] RETURNS HOME AND PARKS THE CAR. IP LEAVES THE CAR HEARS A NOISE AND LOOKS BEHIND AND SEE'S 6 MEN RUNNING TOWARDS HIM. IP RUNS TRIES TO MAKE A NOISE BUT IS GRABBED BEATEN DRAGGED TO ENTRANCE TO FLATS DRAGGED UPTO IP'S FLAT. TAKEN TO KITCHEN WHERE KEYS FOR\_A BLACK RENAULT MEGANE [VRM] ARE DEMANDED AND HANDED\_OVER. IP THEN FORCED INTO A BEDROOM AND OFFENDERS LEAVE. BOTH CARS STOLEN. 6 MEN WEARING BALACLAVAS. VEHICLE [VRM] FOUND [NAME OF] LIBRARY CAR CAR\_PARK

In another permutation, offenders gained entry to a dwelling via threats or force whereupon a robbery against the person took place resulting in the theft of a vehicle, for example:

BETWEEN 1515 HRS AND 1230 HRS TWO UNKNOWN MALES HAVE ENTERED ADDRESS VIA INSECURE FRONT DOOR ARMED WITH AN AXE AND KITCHEN\_KNIFE. THREATS WERE MADE TO THE 3 OCCUPANTS DEMANDING CASH AND CAR\_KEYS. ALL 3 OCCUPANTS WERE THEN TIED\_UP BEFORE AN UNTIDY SEARCH WAS MADE OF THE ADDRESS AND CASH JEWELLERY AND CAR\_KEYS STOLEN. OCCUPANTS GRANDSONS CARSTOLEN BMW [VRM] CAR KEY CRIME VEHICLE\_RECOVERED [NAME OF ROAD AND TOWN] ON AND RECOVERED TO [RECOVERY FIRM]. DWELLING

Sub-cluster 9b shares features with the burglary dwelling cluster (involving forced entry to dwelling) and the sub-cluster 9d where victims are followed home.

Although it is not directly evident from the data this appears to be one of the few scripts with evidence of planning or prior preparations, the MO descriptions mention waiting for targets and the offenders appear to be acting on at least some information regarding in which dwellings desirable cars are located. It should be remembered that these are only indicators of planning and there is no concrete evidence of planning in the data.

### ***Summary of cluster characteristics***

This cluster was characterised by the inclusion of cars, but the role that the vehicle takes in the crimes varies according to different offence types. The offences in this cluster appear to fall into three distinct tracks: 1) carjackings 2) theft from persons in motor vehicles and 3) thefts of vehicles from dwellings, the cluster analysis only directly identified two groups and did not distinguish between carjackings and thefts from people within vehicles. All tracks had in common the use of bold strategies characterised by verbal aggression and demands. This cluster was unusual in



**Table 8-13 Examples of MO descriptions assigned to cluster 10: robbery involving young people. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

Cluster Membership Score			
	A) High	B) Average	C) Low
i	<p>between 1835 and 1845 hrs on 19/10/08 ip was with 2 friends when he was approached by 3 male offenders. they have shown a knife to the ip and said give us what's in your pocket. the ip then handed_over his mobile_phone. the offenders have then made_off. park</p> <p>CMS = 61%</p>	<p>"ip was stood by gates of [xxxx] park with friends saw a group of approx 20 ic3 youths walking towards them fear something may occur ip began to walk_away but chased by group along [xxxx] road. ip was slowest runner and stopped and was surrounded by group. offender said give_me your phone"". ip then handed_over his phone and offender plus group made good thier escape."""park</p> <p>CMS = 32%</p>	<p>at material_times and location u/k offender jogged towards ip and asked_for watch then opened jacket and exposed half of akitchen_knife. property taken offender made_off by jogging toward direction in which he came. park</p> <p>CMS = 10%</p>
ii	<p>"whilst in park with friends ip was approached by a group of males. ip was asked_for_time. offender then said give_me your phone"". ip then forcibly searched and property stolen. park</p> <p>CMS = 61%</p>	<p>unknown male_offender has approached ip pushed ip and demanded the ip's phone. ip tells offender he hasn't got one offender has then searched ip's pockets. no items stolen. offender has then made good his escape in unknown direction park</p> <p>CMS = 32%</p>	<p>between 1600 and 1630 hours on thursday 9th july 2009 unknown offender approached ip asked to look at his i_pod. offender then reached into his pocket and flashed a metal object at ip. ip then handed_over his i pod</p> <p>CMS = 10%</p>
iii	<p>ip and friends have_been approached by two unknown offenders. one offender has said this is what it is park</p> <p>CMS = 60%</p>	<p>"offender/s uk have approached ip in street offender 2 has said to ip what have you got_for_me""then searched ips pockets offender 1 stood in_front_of ip when offender 2 found nothing walked_off offender one then notices ips gold chain and said""he's got a chain""ip fearfor his own safety let offender remove it offenders tan off and joined a group of other males""road</p> <p>CMS = 61%</p>	<p>[xxxx] school pupil was surrounded by a group of possibly holte school pupils as he went to local shops in lunch hour. 2 x members of group grabbed and searched ip who surrendered his wallet and contents fear_violence would result from non_compliance. road</p> <p>CMS = 10%</p>

Words highlighted in red are significantly characteristic of this cluster (Chi-Square  $p < 0.001$ )

Cluster 10 includes 1262 cases which are characterised largely by the offences' wider environment in which the crime takes place and the nature of the victims and offenders. The tokens that are significant in the cluster suggest that the offences are robberies, this is confirmed by the classification, in the police data, of 95% of these cases as robberies of person property. The offences described in this cluster appeared to involve threats made towards victims. Frequently (and significantly) victims were with *friends* or *with company* at the time of the offence, a feature quite distinctive to this cluster. The token *group* was significant to this cluster, and on inspection of relevant descriptions, relates both to groups of offenders and groups accompanying the victim(s). *Youth* was also significant and although used mainly to describe offenders it is clear that in this cluster both the offenders and victims were young.

Two sub-clusters were identified by the analysis. These sub-clusters can be regarded as separate tracks that are distinguishable by different wider environments. Cluster 10a was characterised by robberies in *parks* which was the most significant token in cluster 10 overall and cluster 10a specifically. The track identified in cluster 10b was characterised by environments including *footpaths*, *alleyways*, *sports fields* and *schools*. Again, the location tag included in the MO descriptions provides a considerable level of consistency in the recording of environments which helps the clustering process.

In both sub-clusters, a number of significant tokens relate to the victims' actions prior to the offence, these include *listening to music* (which might be indicative of inattention and also of the presence of a valuable target for theft) and *playing football*, *walk* and *bench* (i.e. sitting on). There was an explicit *approach* by the offenders towards the victim, which was reported and recorded, *cornered* was significant in this cluster, indicating a key mechanism through which offenders exert control over victims. The opportunity to corner the victim may be enhanced by offenders operating in groups. Overall, the offenders are generally seen, as these are offences are bold rather than relying on secrecy. However, *bush* was mentioned in a very small but significant number of cases (22) which points to an element of surprise.

In both tracks, the most common property stolen was *mobile\_phones* (most frequent) and *music\_player* (most significant), but *money* was also significant in these tracks. The target enclosure was frequently the victims clothing and this was breached through different permutations in that some cases involved victims handing over property while others involve the offender searching the clothing of victims to *retrieve*. The token *search* was significant within both tracks, searching for clothing property on a victim may be a more feasible strategy when there is more than one offender, one to search one to 'look out'.

There was evidence that approach often began with non-threatening conversations indicated by significant tokens such as *verbal* and *cigarette* (typically asking for a cigarette). However, much of the communication was threatening, demonstrated by tokens such as *threat*, *hand\_over* and *demand*. Offences in this track rarely appeared to escalate to violence, and *assault* and *injury* are not significant in this cluster. Tokens referring to victim resistance were also not significant in this cluster. The threats made, were often reinforced with the weapons, particularly *knives*, suggesting similarities to cluster 7, Threats. However, the key differences relate to the environments in which offences take place (Parks vs. Roads/Alleyways), the property stolen (music players and mobile phones vs. money) and the age of offenders and victims. The park offences were more likely to involve groups of young people, and for the offender and victims to be in the company of friends at the time of the offence.

### ***Summary of cluster characteristics***

Cases in this cluster were characterised by the environment, and parks were predominant. Offences were frequently (but not exclusively) conducted by multiple offenders and could involve victims who were with company at the time of the offence. The nature of the offence was characterised by an explicit approach and attempts to dominate the victim, first by cornering and then with threats. The exchange of property is, therefore, bold but it is threatening rather than physical. Although weapons are shown, they are not used. There is some physical contact between offender and victim when offenders search clothing for property. Property stolen is most frequently small valuable electronics i.e. phones and music players. Victims rarely resist or refuse the offenders demands.

### ***Cluster 11 Theft from bags at Licensed Premises***

**Cluster 11 was mapped to one script track relating to theft from bags at licensed premises.**

Cluster 11 contained 689 cases which relate to a track for thefts from bags (usually a handbag), in the main, while the victims are attending licensed premises or similar venues. The police data classified the majority (93%) of these cases as thefts from the person. The tokens which characterised this track (seen in Figures 8.23 and 8.24), relate to the environment in which the offences take place, *licensed premise*, *club* but also *market*. These locations were commonly described as *crowded* and *busy*. The track was further characterised by tokens relating to the target enclosure, *bag* and the elements of this enclosure, namely *zip* and *clasp*, which the offender must *undo* in order to access the desired property. *Body* is significantly overused in this cluster; in this group tokens relating to body (predominantly *shoulder*) refer to the location of bags at the time of the offence i.e. over the victim's shoulder. Given the focus on handbags, it is probable that the

majority of these victims are female, however, the token *female* was not significant in this cluster. Additional victim variables, not available to this analysis would need to be inspected to confirm or refute this possibility.



Figure 8-23 Significant tokens in cluster 11. (font size is relative to chi-square score).



Figure 8-24 Significant tokens in cluster 11. top 5 tokens removed

Several tokens relate to the victim’s activity before the offence, including *dance\_floor* and *stood\_at\_bar*. It is not clear whether this description of activity is causally linked to the offence. Despite the environment in which offences take place, tokens relating to intoxication are not overly represented in this cluster.

In this track, transfer of property was secretive with *dip* and *remove* being significant tokens associated with the surreptitious transfer of property, these methods stand in contrast to grasping, grabbing or snatching in clusters 2 and 5. The secrecy of the offending is further confirmed through the significance of tokens such as *offender not seen*, *unseen*, *nothing felt*, *no evidence*, *notice* and *missing*. In keeping with the secretive nature of this script, there was no verbal or physical contact between victim and offender, other than instances of a *light\_touch* and a *tug*. Property stolen relates to items commonly stored in hand bags while on a night out including *phones*, *wallets* and *digital cameras*.

The cluster analysis did not split cluster 11 into any sub-clusters, and no alternative tracks were identified within this cluster.

**Table 8-14 Examples of MO descriptions assigned to cluster 11: theft from the person in licensed premises. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

				<b>Cluster Membership Score</b>		
				<b>A) High</b>	<b>B) Average</b>	<b>C) Low</b>
i	<p>ip <b>took purse</b> out_of <b>handbag</b> to pay to enter location then returned the <b>purse</b> to <b>bag</b> and <b>zipped</b> up top_of <b>bag</b> ip then kept <b>bag</b> on <b>shoulder</b> using <b>strap</b> whilst <b>dancing</b> ip went to pay for <b>drinks</b> <b>opened bag</b> and <b>found property stolen</b></p> <p><b>nothing_seen_or_felt</b> bag was <b>zipped</b> up <b>public_house</b> licensed premises</p> <p>CMS= 65%</p>	<p>offender unknown approached ip who was <b>stood</b> in a <b>crowd</b> <b>watching</b> a <b>concert</b> offender <b>opened</b> ips <b>bag</b> and <b>removed mobile_phone</b> and other <b>items</b> without ip <b>noticing</b> and made good escape. <b>nightclub</b></p> <p>CMS= 29%</p>	<p>on the <b>relevent date</b> <b>between</b> the <b>relevent</b> times ip has_been attacked <b>inside</b> take_away <b>resturant</b> by one named offender and two unknown offenders and whilst <b>being</b> attacked a <b>bag</b> <b>containing</b> cds has_been snatched from him and <b>stolen</b> <b>restaurant/cafe</b></p> <p>CMS= 6%</p>			
ii	<p><b>bmt</b> offender has <b>stolen</b> ips <b>mobile_phone</b> from her <b>handbag</b> whilst <b>handbag</b> was over ips <b>shoulder</b>. offender made good thier escape. <b>public_house</b> licensed premises</p> <p>CMS= 59%</p>	<p>unknown person has <b>taken items</b> from a <b>shoulder_bag</b> whilst the ip was <b>wearing</b> it person has mge <b>direction</b> unknown. <b>nightclub</b></p> <p>CMS= 29%</p>	<p>the ip in this matter attended [xxxx] <b>nightclub</b> at the material_time and date. the ip <b>states</b> he was <b>heavily intoxicated</b> and sat down on chair in the <b>club</b> and fell asleep. the ip <b>states</b> he looked <b>at</b> his <b>watch</b> prior_to his sleep and was <b>aware</b> it was 0030hrs. when the ip awoke <b>at</b> 0115hrs his <b>watch</b> had gone. the ip believed the <b>watch</b> was <b>stolen</b> as it was <b>buckled</b> up and <b>could</b> not have fallen off. the ip <b>saw no</b> offender/s. <b>nightclub</b></p> <p>CMS= 7%</p>			
iii	<p><b>at</b> relevant_times and location ip has_had her <b>bag</b> on her <b>shoulder</b> and her <b>mobile_phone</b> has_been <b>stolen</b> from the <b>bag</b>. offenders unknown made good escape in unknown <b>direction</b>. <b>public_house</b> licensed premises</p> <p>CMS= 59%</p>	<p><b>between</b> the material_times stated ip has_been <b>inside</b> the <b>club</b> with her <b>handbag</b> on her <b>shoulder</b> <b>at_all</b> times. <b>at</b> approx 0030hrs ip has used her <b>purse</b> to buy a <b>drink</b> then has put the <b>purse</b> <b>back_into</b> her <b>bag</b> which may have_been left <b>unzipped</b> <b>accidentally</b>. approx 1 <b>hour</b> later ip has <b>noticed</b> her <b>purse</b> and <b>camera</b> have_been <b>removed</b> from the <b>bag</b>. <b>nothing</b> <b>witnessed</b> by the ip and <b>no</b> offender/s identified.<b>clubsocial</b></p> <p>CMS=29%</p>	<p>ip has_been be_friended by offender in [xxxx] <b>public_house</b> they have then left together and has_been shown west [city sights] offender has then ran away ip has then <b>realised</b> his wallet has_been <b>taken</b> from his trouser pocket ip is visiting us citizen on a business trip returning to <b>usa</b> on 27. 10. 07 road</p> <p>CMS= 7%</p>			

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

### ***Summary of cluster characteristics***

The offences in this cluster relate to a clear script track that is distinguished principally by the environment in which the offence takes place. There is a clear target enclosure, a bag, and this is breached by the offender surreptitiously via dipping. The offender does not make a visible approach to the victim, rather he or she takes advantage of the crowded nature of the environment. There is, therefore, no verbal or physical engagement between victim and offender during these offences. There are clear similarities between this track and the tracks 5a and 5b relating to thefts from bags in shops.

### ***Cluster 12 Robbery with Assault***

**Cluster 12 was mapped to three robbery of personal property tracks: robbery with assault and injury, robbery with assault and no injury and robbery with domestic assault.**

Cluster 12 contained 2843 cases which are examples of the most physical and violent offences identified in the data. The textual analysis identified that the cases were robbery offences and this is confirmed by the police data with 98% of these cases classified as robberies of personal property. The cluster analysis identified three sub-clusters which could be considered as separate robbery tracks: 12a robbery with assault and injury, 12b robbery with assault and no injury and 12c robbery with domestic assault. However clusters 12a and 12b could be regarded as the same track but with different outcomes.

These offences typically occurred on *roads, alleyways, footpaths*. The exception to this was track 12c offences which were more likely to occur in *dwelling*s. Although there is little to indicate the vulnerability of the targeted victims, victims were frequently *walking* prior to the offence and assaults did not take place in crowded areas. *Behind* was significant to the cluster overall, as was *bushes*, both providing steps that help offenders to avoid advance signs of the attack. As with the script 10 robbery of and by young people, this script includes significant mentions of *groups*, although in this script, the groups are more likely to be offenders. Tokens explaining whether victims were alone with friends or with company were not significant. In contrast to script 10, *youth* was not a significant token for this script.



Figure 8-25 Significant tokens in cluster 12. (font size is relative to chi-square score).



Figure 8-26 Significant tokens in cluster 12. top five tokens removed

**Table 8-15 Examples of MO descriptions assigned to cluster 12: robbery with assault. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

		Cluster Membership Score		
		A) High	B) Average	C) Low
i	<p>unknown offenders have <b>approached</b> ip <b>male</b> 1 offender has <b>punched</b> ip to <b>face causing</b> ip to <b>fall to floor</b> other offenders approx 10 in total have <b>hit</b> ip to <b>face and body by kicking and punching</b> him offenders have then <b>searched</b> ip whilst on <b>floor stealing property</b> from his <b>pockets</b> offenders made good thier escape ip has <b>bruising and swelling</b> to his <b>face and cut to his mouth and nose road</b></p> <p>CMS=82%</p>	<p>ip <b>walking</b> to work in <b>early_hours</b> when 4 youths started to follow him. as he <b>walked</b> along <b>footpath</b> leading to main <b>road</b> one youth <b>hit</b> him on the side of <b>head and</b> said give us your phone or i will <b>stick_you</b>. ip <b>ran_off</b> towards home. <b>youths</b> gave <b>chase and kicked and punched</b> him further before making_off. i/p <b>phoned_police</b> had <b>bruising swelling</b> above left <b>eye</b> and around temple <b>public footpath</b></p> <p>CMS=32%</p>	<p>at material_time named offenders have asked to have a go_on ips moped. ip has declined and offenders have stated they will take it anyway otherwise they will batter ip. moped was recovered by ip later the same day. damage caused to front of moped and parts removed. <b>road</b></p> <p>CMS=6%</p>	
ii	<p>offenders unknown have <b>approached</b> ip. have <b>kicked</b> him to the <b>floor</b>. have taken his <b>mobile_phone</b>. offenders have then <b>punched</b> ip to the <b>face causing</b> a <b>swollen lip</b>. offenders have then <b>ran_off</b> along [xxxx] <b>road towards [xxxx] road. road</b></p> <p>CMS=79%</p>	<p>offenders have <b>approached male</b> ip <b>from_rear</b> snatched his <b>mobile_phone</b> from <b>jacket pocket</b>. ip ip <b>swung arm</b> round and <b>knocked</b> phone out_of offenders hand onto <b>floor</b>. ip has picked phone off the <b>floor</b> and offender stated 'give us your phone '. at this both offenders have <b>begun to punch</b> ip to his <b>face</b> eventually made_off with connector from phone. <b>road</b></p> <p>CMS=32%</p>	<p>offenders whilst in process of committing an <b>armed</b> robbery [crime ref] were interrupted by ip who happened to <b>walk</b> into petrol_station. ip then grabbed <b>searched and property stolen</b> from him prior_to offenders <b>maaking</b> off on_foot commercial</p> <p>CMS=6%</p>	
iii	<p>between stated_times 2 u/k offenders have <b>punched and kicked</b> ip to the <b>face and stolen</b> items_listed from his person. ip has <b>swellingand bruising</b> to his left <b>eye. road</b></p> <p>CMS=78%</p>	<p>offender has <b>approached</b> ip asked him for <b>money</b> then grabbed ip at the <b>back_of</b> the <b>neck</b> he threatened him until ip handed_over 50 pence offender then asked him for #1 <b>ip_refused</b> so offender pulled ip to the <b>floor hit</b> ip over the <b>head</b> with his hand offender then threatened_ip <b>again before kicking</b> him on the <b>backside road</b></p> <p>CMS=32%</p>	<p>pupil known to teacher went to ask teacher a question as teacher was giving explanation pupil removed <b>wallet</b> from outside <b>right_handpocket</b> of teacher's <b>jacket</b>. pupil returned to his seat at rear of class teacher saw pupil with his <b>wallet</b> pupil <b>threw wallet</b> to <b>floor</b>. <b>nomoney stolen. educational</b></p> <p>CMS=4%</p>	

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

In sub-cluster 12c, victims and offenders were known to one another indicated by tokens such as *family* and *partner*.

The token indicating victim resistance, *victim refuse*, was significant in this cluster. As discussed earlier, the significance of this individual token does not explain the sequential consequences of resistance. However, the significant inclusion of *without\_provocation* suggests *assault* was not only used as a reaction to victim resistance but was also used to initiate engagement with the victim. Further, *threat*, *verbal aggression* and *verbal* are not significant in this cluster, indicating that victims do not receive forewarning before the assault. *Demand* is, however, significant although offenders also conducted forceful searches of *pockets* and *clothing* to obtain goods.

All of the tracks were characterised by the violent and forceful nature in which property was transferred from victim to offender. The most significant token for each of the sub-clusters was *assault*. Tokens relating to the violent nature of the assault, including *face*, *cut*, *injury*, *ground*, *treatment* and *repeated* are also significantly overused in tracks 12 a and 12b of this cluster. In the whole database the attendance of an ambulance was only mentioned in 20 cases but 70% of these cases were located in this cluster. The use of weapons was not statistically significant in track 12a but significant for tracks 12b and 12c. The property stolen covered the most of the items commonly stolen across the dataset including *money*, *wallets* and *mobile\_phones*.

### ***Summary of cluster characteristics***

This cluster was characterised by the physical nature of the offences. The offenders met their objective of transferring property into their ownership with a bold, physical and violent strategy. The offences did not rely on stealth or secrecy, although, they appear to incorporate speed and surprise into their strategy moving quickly from the initial approach to physical contact. There is little detail on the approach itself, although it is frequently from behind to provide a degree of surprise. There is little in the way of verbal communication other than direct demands. No specific resources or skills are employed. The property stolen is not specific and includes all valuable items normally carried on the person.

There are no specific provocations, victims are attacked without provocation, they are commonly in non-crowded areas, and are not accompanied by acquaintances. Frequently reported environments for this type of robbery are roads, footpaths and alleyways. A distinct track relates to offences committed against people known to the offender, these offences tended to take place in domestic settings, although, the methods of robbery were otherwise broadly similar.



**Table 8-16 Examples of MO descriptions assigned to cluster 13: mobile phone theft. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

Cluster Membership Score			
	A) High	B) Average	C)
i	u/k offenders approached ip and asked to use mobile_phone have then snatched phone and made good escape road  CMS=80%	"offender was <b>approached</b> from side by 3 unknown males who <b>asked</b> if he had a <b>phone</b> . i. p. <b>said</b> yes and are you going to rob me?"". offender 1 then <b>took</b> i. p's <b>phone but returned sim_card</b> . offender 2 <b>said</b> "alleyway  CMS=31%	known offender <b>took</b> computer_game belonging to another and <b>sold</b> it. ip states game <b>taken out_of his hand header</b> amended <b>road</b>  CMS=9%
ii	unknown offender approached ip from_behind whilst ip was walking and on his mobile_phone. offender snatched phone from ip's handand ran_off making good escape. road  CMS=67%	<b>bmt</b> the ip was <b>walking</b> alone at the offence_location when 2x u/k offenders have <b>approached</b> him and <b>snatched</b> the ips <b>phone out_of hishand</b> offenders have then <b>made_off on_foot along [xxxx] road</b> towards [xxxx] <b>road</b> where the ip has then <b>lost_sight</b> of them <b>road</b>  CMS=31%	known offender entered premises and it appears <b>following</b> a dispute <b>took</b> ip's <b>mobile_phone from_his_hand</b> . full details are yet to be obtained ip is not in the country for another 7 days ( 55766 ) <b>community</b> centre  CMS=9%
iii	bet_mat_times offender approached ip asked to see his mobile_phone then snatched it from him and made good thier escape road  CMS=67%	<b>bmt</b> whilst ip was <b>walking_home</b> an unknown male <b>approached</b> ip on [xxxx] <b>lane</b> . male <b>grabbed</b> ip with one <b>hand</b> stating'have you got a <b>phone</b> '. ip struggled <b>escaping</b> males <b>grip</b> and ran <b>away</b> . ip's <b>school</b> blazer had_been ripped. offender made good <b>escape road</b>  CMS=31%	between the stated_times and dates whilst the ip""has_been <b>walking along</b> billesley common " <b>public footpath</b> "  CMS=8%

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

Cluster 13 contains 1355 cases which relate to the theft, of mobile phones from the person. This cluster is largely characterised by tokens relating to the property stolen i.e. *mobile phone* and *phone*. Importantly, the token *hand* and *using\_phone* is highly significant, indicating, along with *phone\_call* and *victim\_holding\_property*, *listening\_to\_music* and perhaps *message*, that the target property for the theft was on view and accessible to the offender at the time of the offence. Conversations between offenders and victims were significantly over-represented in this cluster as a whole but this did not occur in all offences. In fact, the cluster analysis identified two sub-clusters, separated by the

significance of communication in cluster 13a and the lack of significance for this token in cluster 13b. The clusters involving communication included a higher proportion of robberies (69%) compared to the cluster without communication in which theft from the person was more prevalent (66%) but both clusters appear to include a mixture of the two crime types. When communication occurred it included direct requests relating to mobile phones, but also, general low risk *conversation* including asking for the *time* and asking to *borrow* a phone and direct demands. Low risk conversations provide a means for the offender to gain proximity to the victim and perhaps, to inspect property and judge the resilience of the victim. Cluster 13a also included significantly more occurrences of *pocket* and *search* (there may be a need to engage the victim if the property is hidden). Cases in cluster 13b were more likely to involve the *victim holding* the property. Cluster 13b also included offences committed by offenders cycling or running *past* the victim, again this strategy may be dependent on the property being visible and accessible rather than enclosed.

In both tracks the property is exchanged via speed and or surprise as indicated by the significant token *grab. No\_threat* and *no\_force* were both significant for this cluster highlighting that the offences relied first, on gaining proximity and then, on speed rather than on direct physical force. These offences commonly took place on *footpaths* and *roads*. Victims were commonly *walking* and *alone* at the time of the offence. Note that *sim\_card* was a significant token, some MOs describe the offender removing the sim card before taking the phone.

Victims did not offer any resistance to the crimes (to any significant level) but from the offenders' point of view, the surprise nature of the grab phase of the offence probably serves to reduce this possibility.

### ***Summary of cluster characteristics***

The cases in this cluster are characterised by the property that is stolen. However, mobile phone theft is also a feature of a number of other clusters. The factors that distinguish this cluster are the combination of initial approach and conversation, followed by an unexpected grab. There was, therefore, verbal engagement between the offender and victim; this can be a fairly innocuous conversation that poses low risk to the offender initially but also included direct demands. There is little physical contact and no force used by the offender. Offences take place in public locations but where victims are likely to be alone.

Cluster 14 Public Transport

Cluster 14 mapped to four tracks relating to different types of public transport location

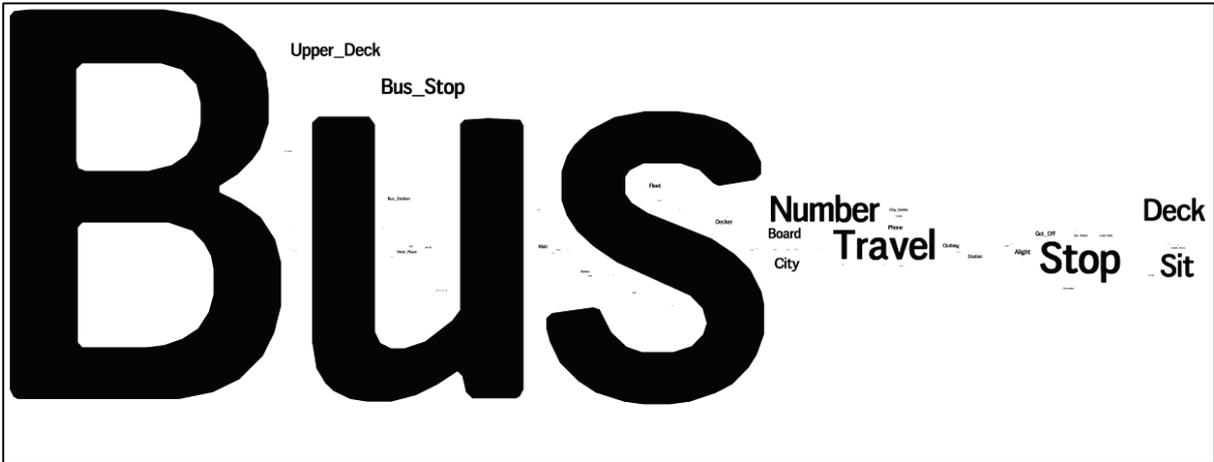


Figure 8-29 Significant tokens in cluster 14 (font size is relative to chi-square score)

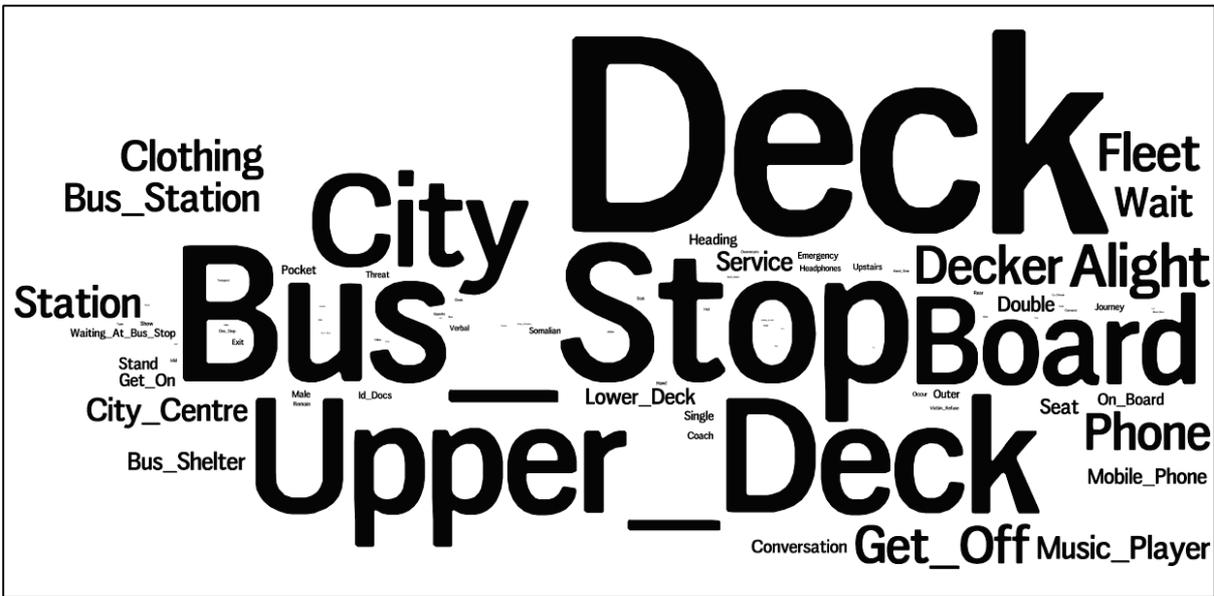


Figure 8-30 Significant tokens in cluster 14. top 5 tokens removed.

**Table 8-17 Examples of MO descriptions assigned to cluster 14: thefts on buses and at bus stops. Includes MO descriptions that had (a) high (b) average and (c) low cluster membership scores.**

Cluster Membership Score			
	A) High	B) Average	C) Low
i	<p>ip with <b>friend travelling</b> on <b>upper_deck</b> of <b>no. [XX] bus</b>. only other persons on same <b>deck</b> are the 7 offenders. one offender hits ip on head another offender then <b>demands</b> ip's <b>phone</b> which ip <b>hands_over</b> to him. offenders <b>also</b> take ip's <b>friends</b> bag taking personal_property. ip has discomfort to left_side of head. <b>bus</b></p> <p>CMS=65%</p>	<p>bmt at offence_location ip was <b>travelling</b> on a <b>bus number [XX]</b>. ip was <b>sat</b> in <b>upper_deck</b> of the <b>bus</b> he was approached by three <b>male</b> offenders one of the offenders <b>clenched</b> a fist causing ip to <b>fear</b> that he would_be assaulted. offenders have stolen the ips black samsung <b>mobile_phone</b> offenders made good escape off the <b>bus bus</b> road</p> <p>CMS=32%</p>	<p>between the times stated <b>at</b> the offence_location u/k offender approached ip &amp; had a verbal_altercation about the ownership of a pram which ip had offender stated the pram the ip had is hers therefore the offender has the rights to take the pram off ip offender took the pram ip had to take the child out_of the pram who was distressed <b>bus stop</b></p> <p>CMS=7%</p>
ii	<p>ip has_been <b>travelling</b> into <b>city</b> on [XX] <b>bus</b> route [XX] ip has_been <b>sat</b> on <b>upper_deck</b> as offender <b>boarded bus</b>. offender snatched <b>phone blackberry from_ips_hand</b> and made good his escape off <b>bus</b>. no_threats_or_force used. <b>bus</b> road</p> <p>CMS=64%</p>	<p>between material_times u/k offender has punched the ip whilst <b>travelling</b> on the <b>number [XX] bus upper_deck</b> before snatching his bag making good escape. <b>Bus</b></p> <p>CMS=32%</p>	<p>details created from btp log [ref no] ip had his <b>wallet</b> removed from his <b>rear pocket at</b> location did not realise until he arrived at [name of ] railway <b>station</b> is convinced it was taken <b>at</b> location stated reported to btp originally limited details available from thier log. <b>no suspects/cctv</b>. ip was visiting his sister's grave in [name of]. road</p> <p>CMS=5%</p>
iii	<p><b>bmt</b> ip was <b>sat at rear</b> of <b>upper_deck</b> of [XX] <b>bus travelling outercity</b>. when approached by x4 somalio <b>males</b> who punched ip to left ear and stole <b>mobile_phone</b> from ips left <b>jacket pocket</b>. offenders then <b>alighted</b> the <b>bus</b> and made good thier escape on_foot <b>no</b> direction_of_travel <b>bus</b></p> <p>CMS=63%</p>	<p><b>bmt</b> ip was <b>followed</b> off [XX] <b>bus by</b> offenders. <b>demands</b> made for property. ip was <b>searched</b> and money stolen. offenders made good thier escape x_ref with [XX] /<b>road</b></p> <p>CMS=31%</p>	<p>the offender in this matter has <b>sat</b> alongside the ip who was workingat a computer <b>station</b> within the public <b>library</b>. the defendant has then removed the ips <b>passport</b> from his <b>jacket pocket</b> which was draped over the back_of his chair offender made good his escape.<b>library</b></p> <p>CMS=4%</p>

Words highlighted in red are significantly characteristic of this cluster (Chi-Square p<0.001)

Cluster 14 contains 2294 cases which are characterised, principally, by the environment in which the offences take place. The token *bus* is by far the most significant token in this cluster, along with related tokens including *bus\_stop*, *deck*, *upper\_deck*, *bus\_station*.<sup>51</sup> The bus environment, which could also be considered an enclosure, appears to have provided the necessary proximity between offender and victim in most cases, although the variable *cornered* is significant as a means of approach/control. The cases in this cluster involved a significant level of verbal engagement between offender and victim, this included low risk (from the offender's point of view) *conversation*, as well as *threats* and *demands*. The extent of threats in this cluster is illustrated by the significance of the token *fear*. Property stolen in these cases includes *mobile\_phones*, *music\_players*, *wallets* and *money*. The results of the cluster analysis, alone, are unclear on how property is transferred in this cluster. Tokens such as *dip*, *grab*, *remove* and *search* are not significant. Rather, the token *hand\_over* is significant, suggesting that victims are forced by threats to surrender their property. The token *victim\_resist* is statistically significant. Physical contact is not described to a significant degree across the cases in this cluster, although the extracts in Table 8.17 show that such crimes are included in this cluster. Although *knife* is not significant *stab* is, as seen previously this may relate to a threat rather than the use of a knife.

This cluster was divided into a total of five sub-clusters which can be interpreted as four different tracks and a further group acting almost as an 'other' categories with public transport offences that were a poor fit elsewhere. The four script tracks are largely split by the wider environment in which the offence occurred with the methods of offending remain similar across the sub-sets.

Track 14a, with 123 cases, was characterised by offences that occurred while the victim is waiting at the *bus\_station*. Track 14b is similar, but involves victims who are waiting at a *bus\_stop*. Track 14c involves offences that occur on the *bus*. Track 14b and 14c were the largest sub-clusters in this group with 414 and 1467 cases, respectively. Track 14d principally involves offences that occur when the victim is *walking\_away* from the bus. However, the ambiguity of terms such as *follow* means that this cluster also includes cases where, for example, the victim has surrendered property *following* demands made by the offender. Sub-group 14e is a rather incoherent cluster which appears to simply contain offences that were poor fits elsewhere. This includes some offences where the

---

<sup>51</sup> Note that offences committed on trains and at train stations are reported to British Transport Police and therefore do not feature in this data.

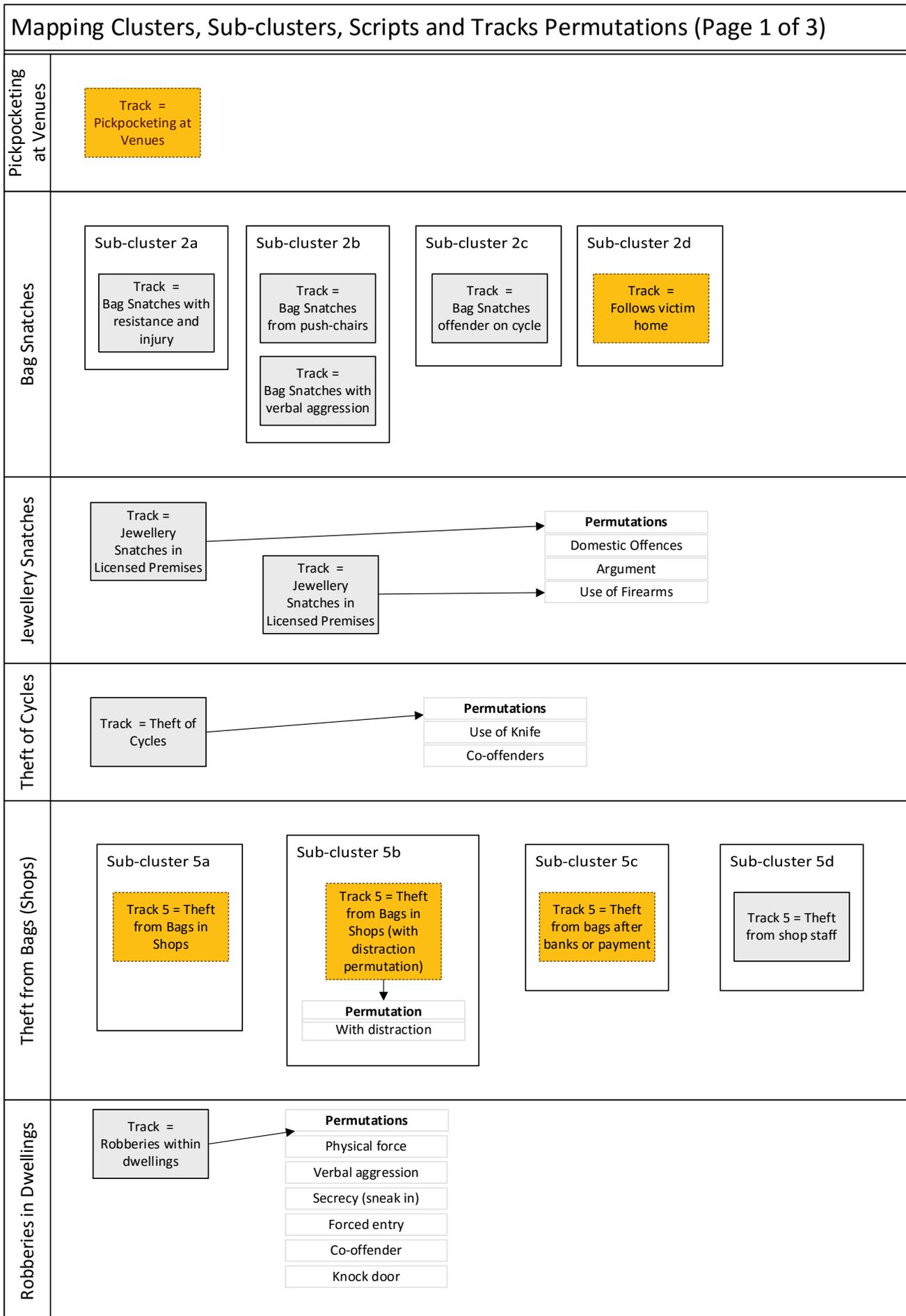
offender was believed to have made his escape on a bus, but overall, there is generally no fixed pattern to the offences included in this group.

### ***Summary of cluster characteristics***

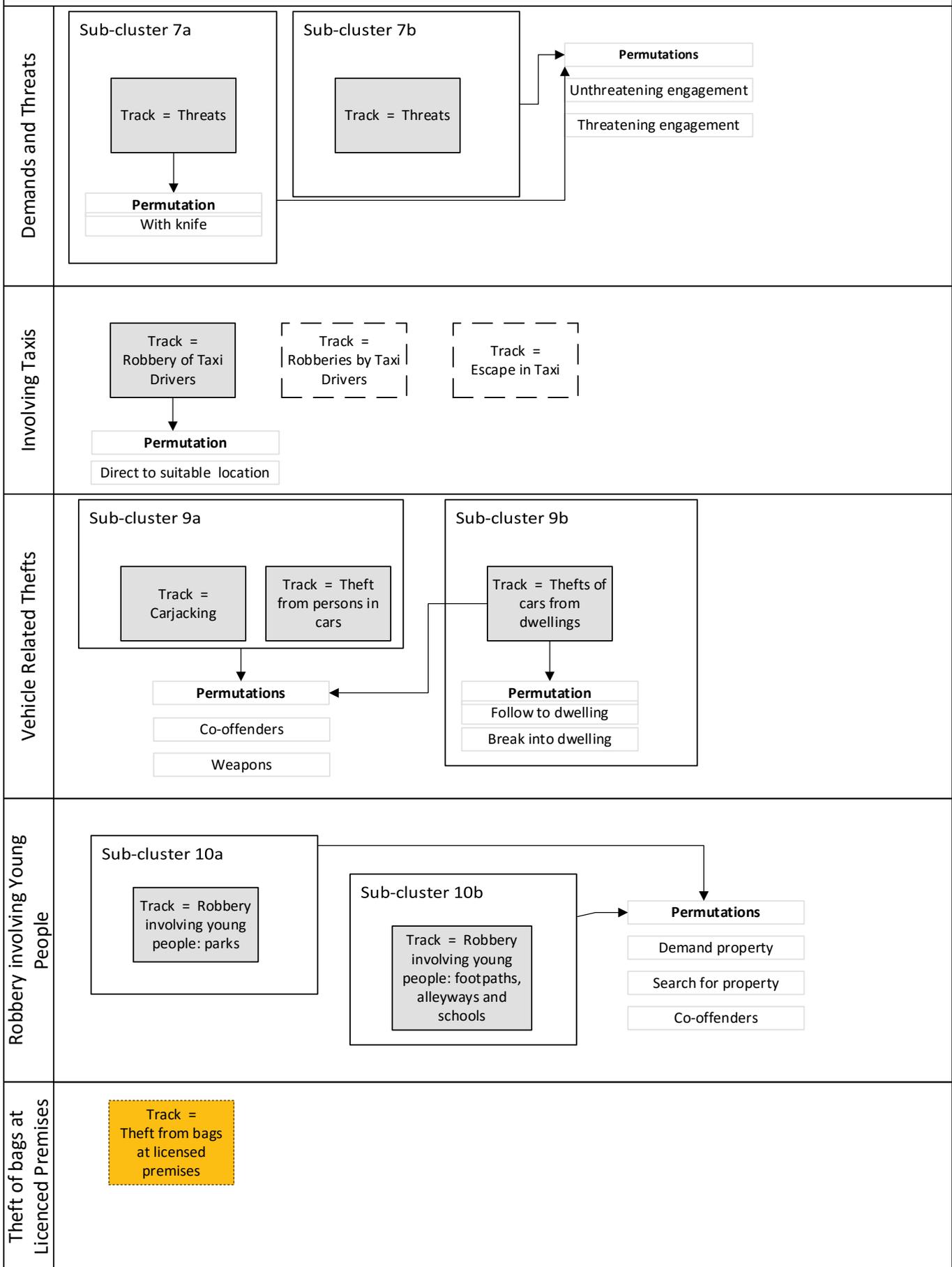
This cluster is characterised by the environments, public transport, in which the offences takes place and the individual tracks are characterised by sub-sets of that environment. There appears to be some variation in methods with individual transfer methods (e.g. snatch thefts, secretive dips, and thefts following assaults) not showing statistical significance. In other words, all of these types of crime were present in the cluster but none dominated. There is evidence that offenders engage victims in verbal communication to a significant degree, ranging from low risk communication to threats and demands. Physical contact was not statistically significant. The location code which provides a basic categorisation of the offence site was clearly instrumental in helping to define and demarcated these sub-groups. However, tokens in the MO descriptions themselves frequently repeat this location information (particularly for this cluster). This provides some confidence that these offences would be successfully clustered in data where a location code was not available.

The above section has described the characteristic features of clusters and where possible related mapped them to script tracks and permutations. The following chapter will review the clusters, draw comparisons between them and provide an assessment of their validity.

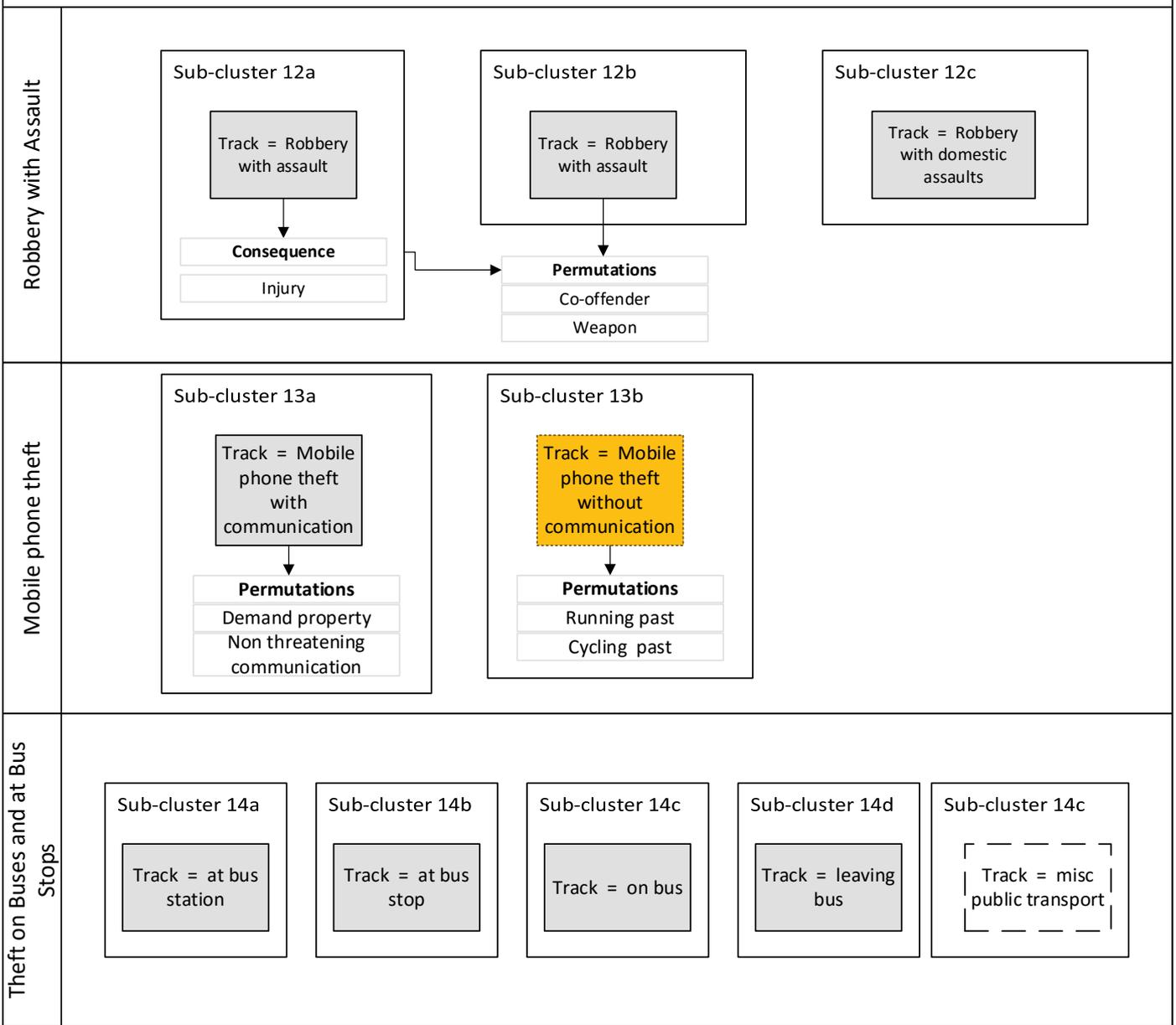
Figure 8-31 Mapping clusters and sub-clusters to scripts, tracks and permutations



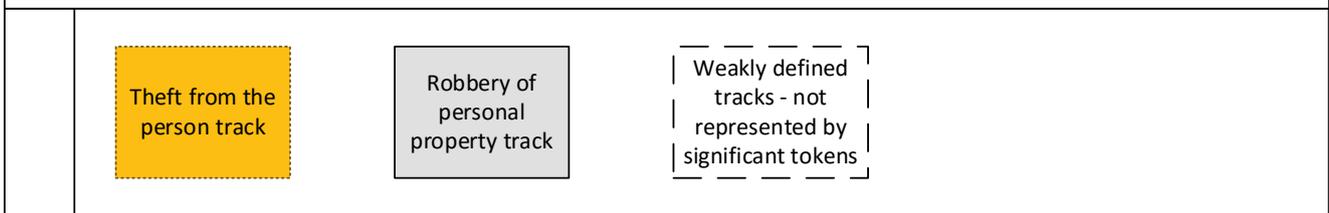
# Mapping Clusters, Sub-clusters, Scripts and Tracks Permutations (Page 2 of 3)



# Mapping Clusters, Sub-clusters, Scripts and Tracks Permutations (Page 3 of 3)



## Key



## Chapter 9 Reviewing and Validating the Clusters

### *9.1 Comparing Scripts and Tracks*

Based on the analysis in Chapter 8, a script template was created which depicts all of the empirically observed tracks and permutations in the data (Figure 1 Appendix 2). This diagram is repeated for each cluster (Figures 1-14 Appendix 2), with each version highlighting the significant elements that were relevant for that specific cluster.

On the left-hand-side of each diagram, in blue boxes, the acts and scenes of the script are listed broken down into specific steps identified in tracks and permutations. For each cluster the relevant elements are highlighted in yellow (i.e. those that occurred in the cluster to a statistically significant degree). Each scene is summarised, in italic text, by the goal of scene completion – a milestone at which a scene can be considered successfully completed. At the right-hand-side of the diagram, the lilac boxes represent CCO elements that shape offenders' perceptions and actions in the script. In recognition of the fact that CCO elements are dynamic and can change, particularly as a result of actions taken by actors in the situation, these elements include some CCO features that are changed by actions of crime preventers and crime promoters in the situation.

Inspection of these diagrams enables the comparison of different scripts and tracks. It was also intended that comparison of these diagrams would help identify scenes that are essential (although there may be multiple ways of completing them), in contrast with scenes that are optional depending on the track taken (in a similar manner to the analysis of script stages in the literature, presented in Table 5.1). However, from the data alone it is not possible to ascertain whether the absence of details about a script scene shows that this stage has not occurred and is perhaps not necessary or whether there simply is no available information about this scene.

#### **9.1.1 Pre-and-post script conditions**

The data analysis provides little information to shed light on the preparatory stages of the offences. In particular there is no information relating to the nature of readiness or the formation of, or the influence on the situation of, goals, motivations or morals. This is not to say that goal formation did not take place. Previous studies have highlighted that some types of victims may be regarded as more acceptable to certain offenders than others. This perception may be based on the victims' characteristics (Deakin et al., 2007) or their actions (Jacobs & Wright, 2008; Topalli et al., 2002;

Topalli & Wright, 2004). For some crime types, such as hate crime<sup>52</sup>, it would be possible to use MO descriptions to infer from the data the moral or normative decisions that shape factors such as target selection. This type of information has not been revealed in this analysis. The data analysis did not provide very much information on the planning and preparations undertaken by offenders. This does not necessarily mean that planning did not take place, but that the data does not provide any information on this either way. Existing research evidence suggests that preparation and planning may be minimal (Copes & Cherbonneau, 2006; Feeney, 1972; Hochstetler, 2001; Jacobs, 2012) although the extent and nature of preparations may be shaped by factors such as the nature of the intended victim (Copes et al., 2012; Jacobs & Wright, 2010).

Information relating to planning was restricted to details of resources utilised to complete certain script tracks. It is not clear whether offenders deliberately 'geared up' with these resources or whether availability of these resources was part and parcel of offender routine activities. There is little in the literature to support an explicit 'get resources' scene. The only resources that were identified from the MO descriptions were weapons, vehicles and co-offenders, but tracks were available that did not rely on these assets. The option to offend with others was significant for six tracks, theft of cycles (4), robberies within dwellings (6), carjackings (9a), theft of cars from dwellings (9b,) robberies involving young people (10) and robberies with assault (12a and 12b). There is little to suggest that there was a planned 'meet with co-offenders' stage in preparation for the offences and for most script tracks, again this is supported by the literature (Hochstetler, Copes, & DeLisi, 2002) it seems more feasible that offending emerged from a joint recognition of an opportunity while offenders were already in each other's company. One exception might be the vehicle related scripts where evidence of some shared planning is suggested by co-offenders turning up together at a specific dwelling or location of an attractive vehicle. However, this is a logical inference rather than a finding directly extracted from the data.

A similar group of tracks involved the use of weapons to a significant degree: jewellery snatches (3), robberies within dwellings (6), demand and threats (7b there was also a strongly apparent track, 7a that did not utilise weapons), vehicle related offences (9), robberies involving young people (10) robberies with assault (12b). It should be noted that these tracks are all robbery tracks, they also involved a greater level of engagement between offender and victim than the remaining scripts for which weapons were not significant. Although all of the script tracks can be conducted without

---

<sup>52</sup> The author has conducted [unpublished] data analysis on hate crime using same techniques of textual analysis and through dictionary development was able to bring out features relevant to different hate crime types (homophobic, racist etc.)

these resources, it seems that weapons are particularly beneficial when there is direct contact with the victim. Possession of weapons during a robbery script did not always mean direct use of that weapon to physical harm the victim: weapons were frequently used to gain victim compliance via fear (e.g. track 7a threats with knives). It was observed across a number of tracks, (particularly 12a and 12b robberies with assaults, with and without injury) that tracks using weapons to threaten victims were associated with fewer injuries than those without brandishing of weapons. Similarly, there was less evidence of victim refusal or resistance in tracks involving weapons (for example this is apparent in comparisons of cluster 7a with 7b). Again there are parallels here with the literature which point to the use of weapons by offenders to communicate a threat and forestall and reduce the possibility of physical confrontation (Barker et al., 1993; Deakin et al., 2007).

### 9.1.2 Target selection

#### *Search for target*

The analysis provided clear evidence that offenders committed their offences in locations that were conducive to theft, although the specific features of those environments vary. For example, some tracks take advantage of crowded locations while others are more likely to occur in quiet or secluded environments. There is an identifiable distinction between stealth based tracks for which crowded environments, or at least environments with other distractions were more significant compared to tracks that involve aggression or physical contact that were associated with more secluded environments, this is possibly because it would not be possible to escape the notice of other witnesses and potential crime preventers in a busier environment. Other tracks can be identified where the availability of attractive targets appears to have been a factor. In a similar vein to the findings from Bernasco and Block (2011), Bernasco et al. (2013) and Holt and Spencer (2005) such as locations included those where cash transactions take place (5c theft from bags and 8 involving taxis). There is no evidence in the data to either support or refute whether offenders travel to these locations specifically for the purpose of committing theft (i.e. whether the locations are crime generators or crime attractors). In the pickpocketing from entertainment venues track it is feasible that offenders are attracted to venues for the available opportunities and may have made explicit preparations to travel there. In contrast, the association between the wider environment of parks and young people may suggest that offences occur here because this is where both offender and victims<sup>53</sup> spend their time recreationally. As noted in Section 5.2, Jacobs (2010) argued that

---

<sup>53</sup> These individuals can change of course switch roles in different occasions and circumstances.

offenders' lifestyles regularly placed them in situations where opportunities for theft were likely to emerge, but that they also used their skills to identify and capitalise upon opportunities when they arose. Thus the offender is both attracted to the situation due to the potential for crime opportunities, but at the same time the location helps to generate crime opportunities. This seems to blur the crime pattern theory (Brantingham & Brantingham, 1975) distinction between crime attractor and crime generator, which is perhaps a consequence of the static treatment of offenders in that perspective.

The tracks also reveal how offenders manipulate the crime situation to create suitable opportunities, rather than passively react to them. This is evident in the small number of scripts where offenders actively move victims from the point of selection to a location more suitable for the transfer of property. This can be seen in tracks 2d bag snatches following victims home and tracks involving theft of taxi drivers in cluster 8. This is a permutation that has been highlighted in research in to sexual offences (Beauregard et al., 2007).

### *Target selection*

As with the wider environment, features that lead to victim vulnerability can be identified but vary across script tracks and were the result of interactions with the nature of the location and the property that victims were carrying. This supports the assertion that victim attractiveness is situationally defined (Topalli et al., 2015). Specific information on victim characteristics within the MO descriptions is patchy and the analysis would benefit from marrying descriptions with specific data on victim (and offender) demographics. There was evidence of script tracks that were associated more significantly with younger victims (theft of cycles, robberies involving young people) and offenders while other tracks older victims were over-represented (bag snatches including following victims home). Although in some scripts victims were alone, not all tracks were dependent on this, particularly those tracks involving either stealth, surprise, or co-offenders. It is very important to note that although the data provides some insights into the reasons for the selection of offence location and target selection within that location, the data does not reveal anything about *how* the offender makes these selections, an issue highlighted earlier with Bouhana's (2013) critique of situational crime theory.

### **9.1.3 Transfer**

The remaining script act, transfer of property, and its component scenes, approach, engage victim, search for property and transfer, are covered in much greater detail and consequently populated with more information. It is apparent that there were numerous tracks and permutations through

which these scenes were accomplished. Thus, these scenes could be completed using equifinal actions leading to the same end result. These equifinal actions were not completely interchangeable, as the actions available are dependent on the specific situational conditions.

### *Approach*

In all the tracks identified it is necessary for the offender to gain proximity to the victim. In most tracks this involves some form of approach although the victim may not be aware of this. One consequence of the often secretive nature of approach is the frequency with which it is not detailed in the descriptions.

In several script tracks the environment facilitates proximity. In crowded locations (licensed premises and entertainment venues) bumping into other people is not out of the ordinary, and this degree of contact is all that is often necessary for dip or pickpocket offences. In other tracks the offender and victim share an enclosure in the course of routine activities, particularly when taking a bus or a taxi (these enclosures can also make it difficult for a victim to escape).

As noted in Chapter 5, the nature of approach and the nature of transfer are inextricably linked, therefore transfers achieved by stealth involve a less obvious approach. This also relates to the target enclosure, enclosures that can be easily breached without the cooperation of the victim do not require the offender to approach or engage with the victim.

In tracks where securing victim compliance is required (these are predominantly robbery tracks), the methods of approach identified appeared to be designed to reduce the victims' range of options for action. Thus offenders cornered victims, or blocked their path (found in tracks 4, 9, 10, 14) presumably in an attempt to reduce the victim's options to distance themselves from the offender. Other track permutations used surprise: this included the more aggressive tracks but also those that relied on snatching property. In both cases surprise reduces the victim's options for initiating defensive action (Copes et al., 2012). Characteristic approaches of these tracks include jumping out at victims, sometimes from a hiding place (seen in tracks 10, robberies involving young people and 12, robbery with assault) and approaching the victim from behind (2, bag snatches and 12, robbery with assault).

It should be noted that although gaining compliance might be thought of as a step more associated with victim engagements and transfer, the data suggests that offenders are working towards compliance from the earliest stages of the offence. This supports the findings of Jacobs (2012) who

demonstrated the role of approach not only as a step to gain proximity to the victim but crucially also in gaining control over victims.

Other track permutations for approach that provide surprise via speed included running past and cycling past the victim (2, 13b)

### *Engagement*

Engagement is one of the few scenes that is not necessary for all tracks. Where engagement occurred it was kept to a minimum, particularly any contact or communication that precedes attempts to transfer property; again this relates to the benefits of surprise. Thus engagement with the victim was either avoided, restricted to brief but innocuous conversation or was immediately aggressive – blurring any distinction between engagement and transfer. Tracks where engagement tended to be either minimal (such as a bumping into or brushing past the victim) or non-existent included 1, pickpocketing at venues; 2, bag snatches; 5, theft from bags in shops; and 11, theft from bags in licensed premises. Tracks that involved victim contact in the form of innocuous conversation were very common and included pickpocketing at venues (1), this included configurations involving short conversations but also hugging, thefts related to banks and monetary transactions (5c), mobile phone thefts (13), tracks involving public transport (14) and robberies of shop staff (5d). It is notable that some of these tracks require exceptionally close proximity to the victim, particularly where the target enclosure is a person's clothing or their hand. It may be that it is the need to directly invade someone's personal space that is being accomplished using conversations. However, the literature also points to the role of conversation in assessing the victim (Katz, 1988).

The remaining tracks where engagement was immediately aggressive included jewellery snatches (3), thefts of cycles (4), robberies within dwellings (6), demands and threats (7), and carjackings (9a), thefts of vehicles from dwellings (9b), robberies involving young people (10) and robberies with assault (12). Although tracks existed that were initiated with aggression, no tracks were initiated with direct assault; again this parallels the literature (Copes et al., 2012).

The remaining strand of engagement was deception in which the offender aims to conceal their intentions for as long as possible rather than conceal themselves (Luckenbill, 1981). The two main forms of deception were posing as an official (6, robberies within dwellings) and making an apparently legitimate request for the property, such as requesting to borrow a phone (13a, mobile phone thefts with communication). There is nothing in the data to suggest that victims were taken in by these bogus claims, these interactions were still followed by other forms of transfer (snatches, threats etc.), but the claims perhaps are sufficient to temporarily wrong foot the victim.

### *Search for property*

Previous studies have identified different strategies employed by robbers to locate valuable property on the victim: this included offenders demanding that victims turn out pockets or bags to reveal the property and offenders conducting searches of victims' clothing (Smith, 2003). This distinction was also identified in this data, although all robbery tracks included both permutations. Also in line with the literature, there was some indication that searching for property by an offender may be more likely when he/she is operating with co-offenders, although this is not conclusive. In addition to the offences where property was enclosed in the victim's clothing, these two track permutations were also identifiable in track 6 robberies within dwellings, in which offenders either searched the dwelling themselves or demanded that those residents relinquish property.

In many tracks searching for property is not necessary as it is already visible to the offender. These tracks tended to relate to theft from the person scripts but also include robbery scripts. They include tracks that take advantage of situations where the target enclosure is the victim's hand or body (2, bag snatches; 3, Jewellery Snatches; 4, theft of cycles; 13, mobile phone theft) or a pushchair or shopping trolley (2b).

### *Transfer*

The nature of property transfer was inextricably linked to the nature of approach and engagement. Thus the tracks with the least victim contact tended to be completed by dipping or pickpocketing (pickpocketing at venues 1, theft from bags in shops 5 and theft from bags at licensed premises 11). Similarly 'snatching and grabbing' of property was associated with limited contact, the speed and surprise elements of these tracks were also characterised by features such as the offender running, approaching on a cycle or jumping out at the victim (for example in tracks 2 and 13). Although contact with victims in snatch and grab tracks was limited, they did include configurations with combinations of aggression, physical contact and force. Thus, bag snatches included the use of force directed towards the primary target (pulling at and snapping of bag straps<sup>54</sup>) but could also involve directing force towards the victim such as pushing the victim to the ground. There were indications across a number of clusters that force was more likely to be used in snatch and grab tracks when victims offered resistance to the removal of their property, this was the case in tracks 2a, bag

---

<sup>54</sup> Slashing was not identified (Johnson et al., 2010)

snatches with resistance and injury and 3, jewellery snatches. The role of victim resistance highlights that the track permutations adopted in an offence are not entirely of the offender's choosing.

The tracks that were initiated with aggression continued in that vein and transfer was achieved by means of this aggression, which manifested through combinations of demands, threats and physical violence. The success of these aggressive strategies is demonstrated in the significance of (victims' feelings of) 'fear' in these tracks, this token was significant in tracks 4, theft of cycles; 7, demands and threats; 10, robberies involving young people; and 14, thefts on public transport. Not all of these scripts involved physical violence towards the victim, for example, track 7, demands and threats, did not. Where force was used this was a combination of forceful efforts to access property (e.g. 4, theft of cycles) but also as a means of obtaining victim compliance through fear (e.g. track 9a, carjackings). The level of force used was generally greater than that seen with snatch and grab theft tracks.

### *Escape and disposal*

The MO data provides very little insight into the escape and disposal acts of theft from the person and robbery of personal property. This reflects the factors which those reporting and recording offences regard as most important but also shows that the combinations of speed and stealth utilised in these tracks means that this information is often unknown. Details of escape are limited to characteristics such as 'on foot' which was most significant for tracks involving snatch and grab. Escape was mentioned to a significant extent in the vehicle related tracks (9a and 9b) and the theft of cycles track (4). In both of these tracks the target of theft, car or bike, was itself used as a resource for offending – the means of escape. Factors relevant to disposal were only mentioned in the vehicle related crime scripts, and related to the subsequent recovery of stolen vehicles and the use of false registration plates to disguise vehicles.

### **9.1.4 Summary**

The results of the cluster analysis provided support for Cohen et al.'s (1995) classification of expropriative strategies, with examples of cryptic, deceptive, bold and surprising theft strategies. Although the distinctive nature of each strategy is most apparent in the transfer scene, the earlier scenes serve to establish the necessary groundwork for the strategy to be successful. The available strategies are also mutable, providing different routes or permutations, through which they can be performed and have resistant counter strategies embedded with them, including methods of reducing the victim's awareness of the offence occurring and controlling the victim to gain compliance.

There is no evidence in the data to comment on the adaptation of offender strategies to counter competing strategies or the contagion, stimulation or mobility of strategies. The absence of the need for specific resources would aid contagion and the evidence of similar strategies in the findings of studies in multiple countries lends support for their mobility. Previous research has pointed to the expressive nature of theft, in addition to its purely economic value, which provides evidence that the act of theft as well the outcomes can be rewarding for offenders (Shover & Honaker, 1992).

The script tracks identified also reflect procedural characteristics common across all criminal behaviour (Eckblom & Gill, 2015) including foraging, conflict between actors, violent and threatening behaviour, stealth, surprise and deceit. The descriptions also reveal collaborative behaviour and imply trust between co-offenders, although in isolation they do not show the co-evolution of tactics and strategies.

## *9.2 Evaluating and Validating the Clusters*

Cluster analysis is an exploratory tool that aims to identify groups of similar cases and in doing so reflect patterns in the real world. Interpreting the meaning of statistically-derived clusters and deciding whether that interpretation genuinely reflects real-world patterns has always been a problematic issue, and this study is no exception. If instructed to do so, a cluster analysis algorithm will find clusters in randomly distributed cases. It is therefore important to ensure that clusters derived in an analysis do not simply represent random variations and to assess whether clusters are a fair presentation of the world which they are meant to be classifying. While this relationship with real world patterns can never be fully ascertained there are a number of ways in through which greater confidence can be assigned to the validity of clusters. These include:

- Comparing cluster findings to the findings from previous research and existing theories.
- Comparing the cluster results with other variables.
- Comparing the stability of cluster results over time.

This remainder of this chapter will adopt these strategies to explore the validity of the clusters identified in Chapter 8. Section 9.1 has already identify similarities between the findings from the cluster analysis and the literature providing a theoretical support for the validity of the clusters. Section 9.2.1 will compare the cluster results to the official police classifications of crime types found in the raw data. Section 9.2.2 will assess the stability of the clusters by exploring the change in their shape and size over time, as recommended by Clatworthy et al. (2005). Although some changes in the nature and size of clusters is expected, change that is greater than expected or difficult to explain, given knowledge of the domain of interest, places the validity of clusters under question.

### 9.2.1 Comparing the cluster results with other variables

One criteria to judge whether the patterns in cluster solutions reflect real world patterns is to compare the clusters to variables, other than those that were used in their composition, with which one would expect to find a relationship. In the example below, each cluster is compared to crime classifications found in the original police data. Given that the nature of transfer, (i.e. forceful or otherwise), is only one dimension of theft on which theft from the person and robbery of personal property scripts have been shown to vary, it should not automatically be assumed that there would be correspondence between clusters and crime types. However, in the interpretations of clusters above, the nature of transfer was frequently a defining characteristics of a cluster or sub-cluster. Each of the clusters were interpreted *before* cross-tabulating the cases with the crime classifications. It is, therefore, useful to consider whether the textual analysis of clusters has organised cases in a way that seems logical given the classification of crimes in the data. If aspects of force and/ or threats of force, as per the legal definition of robbery, are significant within a cluster, it would be concerning to discover that the majority of offences in that cluster were thefts from the person.

In Table 9-1 there appears to be a high degree of agreement between characteristics of the clusters and the police crime classifications. Clusters that are characterised by force or threats of force (i.e. Cluster 3, 6, 7a and 7b, 8, 9, 10, 12) contain an overwhelming majority of cases classified as robbery. In contrast, theft from bags in licensed premises is predominantly theft from the person, corroborating the textual analysis. Given the nature of cluster 1 pickpocketing at venues, the proportion of theft from the person offences in the cluster, although still the majority, seems surprisingly low. Given the significant characteristics of the cluster, there does not seem to be a logical reason to find robberies within this cluster and this suggests that there some of the cases assigned here are poorly matched.

The fact that the analysis suggested that there may be two distinct tracks within sub-cluster 2a may explain why there is a 70/30 split between the two crime categories, rather than a more marked majority for either type. Thefts from pushchairs are more likely to be thefts from the person while thefts involving verbal aggression, if threatening, can be classified as robbery.

Cluster 5d, being comprised mainly of robberies, presents different proportions of the crime types to clusters 5a, 5b and 5c which are dominated by thefts from the person. This difference provides support for the interpretation that cluster 5d constitutes a different and unrelated track to the others in this cluster.

**Table 9-1 A cross tabulation of cluster solutions and official crime classifications**

	Cluster	No. Cases	% Theft from the person	% Robberies of Personal Property
1	Pickpocketing at Venues	692	88	12
2	Bag snatches	2392	32	68
2a	<i>including resistance and injury</i>	110	18	82
2b	<i>from pushchairs and including verbal aggression</i>	144	31	69
2c	<i>offender on cycle</i>	1583	14	86
2d	<i>including following victim home</i>	555	65	35
3	Jewellery Snatches	682	6	94
4	Theft of Cycles	757	11	89
5	Theft from Bags (Shops)	2185	74	36
5a	<i>theft from bags (Shops)</i>	1388	91	9
5b	<i>theft from bags (with distract)</i>	285	91	9
5c	<i>including banks, cards and withdrawals</i>	424	91	9
5d	<i>thefts from shop staff</i>	88	22	78
6	Robberies within Dwellings	1355	4	96
7	Demands and Threats	1472	3	97
7a	<i>not involving knives</i>	1010	1	99
7b	<i>involving knives</i>	462	5	95
8	Involving taxis	313	5	95
9	Vehicle related crimes	2318	10	90
9a	<i>Carjacking and thefts from cars</i>	1645	10	90
9b	<i>Car thefts from dwellings</i>	673	11	89
10	Robbery involving young people	1262	6	94
10a	<i>in parks</i>	378	5	95
10b	<i>on footpaths, alleyways and in schools</i>	884	6	94
11	Theft from bags at Licensed Premises	689	93	7
12	Robbery with Assault	2842	2	98
12a	<i>with assault and injury</i>	837	3	97
12b	<i>with assault no injury</i>	204	2	98
12c	<i>with assault domestic</i>	1801	2	98
13	Mobile phone theft	1355	48	52
13a	<i>with communication</i>	989	31	69
13b	<i>without communication</i>	366	66	34
14	Thefts on buses and at bus stops	2294	36	74
14a	<i>at the bus station</i>	133	48	52
14b	<i>at the bus stop</i>	445	22	78
14c	<i>on the bus</i>	1578	23	77
14d	<i>leaving the bus</i>	22	26	74
14e	<i>misc</i>	116	12	88
	<b>Total</b>	<b>20608</b>		

The two mobile phone theft sub-clusters (13a and 13b) present almost opposing proportions of the two crime types. This reflects the greater use of force and threats when there is communication between the offender and the victim (13a). In the two sub-clusters, the majority crime type was not as dominant as seen in other clusters (where the majority was frequently 90% or higher) highlighting that there remains some crossover of transfer types between the two sub-clusters.

Overall, Table 9-1 provides support for the cluster solution. With the exception of cluster 1, there is little to contradict the textual analysis. It should be noted that poor matches between text interpretations and crime classifications may be the result of problems in the police data rather than problems in the cluster. The approach of cross-tabulating results could be used to guide the audit process and to point, either to specific clusters or specific cases where matches are contrary to expectations. These areas of the data could then form the focus of auditing.

### *K=2 Cluster solution*

Another way to inspect the correspondence between clusters and crime classifications would be to review the results of a two cluster solution ( $K=2$ ). This can be done by re-inspecting the dendrogram Figure 8.1 in Section 8.2. As noted in the methodology, the hierarchical structure that results from BKM cluster analysis allows the comparison of solutions with a greater or smaller number of clusters. Figure 8.1, reproduced below as Figure 9.1, demonstrates how cases would be organised in a solution with only two clusters, referred to as cluster A and cluster B. The clusters from the original analysis that have been subsumed under cluster A are highlighted in blue, those that have been subsumed under cluster B are highlighted in red. Cross referencing the dendrogram with Table 9.1 above reveals that, with the exception of the original cluster 1 (pickpocketing at venues) all of the original clusters that fall with the cluster A contained a high proportion of robbery against the person offences. In contrast, the clusters that were subsumed under cluster B include those that contain a high proportion of theft from the person offences (e.g. Cluster 11 theft from bags at licensed premises) – but, significantly, also clusters that included a high proportion of robbery of personal property (e.g. Cluster 9 Vehicle related crimes.) The proportion of crime types in cluster A and B is summarised in Table 9.2. This table shows that a  $K=2$  solution would have produced two clusters, cluster A with 13024 cases and cluster B with 7584. Cluster A is dominated by cases classified as robbery of personal property (84% of cases). However the second cluster contains a mixture of theft from the person (45%) and robbery of personal property cases (55%) cases. The nature of transfer, which is reflected in the crime classification is clearly influential in distinguishing the two clusters, but attributes relating to transfer are not the only factor splitting the clusters.

Table 9-2 Clusters in a two cluster (k=2) solution

	Number of cases	% Theft from the person	% Robbery personal property
Cluster A	13024	16	84
Cluster B	7584	45	55
<b>Total</b>	<b>20608</b>	<b>100</b>	<b>100</b>

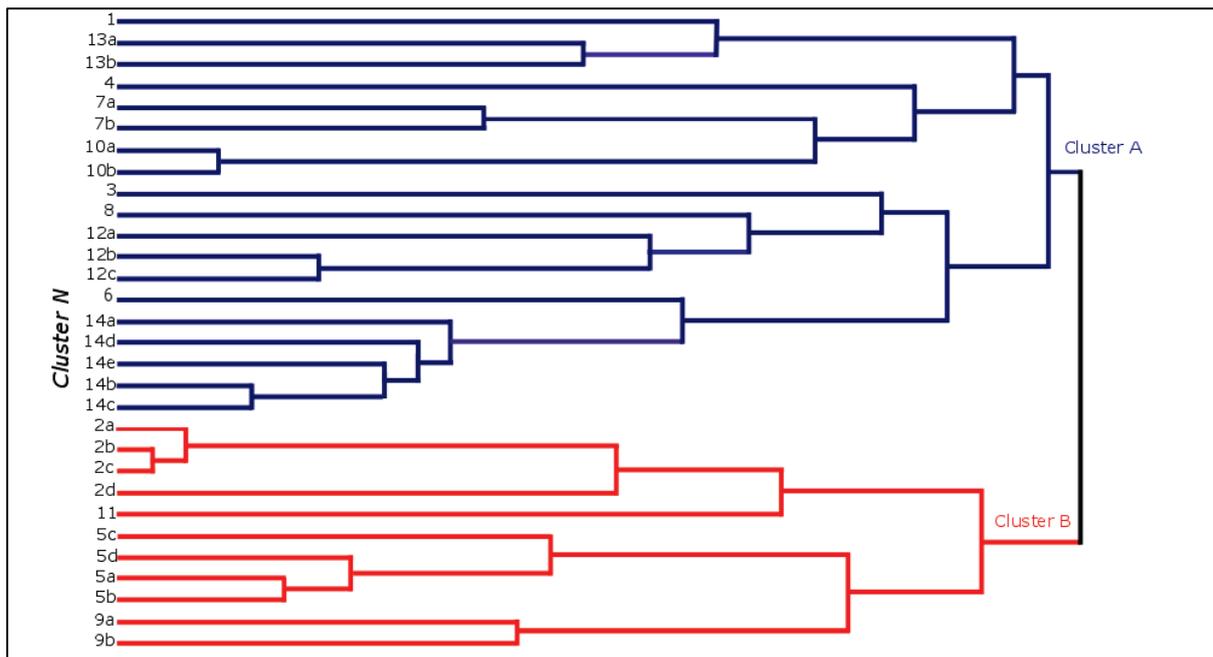


Figure 9-1 Dendrogram highlighting a two cluster solution

Figures 9.2 and 9.3 below show the significant tokens which characterised the two root clusters. Figure 9.1 shows that dominant tokens within cluster A included *face* and *assault* which suggest violent methods of transfer sometimes resulting in *injury*. This use of violence reflects the high proportion of cases in the cluster that are classified as robbery of personal property. Other dominant tokens relate to property stolen, particularly *phone* and *mobile phone* and also *jewellery*. The tokens *car* and *dwelling* are significantly over-represented in this cluster reflecting the fact that robberies taking place at dwellings (original cluster 6) and relating to vehicles (originally cluster 9) are subsumed within this larger cluster. There is also evidence of engagement between offenders and victims with the significances of tokens such as *hand over* and *verbal*.



Figure 9-2 Significant tokens in cluster A of a two cluster solution (font size relative to chi-square score)

Figure 9.3 below shows that the most over-represented token in cluster B is *bag*, followed by *purse* (also including wallets) and *shop*. These tokens relate to targets, target enclosures and wider environments. Figure 9.4 repeats the same results but with three most over-represented tokens removed. This wordcloud highlights features of the nature of transfer in this clusters which are characterised by secrecy and stealth as indicated by tokens such as *missing* and *notice* (which relate to victims noticing later that items are stolen). *Offender not seen*, *bump into* and *nothing felt* are also significantly over-represented. These tokens are indicative of theft from the person as is the absence of tokens relating to violence and the threat of violence.



force. The validation has shown that the use of force is a key element in the determination of the clusters but there are additional elements that help to form the distinction between these types of crime.

### 9.2.2 Comparing the stability of clusters over time

Section 8.2.1 created a selection of clusters based on a dataset spanning the whole period January 2007 to December 2009. In the analysis below, two subsamples have been extracted from the larger dataset; these are for Q1 (Jan to March 2007) and Q9 (Jan to March 2009). These two quarters were selected because they allowed the comparison of a quarter at the beginning of the analysis period with a quarter towards the end of the time period while comparing Jan-March in both years, avoiding skewing the comparison with seasonal differences. Table 9.2 below maps the similarities and differences between the cluster solutions produced for the three samples.

#### *Quarter 1*

The sub-sample for Q1 contained 2,941 cases. With fewer cases, a smaller number of clusters was produced compared to the analysis of the whole sample. All of the 2,941 cases were successfully matched to one of the 11 clusters. As shown in Table 9.2 these clusters closely matched those that were produced for the sample as a whole. This finding provides confidence in the clusters that were identified. Clusters from the whole sample that were very closely replicated in the Q1 sample included:

- 3. Jewellery Snatches
- 6. Robberies within Dwellings
- 12a Assault with injury
- 12b Assault without injury
- 12c. Robberies within Dwellings: Assault

Clusters that were comparable, but with some differences to the whole sample solution, included:

- 7. Demands and Threats – there was no separation of threats with and without knives.
- 9. Vehicle Related Crimes – were all combined within one cluster.
- 14. Public Transport – only one cluster identified no distinction between transport locations.

Other Q1 clusters were not directly comparable but could be mapped to the whole sample of clusters in a straightforward manner. There were two Q1 clusters relating to bag snatches, which

included one which replicated 2b Bag snatches from pushchairs and another, which appeared to be a combination of 2a Bag Snatches with resistance and injury and 2c, Bag Snatches involving bikes.

Two clusters in the Q1 sub-sample related to dips and pick pocketing. There was no straightforward mapping between these clusters and the six related clusters in the whole sample (i.e. clusters 1, 5a, 5b, 5c and 11). Thus, the cluster solution for Q1 did not distinguish between pickpocketing and thefts from bags or between thefts in shops and thefts in licensed premises (including entertainment venues). The two clusters identified in the Q1 sample were very similar. One included significant occurrences of *verbal* communication, the second included significant mentions of witnesses and the absence of CCTV.

Clusters identified in the whole sample which were not identified in the Q1 sample included:

- 2d. Bag Snatches after following victim to dwelling
- 4. Theft of Cycles
- 5d: Robberies of shop staff – these offences appeared to be subsumed within the bag snatch clusters (as was the case in the whole sample before further partitions were imposed)
- 8. Involving taxis
- 10 Robbery involving young people.

It is not surprising that 5d was not identified in this sub-sample as no cases with this categorisation occurred during this period.

Three Q1 clusters related to the theft of mobile phones. The first of these involved mobile phones snatched from the victim's hand. There were two principal offending pathways within this cluster. In the first, the victim held the phone in their hand (making a call or listening to music) enabling the offender to snatch the property. In the second, the offender makes a request as part of a conversation which requires the victim to get out the phone (e.g. asks for the time, asks to look at phone). The other two mobile phone theft clusters were very similar to each other. In both these clusters, the target enclosure was most frequently the victim's pocket, as with thefts from the victim's hand there were different means through which the offender overcame this target enclosure. Thus, offenders either made a request that required the victim to get out their phone, or the offenders made threats and demands or directly searched the victim's pockets. The key difference between these two clusters was that the first included significant overuse of the token *groups*, which, on inspection, refers to groups of offenders.

The manner in which theft from the person and mobile phone theft have been subdivided in different ways, during different periods, may highlight some concerns about the validity of the partitions. The different sets of clusters have highlighted that there are different ways in which these offences can be classified and there may be a risk of over-simplifying the various combinations in which offences can take place. For example, the whole sample solution divided mobile phone thefts according to the nature of engagement with victims (demands/search) and the nature of expropriation (snatch). The Q1 subset, segmented mobile phone thefts according to the target enclosure (hand vs. pocket). In reality, both sets of partitions are likely to be important but there may not be enough cases of each set of combinations, to produce clusters representing every configuration. It should also be remembered, that, mobile phones also feature in other clusters as objects stolen. It is, therefore, strongly recommend that cluster analysis be used as a means to explore and identify key dimensions within the data, including dimensions along which crime may be changing, not to promote one form of dimension as more important than another.

### *Quarter 9*

A total of 2,378 cases were reported during Q9 (Jan-March). All of the cases were assigned to one of 19 clusters. It is apparent that more clusters from the whole sample are directly replicated in the Q9 sample that was found to be the case in Q1. Clusters from the whole sample that were very closely replicated in the Q9 sample included:

- 1. Pickpocketing in Venues
- 2b. Bag Snatches, from chairs /verbal aggression (no resistance and or injury)
- 3. Jewellery Snatches
- 4. Theft of Bikes
- 5a Theft from Bags (shops)
- 5b Theft from Bags – relating to banks
- 6. Robberies within Dwellings
- 7a Threats with knives
- 7b. Threats without knives
- 11. Theft from bags: Licensed Premises
- 12b Assault without injury
- 12c. Robberies with domestic assault
- 13a. Mobile Phone Theft: Demands/Search
- 13b:Mobile Phone Theft: Snatch
- 14b. Bus stop

Bag snatches were divided across two clusters. As noted above, in the cluster, bag snatches with pushchairs, verbal aggression but not significant resistance or injury was replicated in Q9. The other bag snatch cases appeared to be combined within another separate cluster (i.e. the original clusters 2a and 2c). Involving bikes and at dwellings did not form their own clusters in Q9. In Q9, offences of dipping or theft from bags were spread across three clusters: Theft from bags: licensed premises, theft from bags: shops and theft from bags: relating to 'banks'. These are similar to clusters identified across the whole sample, although, the finer detail of theft involving distraction (5b) was not identified. Further, the robberies of shop staff (5d), which were uncovered from within the whole sample clusters of theft from bags, were not identified within Q9. Thefts relating to motor vehicles were combined within one cluster, rather than split into two as was the case with the whole sample. Although offences involving bus travel were identified as per the whole sample, the clusters did not differentiate between different segments of the journey as successfully in Q9.

Clusters from the whole sample solution that were not identified in Q9 included:

- 2d. Bag Snatches after following victim to dwelling.
- 5d: Robberies of shop staff.
- 8. Involving taxis.
- 10 Robbery involving young people.

In addition to clusters that were directly replicated, some clusters, identified in Q9, provided additional details not uncovered when the sample was analysed as a whole. Two robberies with threats/demand clusters comparable to the whole sample (7a demands with knives and 7b demands without knives) were supplemented by a third cluster in the Q9 sample which identified a small cluster of 24 cases involving threats involving *firearms*. An additional cluster, which related to the cluster 12 robbery with assault, involved assault and threats, but also, the restraint of the victim. In these cases, the token *grab* was statistically significant but referred to the offender grabbing the victim rather than an item of property. Other terms used to a significant level in this cluster include *hug* and *drag*. *ATM* was also significantly overused in this cluster, indicating that a significant number (12) of these offences occurred at a cash machine.

### *Summary*

Comparing two cluster solutions that were prepared for different time periods within the data demonstrated a degree of stability in the clusters over time, adding some confidence that the classifications of the MO descriptions are have sufficient real world reality. The successful

identification of clusters within smaller sub-sets of the dataset also demonstrates that this analysis can be used with smaller datasets.

Many of the clusters produced for the whole time period were closely replicated in Q1 and Q9, more so in the latter quarter. In some cases, the clusters were similar but the manner in which differences between MO descriptions were organised differed, which was particularly the case for the bag snatches clusters. Sometimes clusters were similar but split on the basis of different MO dimensions. Thus, with mobile phone thefts, the nature of communication formed the distinctive feature dividing clusters in the whole sample and in the Q9 sample. In contrast, the divisive element in Q1 was the nature of the target enclosure (pocket or hand). In Q9, robberies involving threats using firearms formed a cluster which was not identified in the whole sample or the Q1 sample. This raises caution about concentrating only on limited dimensions in which the crimes were different; it also shows how different numbers of clusters can produce different interpretations. And it highlights the benefits that are gained from cutting up the data in slightly different ways to explore the impact on interpretations. Thus, although clustering is an automated, computerised process that places cases into the different groups, the output from this process can only assist in interpretation - it does not replace it, it is the human who then decides what factors have created the partitions. It is also humans that need to decide whether those partitions are important for the research questions in hand or the practice problem being faced. This suggests that this approach to data analysis is not one that can be expected to run largely automatically in order to produce results for the practitioner.

Table 9-3 Comparing cluster solutions between the whole sample, Quarter 1 (Jan to March 2007) and Quarter 9 (Jan to March 2009)

Whole Sample Cluster	Q1 Jan - March 2007	Q9 - Jan - March 2009
1 Pickpocketing at Venues	Combined with 5a, 5b, 5c and 11	Replicated
2 Bag snatches		
2a <i>including resistance and injury</i>	Combined with 2c	Combined with 2c
2b <i>from pushchairs and including verbal aggression</i>	Replicated	Replicated
2c <i>offender on cycle</i>	Combined with 2a	Combined with 2a
2d <i>including following victim home</i>	Not identified	Not identified
3 Jewellery Snatches	Replicated	Replicated
4 Theft of Cycles	Not identified	Replicated
5 Theft from Bags (Shops)		
5a <i>theft from Bags (Shops)</i>	Combined with 1, 5b, 5c and 11	Replicated
5b <i>theft from Bags (with distract)</i>	Combined with 1, 5a, 5c and 11	Not identified
5c <i>including banks, cards and withdrawals</i>	Combined with 1, 5a, 5b and 11	Replicated
5d <i>robberies of shop staff</i>	Not identified	Not identified
6 Robberies within Dwellings	Replicated	Replicated
7 Demands and Threats		
7a <i>not involving knives</i>	Combined	Replicated
7b <i>involving knives</i>		Replicated
		<i>NEW! Involving Firearms</i>
8 Involving taxis	Not identified	Not identified
9 Vehicle related crimes		
9a <i>Carjacking and thefts from cars</i>	Combined	Combined
9b <i>Car thefts from dwellings</i>		
10 Robbery involving young people		
10a <i>in parks</i>	Not identified	Not identified
10b <i>on footpaths, alleyways and in schools</i>	Not identified	Not identified
11 Theft from bags at Licensed Premises	Combined with 1, 5a, 5b and 5c	Replicated
12 Robbery with Assault		

Whole Sample Cluster	Q1 Jan - March 2007	Q9 - Jan - March 2009
12a <i>with assault and injury</i>	Replicated	Not identified
12b <i>with assault no injury</i>	Replicated	Replicated
12c <i>with assault domestic</i>	Replicated	Replicated
		NEW! Restraint
13 Mobile phone theft	Replicated but different sub-clusters	
13a <i>with communication</i>	NEW! <i>snatched from hand</i>	Replicated
13b <i>without communication</i>	NEW! <i>Taken from pocket - group offenders</i>	Replicated
	NEW! <i>Taken from pocket</i>	
14 Thefts on buses and at bus stops		
14a <i>at the bus station</i>		
14b <i>at the bus stop</i>		
14c <i>on the bus</i>	Combined	Combined
14d <i>leaving the bus</i>		
14e <i>misc</i>		

### 9.2.3 Determining the number of clusters

The BKM clustering technique applied in Section 8.2 requires the researcher/analyst to pre-select the number of clusters (K) that will be extracted from the data. This decision has the potential to influence the results and their interpretation. Section 8.2 discussed the application of statistical methods that can help to guide the researcher on the optimum number of clusters to select. The dataset for this research was analysed using different cluster solutions ranging from 2 clusters to 30. Each of these solutions was examined and a comparison made of the ratio of between-cluster variance and total variance for each of these cluster solutions (see Table 8.2 above). This comparison suggested that a 14 cluster solution offered the best ratio of between-cluster variance and total variance, in other words this statistical method of determining K suggested that 14 clusters should be extracted.

The ratio of between-cluster variance and total variance provides a form of yardstick that can help decide the number of clusters. However in Section 8.2 this guideline is set aside and the section shows the substantive benefits of increasing the number of partitions to produce 30 clusters. This more detailed solution revealed scripts and tracks that may not have been identified if the partition of the data stopped at 14. For example, the 30 cluster solution enabled distinctions to be made between offences taking place at bus stations, bus stops, on the bus or when leaving the bus – while the 14 cluster solution grouped all of these offences together. Importantly, while the 14 cluster solution subsumed thefts from shop staff (cluster 5d) within a larger cluster of theft from bags in shops, the more detailed 30 cluster solution was able to identify this group of offences and separate them from offences with which they had little in common.

The identification of additional scripts provides justification for increasing the number of partitions but this raises the question of when to stop increasing the number of clusters. The interpretation of the clusters identified that some of the 30 clusters contained distinct sub-groups, these appeared to represent more detailed scripts and tracks. For example, although the 30 cluster solution

distinguished thefts of cars from dwellings from other types of vehicle related offences, the researcher's inspection of the clusters identified two different scripts, carjackings and theft from persons who were in cars, were subsumed within one cluster. It is possible that further partitions of the data with a greater value of K may have been able to formalise these and other groupings. There does appear to be more scope to divide the data further, particularly as the clusters in the 30 cluster solution remained quite large with a mean of 686 cases. However the selection of 30 clusters was partly influenced by practical concerns regarding the ability to fully present and interpret each cluster within the bounds of a reasonable amount of effort. The determination of K is ultimately a subjective decision made by the researcher/analyst which must balance the statistical validity of clusters, the ability to uncover substantive findings and the resources available to interpret the results.

#### 9.2.4 Comparing narrative and N-gram analysis

The methods of text analysis adopted in this research identify the frequency and distribution of different words, or tokens, across the dataset. Words are not analysed in complete isolation, the analysis does account for the differing combinations in which words occur together, but overall the analysis does not directly reveal word ordering or the connections between nouns and verbs i.e. between actions and actors.

Section 4.3.3 highlighted the benefits of a narrative approach to the quantitative analysis of textual descriptions of action (Franzosi, 2010). This method retains clear connections between the actions taking place, the actors performing them and the objects/or actors they are directing action towards. The advantage of this approach is that it offers greater potential for identifying exactly what actions are taking place and what factors are shaping that action. However, as discussed in the methodology section a review of this technique concluded that it was too labour intensive to be used with large volumes of cases.

The alternative approach adopted was a 'bag-of-words' technique. This successfully distinguished between different crime scripts based on the words being used to describe them. This included scripts where a similar token performed a different 'role' in the narrative of the script.

For example 'Bike' was a significant token in both cluster 4 theft of bikes and the sub-cluster 2c bag snatches by an offender on a cycle. While these scripts share some overlapping vocabulary the role of bikes in each script differ. In cluster 4 bikes are the object of theft in cluster 2c the bike was a resource used by the offender. The analysis method was able to distinguish between the different functions of this token due to the different words that were combined with each script and in particular the absence in cluster 4 of alternative tokens related to property stolen. However, elsewhere in the findings the analysis was not able to make these important distinctions drawn from the context in which words were used. This was particularly apparent when verbs were amongst the significantly overused words in a cluster. In such cases it was not automatically apparent which actor (i.e. offender, victim or onlooker) was performing a specific action. For example, to determine which actors were walking (Cluster 1) or arguing (Cluster 3) required further inspection of the cases within a cluster.

The temporal order is another structure present in narrative accounts that is important in understanding how events unfold. This structure is lost when words are analysed as individual tokens. Cluster 3 contained significant use of the term 'argue' but from this finding alone it is not possible to determine which actors were arguing nor whether the argument occurred before, during or after the theft. Sequence analysis was used to identify the tokens in this cluster which typically preceded and followed 'argue.' This revealed that offenders using this script would use an argument as a method of engagement and gaining proximity to the victim. However the findings from sequence analysis of other tokens were not always as conclusive.

The temporal order of tokens within MOs can be crucial in distinguishing between different scripts. This is particularly apparent with the use of violence. The positioning of violence in the sequence of

events can reveal different purposes for which that violence is used. In the examples below, in line with quantitative narrative analysis, MO descriptions have been structured into clear temporal stages. The three examples show that violence plays a different role in each case dependent on where in the sequence it is employed. In the first example violence is used immediately after approach and before an attempt to transfer the property. This is indicative of the offender employing shock tactics to gain immediate compliance from the victim (Smith, 2003). In the second example violence was used simultaneously to the transfer of property, possibly using force to prevent victim resistance. In the third example violence was used to counter victim resistance. The different sub-goals that violence accomplishes within a script can be difficult to identify when words are analysed individually.

Therefore, a consequence of the use of a 'bag-of-words' technique was that the identification of scripts and tracks within the clusters almost inevitably required examination of cases within each cluster. However, this inspection could be conducted on selected sub-sets of a more manageable size than the overall dataset, and importantly these sub-sets have been pre-selected as having some features in common. The textual analysis employed in this thesis cannot automatically extract all relevant information from MO data but it can be used to focus and streamline the analysis and interpretation of patterns in the nature of crime commission.

EXAMPLE 1: UNKNOWN OFFENDER APPROACHED IP FROM BEHIND AND PUT A KNIFE TO THE IP'S THROAT. OFFENDER THEN ASKED FOR IP'S MOBILE PHONE. IP GAVE THE PHONE TO THE OFFENDER WHO THEN RAN OFF WITH SAME.

N1 - OFFENDER - APPROACHED IP FROM BEHIND.

N2 - OFFENDER - PUT A KNIFE TO THE IP' S THROAT.

N3 - OFFENDER -THEN ASKED FOR IP' S MOBILE PHONE.

N4 – IP- GAVE THE PHONE TO THE OFFENDER.

N5 –OFFENDER - THEN RAN OFF WITH SAME.

EXAMPLE 2: OFFENDERS APPROACHED IP WALKING IN THE STREET. OFFENDERS SNATCHED MOBILE PHONE AND KICKED AT IPS BODY. OFFENDERS STOLE MOBILE PHONE AND MADE GOOD ESCAPE.

N1 - IP - WALKING IN THE STREET.

N2 - OFFENDERS - APPROACHED IP.

N3a - OFFENDERS - SNATCHED MOBILE PHONE.

N3b - OFFENDERS - KICKED AT IPS BODY.

N4 - OFFENDERS - STOLE MOBILE PHONE.

N5 - OFFENDERS - MADE GOOD ESCAPE.

EXAMPLE 3: UNKNOWN OFFENDER(S) HAVE ASKED THE IP FOR THE TIME IP HAS LOOKED AT HIS PHONE FOR THE TIME WHEN THE OFFENDER(S) HAVE THEN ASKED FOR THE PHONE IP REFUSED AT FIRST WHEN ONE OFFENDER HAS PUNCHED IP TO STOMACH IP HAS GIVEN THE OFFENDER(S) HIS MOBILE PHONE OFFENDER(S) HAVE THEN THROWN THE SIM CARD OUT OF THE WINDOW.

N1 - OFFENDER(S) - HAVE ASKED THE IP FOR THE TIME.

N2 – IP - HAS LOOKED AT HIS PHONE FOR THE TIME.

N3 - OFFENDER(S) - HAVE THEN ASKED FOR THE PHONE.

N4 – IP – REFUSED.

N5 – OFFENDER - PUNCHED IP TO STOMACH BY OFFENDER.

N6 – IP - HAS GIVEN THE OFFENDER(S) HIS MOBILE PHONE.

N7 - OFFENDER(S) - HAVE THEN THROWN THE SIM CARD OUT OF THE WINDOW.

## Chapter 10 Conclusions

### *10.1 MO Descriptions: Intelligence for Crime Prevention?*

This thesis has explored the question of whether using innovative techniques of textual analysis can produce actionable intelligence for crime prevention from police recorded MO descriptions. MO data provide a description of the circumstances surrounding the commission of criminal offences and the actions taken by the offender to commit them. The focus on the criminal event stems from theoretical perspectives that aim to explain, and prevent offences, by developing an understanding of the situational factors that create opportunities for crime. The first broad question addressed by this research was “*Are police recorded MO data a potential source of actionable intelligence to inform crime prevention?*” It is useful to revisit Ratcliffe’s (2008) distinction between data, information, knowledge and intelligence to answer this question. For MO descriptions to be regarded as intelligence, it needs to be demonstrated that they are relevant to the specific purpose and that once they have been interpreted, they have the potential to guide or inform action.

#### 10.1.1 Relevant for purpose?

According to Ratcliffe’s definition, determining whether MO descriptions can provide intelligence to inform crime prevention requires a consideration of what types of information are *relevant* to understanding the causes of crime. Ekblom’s CCO offers a comprehensive map of the causes of crime, as identified by situational crime theories, and their operation in the proximal situation. Each of the eleven CCO elements offers a set of opportunities to intervene to block the causal mechanisms and prevent crime. Using the CCO to review the findings of the data analysis in Chapter 8 helps to reveal which elements of crime causation are included by MO descriptions and which elements tend to be missing. It can be seen that the analysis of MO data, for theft from the person and robbery, yielded intelligence relating to:

- resources for crime
- offender presence in the situation
- crime targets (including limited information on target person/victim)
- target enclosures
- wider environment
- crime promoters
- crime preventers

The MOs yielded little information on CCO elements from the 'offender' side of the framework, notably:

- predisposition to offend
- lack of skills to avoid crime
- readiness to offend

and perhaps, crucially:

- offender anticipation of risk, effort and reward

Crucial aspects of the 'offender side' of the CCO about which the data did not provide any information were the offenders' moral perspectives and lifestyle factors. From the perspective of situational action theory (Wikström 2006a) it was argued that these issues are central to understanding the actions that individuals take, shaping both the long term predisposition to commit crime, the short term prompts into readiness to offend, subjective interpretations of the opportunities presented by the setting and ultimately whether and how offenders will act upon crime opportunities. Analysing MO descriptions in isolation from other information about offenders risks creating a gap in knowledge around the distal causes of crime and how these remote causes influence the proximate crime situation. It is therefore necessary to seek out additional sources of information that can provide supplementary information, this might include data on offender and victim backgrounds and combining analysis of police records with findings from qualitative studies that have examined the lifestyles and values of offenders.

There are some types of crime that may prove an exception, in that at least some intelligence about moral values and lifestyle factors could be extractable from MO descriptions. Hate crimes for example, are motivated by offenders' prejudices towards certain groups in society, evidence of this prejudice is often revealed during the conduct of the offence.

Overall MO descriptions have clear potential to provide intelligence on some, but not all of the situational causes of crime. The shortage of information on the offender elements of the CCO has implications for the wider interpretation of MO descriptions which will be discussed below.

### **10.1.2 Accurate and complete accounts?**

In addition to providing information relevant to crime prevention, for MOs to provide actionable intelligence they need to provide an accurate, sufficiently complete and reliable account of the crimes they describe. These issues were explored in the data review, Chapter 6. The review of

sample MO descriptions identified a number of issues with the completeness of MO fields. Of crucial importance to the completeness of intelligence from MO fields is the fact that police data only reflect crimes that are reported to the police, and that there is a likelihood of systematic differences between reported and unreported crimes. Administrative processes that could shape data, include 'crimeing' and 'no-crimeing' decisions and the crime classification that is assigned to the event. There is, also, the potential for crimes to have been incorrectly classified or no-crimeed and therefore excluded from this analysis – analysis of incident data and a wider range of crime types could check for this.

The review suggested that reported MO descriptions rarely provide a full picture of all the events leading up to and resulting in the completion of a crime event and frequently present a partial snapshot. This was borne out in the analysis. From the description of script tracks in Chapter 8, it was clear that the intelligence derived from MO descriptions focused predominantly on the stages of offence commission relating directly to the transfer of property. The descriptions provided intelligence on the nature of approach, engagement with the victim, searching for property and transfer of property to the offender. Other than identifying the resources that were used in some of the script tracks and permutations, the MO descriptions revealed little about planning and preparations for offences and could not help to clarify whether planning occurred in any of the tracks identified. Importantly, the MO data does not provide any direct intelligence on the processes of recognising and selecting a vulnerable target. The incompleteness of description reflects, in part, the brevity of MO fields and the prioritising of certain types of information over others, but predominantly reflects the fact that these elements are not known having not been directly witnessed. Where there is no mention of a specific action, it is unclear whether this is because the action did not occur. This can be noted with the theft of bags from shops cluster where one track involved distraction; however, it cannot be known for certain whether the contrasting track was truly one of theft without distraction or theft where distraction was not noticed, reported and/or recorded.

In contrast, the review of existing research provided a more complete picture of offence commission, perhaps with the exception of the scenes that follow transfer, i.e. escape and disposal for which less evidence was available. This provides the potential to use existing research evidence to triangulate findings and populate gaps in the knowledge derived from MO descriptions.

## *10.2 MO Descriptions: Extractable Intelligence?*

The review of data, Chapter 6, also highlighted characteristics of MO descriptions that present challenges for the analysis of these data. The thesis aimed to explore whether techniques drawn from content analysis and natural language processing would help to alleviate these problems. Several issues were successfully resolved through pre-processing techniques.

Frequently occurring superfluous information, particularly information pertaining to the recording process and the auditing of crime records, was identified and removed by the extension of stoplists. Although a small number of these phrases remained in the data analysed, in the event their inclusion had no impact on the analysis because such phrases were evenly dispersed across the data. That is to say, they did not influence the clustering processes and no groupings were formed around these phrases. Problems such as variations of tenses, spelling errors, abbreviations, punctuation, white spacing and multi-word phrases were all at least mitigated by pre-processing of data. Negating phrases (such as 'no force') were identified to an extent, but where the negator was far removed from the related phrase, it is less likely that they have been picked up. Coding dictionary refinements helped to improve the correction of error and variations.

More challenging to address were the characteristics of the data that reflected key features of natural language. The reduction of the flexibility in the data created by the use of synonyms and other related words was time consuming to address, and was the main cause of the multiple (twelve) iterations of dictionary preparation and analysis. Although very time consuming, and far from automatic, the manual input is captured in coding dictionaries that can be used in future research, as can the procedure, developed and documented here.

Disambiguation, the clarification of word meanings when a similarly spelled word can be used in different ways, was also time consuming. This effort is harder to retain in dictionaries because the accurate interpretation of these words is dependent on their context. However, after inspecting the results of the cluster analysis it was apparent that in one sense this analysis technique mitigates, at least to a degree, the problems of synonyms and disambiguation automatically. Both synonyms can be identified by similarities in the context of their use while ambiguous words can be distinguished by differences in their context. When used for textual analysis, clustering produces classes of cases based on similarities in the combinations of words used. Thus synonyms or related words that had not been identified (or where a decision had been made to maintain the distinction) frequently appeared as significant in the same cluster due to the similarities in their co-occurrences with other words. For example, both motor bikes and cycles were grouped into cluster 4 'theft of cycles'. Not

because the cluster analysis picked up the similarity in their meaning, but because these terms tended to be used with similar combinations of words. In contrast, similar words, or identical words that were performing different 'roles' in the MO, were successfully grouped into separate clusters because of differences in the contexts of their usage. This was seen with the involvement of bikes in clusters, where it was possible to distinguish between the bikes as the target of theft and bikes as resource used in theft because, in the former instance, words relating to other types of crime target were absent (i.e. because bikes were the target).

Two problematic issues remained even after pre-processing, both of which relate to the narrative nature of MO descriptions. The bag-of-words technique, which is the start of the process of cluster analysis, loses the connections in the data between actors and their actions. This is important when there is more than one actor in the scene. Thus, where words such as walk, run, hit occur significantly in a cluster, there is a need to inspect the data to determine whether these were the actions of the offender, the victim or some other person. There is a risk of making the assumption that actions relating to assault are always the action of the offender, when they may be indicative of victim resistance or retaliation. The bag-of-words approach also loses the temporal order of actions in the description. This is perhaps less damaging in these theft from the person and robbery of personal property crimes, where the acts described are relatively brief but even so, it still presented difficulties in determining the order of certain actions: for example, determining whether violence preceded or followed victim resistance required further inspection through sequence analysis of cases in the cluster.

Techniques drawn from text analysis disciplines have clearly facilitated the extraction of data from MO descriptions and the process of preparing them for analysis, albeit with some limitations. The techniques offered time efficiencies but were far from being automated. Some of the effort involved in the more manual stages of processing can be preserved in coding dictionaries which could be used in future research, thereby saving time and effort. Working with other crime types might require adjustments to dictionaries<sup>55</sup>, and this opens up the possibility to create a bank of dictionaries of different crime types. The pre-processing did not result in any data loss because a link was retained to the raw data: when inspecting data to further investigate emerging patterns it was possible to do

---

<sup>55</sup> At the time of writing the dictionaries had already been used to conduct analysis of data relating to a different crime type (hate crimes) for a different police force.

this with the original text. It was also possible to check the original words that had been coded into tokens by the dictionary, although there was rarely a need to this.

The data problems that remain, once processing has been completed, relate to actor-action relationships and the temporal sequence of actions, both of which are characteristics of narrative accounts. Addressing these issues required further inspection of the data once cluster analysis was complete.

There are questions about the decisions made by an analyst when coding the data: when a decision is made to treat two words as related, and therefore code them under one token, this is not a fully objective process and different analysts could reach different decisions. There are questions about whether distinctions made between different terms were really necessary and these words could have been coded together (e.g. 'demands' such as 'give me your wallet' were coded separately from 'demands + threats' e.g. 'give me your wallet or I'll stab you' but the analysis placed them in the same clusters). Other terms that were coded together might have been kept separate e.g. distinguishing between serious and minor injuries rather than coding them together. The ability to have flexibility within dictionaries, to be able to easily move up and down coding hierarchies as required by the research/analysis question, would be desirable.

### *10.3 Data Analysis: Identifying Similarities and Differences in Offence Commission*

The analysis demonstrated that it was possible to use the vocabulary of MO descriptions to identify similarities and differences in offence commission. This included distinguishing between crimes that had been assigned different Home Office crime codes (Section 8.1.) This analysis also successfully highlighted the characteristic elements that distinguished the two offences. This included the identification of configurations of wider environments and offence commission, for example, showing that theft involving stealth and secrecy (i.e. theft from the person) were more likely to occur in busy and crowded environments, but thefts involving threats, aggression and violence were unlikely to occur in this type of environment. However, the targets of theft were largely spread across the two crime types with mobile phones, wallets and money etc. being taken in both protocols.

Cluster analysis successfully revealed natural groupings within the data, some of which might otherwise be hidden or overlooked by the analyst or researcher or supplanted by subjective groupings. The cluster analysis was able to group together similarly performed MOs based on helpful

and theoretically relevant distinctions. These distinctions commonly revolved around the wider environment, target property and the nature of transfer. Some of the distinctions were quite subtle, such as separating out threats to stab from real stabbings. Distinctions were frequently based on the absence of other terminology from a cluster.

The assignment of cases to clusters is not precise: assignment is based on the best match rather than specific criteria. This, along with the risk of missing data, suggests that cluster analysis is more suited to exploring the range of ways in which offences are being committed rather than a means to precisely quantify the number of offences that are committed according to each track. It is still possible to identify which tracks predominate in order to prioritise preventative effort and to monitor the impact of intervention through the differential proportions of script tracks, but precise estimation of the number of crimes committed according to a specific track with specific permutations remains difficult using this method.

Methods were used to investigate the validity of clusters. These included comparing the cluster solution to the results of previous research, comparing the stability of clusters over time and investigating the proportion of robberies and theft from the person in each cluster. All three methods provided confidence that the clusters produced reflected groupings in real world offence commission. Investigation of clusters at different time points revealed that some classifications were so strong that they emerged however the data was subdivided. With others, distinctions were only revealed when clusters were cut up in different ways. The use of cluster analysis is strengthened when attempts are made to explore the impact of different cluster solutions on the interpretation of the results.

Although the cluster analysis produces groups with clear statistical similarities it remains necessary to interpret the substantive similarities between clustered cases. This can require further investigations into the cases that have been clustered, such as the sequence and context of characteristic words in the cluster. These investigations can be conducted in a far more focused way than simply exploring raw data, by querying specific anomalies or hypotheses generated from the analysis. Interpreting the clusters requires an additional process that has been defined as sense-making (Weick, Sutcliffe, & Obstfeld, 2005). Sense-making is a process of comprehending results and communicating them in order to guide action. It requires the extraction of cues and the identification of plausible explanations. These explanations can be informed by the researcher's theoretical and empirical knowledge. The importance of this stage should be recognised because although tools for preparation have the potential to provide time efficiencies, the interpretation or sense-making stage still takes time, and can involve following up 'leads' in the analysis results by

revisiting the data and conducting further analysis. Therefore, this approach is not one which can be regarded as a routine process that delivers 'press of the button' interpretations. Although this is an improvement over the manual analysis of MO fields and it allows far larger samples of data to be analysed, the full process may still be too time consuming for practitioners to repeat as part of routine intelligence operations.

## *10.4 Interpretation: The Role of Theory in Understanding Offence*

### *Commission*

#### 10.4.1 The role of conceptual frameworks

This thesis aimed to consider the extent to which conceptual frameworks, the CCO and scripts, facilitated the identification and interpretation of patterns in MOs, specifically for the purpose of crime prevention. The CCO serves to represent situational crime theories, providing a roadmap through the causes of crime identified by these theories. The review of situational crime theories in Section 2.1 argued that, although, these theories have been successful in explaining crime causation, critiques of the theories demonstrate that they would benefit from some refinement. In particular, there have been calls for situational crime theories to provide a fuller consideration of offenders, their motivations, resources and constraints. The CCO is compatible with both the traditional approaches to explaining the crime situation and with the more recent refinements suggested by Ekblom, Wortley, Wikström, Bouhana and others.

The CCO was used to organise knowledge about crime commission, including that derived from the review of existing studies, Chapter 5, and the findings from the cluster analysis of police data, Chapters 8 and 9. Applying the CCO was relatively straightforward, although, one issue identified was the overlap of CCO elements. This is a side-effect of the nature of the conjunction, in that all elements are defined by their interaction with the others. For example, the distinction between target enclosure and wider environment is not always clear (is a bus an enclosure or an environment?) while the research demonstrated the subcultures within which offenders are immersed, can provide both resources to offend and restrict the offender's opportunity to desist from offending.

Using the CCO jointly with script frameworks encourages the researcher to think of situational causes, not as static factors which offenders (and other actors) stumble into, but to show how actors actively manipulate, interact and cope with situations to try and meet their goals. Scripts were introduced to describe a cognitive facility that helps actors to make rapid and effective decisions to

direct instrumental behaviour. Research into theft from the person and robbery has highlighted that a necessary feature of decision making and action in these crimes is that they need to be rapid. The research literature frequently uses scripts, and concepts that parallel or complement scripts, to explain how these rapid decisions are made. For example, researchers have shown how offenders recognise and act upon cues in the environment, how offenders use scripts/or templates to route their decision making, the role of experience in replacing steps in decision making and the incorporation of this experience into learning and future revision of decision templates or scripts.

Ekblom and Gill (2015) argue that there is inadequate linkage between static snapshot models (including the CCO) and the dynamics of scripts. However, using the CCO and script frameworks jointly demonstrates that they can fit together very well. There are two key reasons why applying the CCO and scripts together was advantageous. Firstly, scripts show how situational conjunctions are created dynamically through the interactions of actors and their environments. Secondly, regarding crime commission as a process and paying attention to each act, shows that crime causation is the result of not one, but of a series of conjunctions with the potential for situational factors to vary in each act or scene. Analysing crime commission through a joint script and CCO framework, applied both to the literature and the police data, revealed multiple instances of offenders manipulating rather than merely reacting to the situation, for example, going to a different destination or creating a distraction. The manipulation of the situation to achieve the sub-goal of victim compliance was demonstrated in the use of weapons (both in the literature and the police data), whereby weapons were carried but their use restricted to the minimum required to gain victim compliance.

Seeing scripts as the way in which offenders accomplish goals that line up to their advantage the situational conjunction highlights the steps that are necessary to commit an offence. The CCO can be used to frame the goals that the offender needs to achieve to complete different stages of offence commission, as was shown with the use of CCO elements to describe the necessary conditions for the completion of a script act (as illustrated in Table 5.1 above and the diagrams in Appendix 2). This adds a level of detail to the available intelligence regarding the potential points to prevent offenders from achieving their goals. Thus the analysis is consistent with the notion of goal-directed action that is at the heart of the choice perspectives of offender decision making. The regularity with which the same behaviours reoccur in the data, not only allows patterns to be identified, but also suggests that goal-directed actions can be inferred through consistent patterns of behaviour relationships. Scripts help to routinize decision making and this suggests that behaviour will remain consistent with a decision unless a reason emerges to do otherwise. The consistent relationship between decisions

made and rational action is not explicitly addressed in the rational choice perspectives outline of rational behaviour.

It is important not to interpret all crime outcomes as the result of offender decision making: offender's scripts are not entirely of their own making. The CCO outlines a range of actor roles that influence the crime situation: offender, crime promoter and crime preventer, and this is also in line with Wortley's advice to include consideration of the scripts of victims and other actors in the situation. For this reason the terminology of scripts may be more helpful than that of the MOs, as MO is too closely associated with the actions of the offender.

The research literature provides rich examples of interactions and clashes between actors' scripts. This begins with target selection, in which offenders interpret cues about target suitability from the actions of potential victims. The research literature shows how offenders are keenly aware of the clash between their goals and those of their victims, and they are also aware of the likely means through which victims will seek to pursue *their* goals (especially those of retaining possession of their property). Offenders' own scripts were designed to forestall and counter victims' attempts to protect their property. Thus script scenes such as approach, engagement and transfer were performed in such a way as to circumnavigate victims' counter scripts or gain victim compliance to offender scripts.

MO descriptions contain less direct intelligence about the actions of other actors in the script, and where this existed, it was harder to extract. Viewing the police data in isolation from wider knowledge would make it more difficult to identify the goals behind specific actions, but fortunately the script tracks identified closely resembled those in the literature, allowing such inferences to be made. Therefore, it was possible to identify, with sufficient assurance of validity, the strategies being used by offenders to further their own goals and block those of victims and crime preventers.

Framing offence commission as a process with sequential scenes has some clear benefits to understanding but also some limitations. Thinking about each script act and scene, in turn, acts as a prompt to consider the changing role of CCO elements at different stages of the offence. Through application of script/CCO frameworks, both to the literature and the police data, it was possible to identify distinct track options reflecting different offender approaches. In the police data, some of the script tracks identified reflected the actions of victims and consequent actions of the offender. The analysis concluded that theft from the person and robbery of personal property can be completed using a range of equifinal actions (Greene, 1989). The track options reflect the feature of successful expropriative strategies as outlined by Cohen et al., (1995) and the procedural

characteristics that Ekblom and Gill (2015) identified as common across offence categories. These transfer types were not mutually exclusive and tracks existed that combined, for example both force and surprise. Combining scripts and CCO illustrates how CCO elements can change in each act or scene, with different configurations presenting different alternatives for action. Script analysis can show how decisions and actions taken in earlier stages can act as filters to narrow down the range of options available in later scenes.

The frameworks provided a roadmap for the interpretation of both the literature and the cluster analysis to ensure that all relevant factors were given consideration. In this research, the author used a literal 'checklist' of CCO elements and script scenes as prompts to inspect each cluster. This avoids the research simply homing in on the obvious or most predominant features of a cluster.

The use of the CCO and scripts in organising the literature proved valuable in synthesising the existing evidence. Application of the frameworks helped to draw together available knowledge about each script act and scene. A great deal of consistency was identified across the studies reviewed. The studies reviewed included those that had investigated one act of the script in detail e.g. gaining proximity and securing victim compliance. These provided an in-depth view of how offenders manipulate the situation to achieve goals and sub-goals. These studies also served to highlight how individual scenes that may appear straightforward, such as approaching a target, are in fact very complex achievements.

Using script and CCO frameworks also enabled comparisons to be drawn with other crime types, particularly crimes where the lower level goals are comparable, such as identifying a hunting ground in sexual offences, and to draw up potential hypotheses about how these scripts might apply to the crime in question. Application of the frameworks served to highlight script scenes where there are gaps in existing knowledge. For example, analysis of the literature showed that less is known about post-offence activities such as escape and disposal. This was also the case in the police data as noted above. Although the literature has a lot to say about target selection, this was not directly addressed by the police data, although it can be inferred as noted above.

#### **10.4.2 Applying conceptual frameworks: Some limitations and challenges**

One element that proved difficult when working with scripts was the application of hierarchical levels (scripts, tracks and permutations). Although a number of authors have highlighted this problem, most recently Ekblom and Gill (2015), there does not yet appear to be a solution. In this study the author found it difficult to determine when a different set of methods or offence locations, or resources used, should be classed as a distinct script track or when it should be regarded simply

as a permutation of another script track. Even when setting aside rules established by other authors and endeavouring to develop new ones, the decision appeared difficult and some of the decisions are arguably subjective. The assignment of actions to scripts, tracks and permutations was made harder by the fact that permutations were often interchangeable e.g. 'communication and no weapon' or 'communication and a weapon' meaning that separate tracks for communication and weapons would not make sense, unless one track was created for all combinations of permutations with the end result of producing an uncontrollable number of multiple levels.

The use of script frameworks encourages a hierarchical organisation of the acts and scenes in the offence process, in which the more specific variations in offence commission (tracks and permutations) are subsumed under broader script types. However, this neat hierarchy may not faithfully represent offence commission. There are a number of occasions where during the performance of one script, a change in methods can switch an offence over to a completely different script. This may be due to conscious actions of the offender, or interactions between the offender and changes in the situation. For example, if an offender is confronted with victim resistance while performing a theft from the person script, he/she may decide to use violence to complete his/her goals: this converts the offence from a theft from the person script to a robbery of personal property script thereby representing a horizontal step across the hierarchy rather than a vertical step. Thus attempting to order offences into neat hierarchies may not accurately replicate the inherent flexibility in offending which may be more of a network of related options. The concept of heterarchy (Goldammer, Paul, & Newbury, 2003), which describes an organisational network that allows greater flexibility in the ordering of component elements, has been adopted in disciplines such as anthropology, cybernetics and may prove useful in the future development of script frameworks.

A further problem with the neat organisation of scripts and tracks into sequences is that one of the ways in which different script tracks can vary is in the order of scenes. Thus, although there is a general, causal, 'do this in order to do that' ordering to offences, there are scenes that can be re-ordered, such as identifying the location of property on the victim before or after engagement. Moreover, a number of the studies reviewed in Chapter 5 revealed theft and robbery as fluid activities in which many things happen at once and several goals are achieved instantaneously. For example, target selection, approach, engagement and transfer all being accomplished with one or two actions. The same action can positively further the aims of the offender (e.g. gain proximity with the target) while simultaneously preventing negative outcomes (e.g. injury, arrest, wasted effort).

Scripts revealed genuinely different methods which were available as alternative routes to theft commission and the framework proved helpful in dissecting the goals that offenders are aiming to accomplish. However, assigning alternative script tracks to a sequenced hierarchy of actions often felt like shoehorning complex interactions into overly neat categories. Scripts could perhaps be better thought of as containing a networked menu of interchangeable modules, where the selection of certain modules may promote or prevent the selection of later ones but there is more room for flexibility. Some options may be dependent upon earlier selections, but a degree of interchangeability remains and there is no single 'if this then that', route from the first scene to the last.

Finally it is worth making a comment on Leddo and Abelson's (1986) universal script. In Section 2.3 it was noted that several authors have found this template too difficult to apply including Chainey and Tompson (2008) and Ekblom and Gill (2015) and most authors that have used it have tweaked it before doing so. The notion of the universal script does help to ensure that offence commission is considered from beginning to end, rather than focus just on the 'doing' stages of the offence. However, while there are some fairly broad script stages that are common across all crime types (preparation, pre-activity, escape etc.) on their own, these are not very insightful and they require further inspection. The inclusion of preconditions and post conditions are important in thinking through what is needed to reach a goal but in the analysis in this thesis, it would though be more helpful to consider these conditions at the beginning and end of each scene.

### 10.4.3 Enriching offender accounts

Limitations to situational crime theories were identified in 2.1, with refinements identified from perspectives such as situational action theory (Wikström, 2009). It is important to consider what impact these refinements should have on the interpretation of MO descriptions. Authors seeking to revise situational crime theories have called for a greater awareness that opportunities presented in a crime situation would not be recognised or acted upon by all individuals and called for research to explain the individual differences in the perception and recognition of opportunities. Both the CCO and script frameworks are compatible with these revisions, moreover, they actively demonstrate that opportunities are not merely stumbled upon by offenders. Script frameworks invite the investigation of opportunity recognition and decision making, and the offender elements of CCO help to explore how the crime causes of criminal predisposition, readiness to offend (motivations and provocations) and offender resources shape an offender's ability and readiness to recognise and act upon favourable conditions.

Application of the CCO and script frameworks to the literature showed the benefits of refinements to situational crime theories. The use of these frameworks also serves to refute Sutton's (2006) claim that explaining crime through the situation paints a misleading picture of crime opportunity as an objective phenomenon with causal powers. In Chapter 5, interpreting the literature through script and CCO frameworks demonstrated that the presence of distinct types of predisposition and readiness shaped offender decisions about what type of offence to commit and how, as well as guiding the overall decision about whether or not to offend. However, as noted above it seems that while MO descriptions provided a lot of information on the aspects of the offence that have been traditionally the concerns of situational crime theories, they revealed less about elements that would arguably enrich accounts of offender decision making. The MO descriptions did not provide any intelligence about how offenders perceived, recognised or assessed opportunities: these factors can only be inferred from the outcomes of their actions and by using evidence from the literature to support propositions.

## *10.5 Recommendations*

This thesis has identified that police recorded MO descriptions do provide a potential source of intelligence for the prevention of volume crime and it should be noted that the quality and completeness of the MO field was better than they are often given credit for, demonstrated by the low proportion of missing and incomplete MOs and the ability to use the fields in the analysis. However, during the course of the research, observations were noted about how this potential could be better fulfilled. These observations included recommendations for police practice, and recommendations relating to the use of software techniques and technologies. Recommendations relevant to future research are discussed in Section 10.8.

### **10.5.1 Recommendations for police recording practices**

It is unrealistic to expect dramatic changes to the recording of police MOs in the near future given the current climate of efficiency savings and reducing police bureaucracy. In addition, several forces in England and Wales have recently made considerable investment in crime recording systems, meaning that the available technology is unlikely to change soon. However, the negative attention that has been afforded to the reliability of police crime data in recent years, culminating in the removal of 'official statistics' status, could be used as a lever for future improvements.

However, if feasible, the following changes would help to improve the quality and coverage of MO descriptions. Improvements to the training of those involved in crime recording could include the updating and extension of aide memoires so that they prompt the recording, whenever information

is available, of all stages of the process and all CCO elements. It would be helpful if recorders stated explicitly when information is not known; however, it is unlikely that this information would fit into the brief text fields currently available to record MOs. As a practical proposal this would require a not inconsiderable effort to design both the terminology, the training, and interactive guidance to get the most information, and the best quality and utility information out of police or civilian users undertaking the data collection/entry process, that made narrative sense – and simultaneously served the purposes of immediate response, administrative/legal requirements and preventative intelligence.

The review of data highlighted that information detailing the crime recording and auditing process polluted the MO fields. Where feasible, this information should be recorded elsewhere, if possible through the creation of a separate field. Similarly, training could discourage the inclusion of text that either does not inform the description of what happened or repeats information recorded elsewhere, such as ‘between stated times’.

The role of training in improving police MOs should not be overlooked: it was observed in Chapter 6 that some MO descriptions demonstrated the field author’s skill in providing a lot of information concisely.

### **10.5.2 Recommendations for techniques and technologies**

Given the likelihood that the process of MO recording is unlikely to change considerably in the near future, it seems fruitful to concentrate on methods that can get the most out of MO descriptions in their current state. Throughout the conduct of this research, a ‘wishlist’ of potential developments in techniques, that could have facilitated the approach, was formulated. Some of these techniques may already exist in unrelated fields, and may simply require the application of existing skills from other disciplines to the problem of MO analysis. Others may be dependent on future developments yet to emerge. These provided a technology roadmap of developments that could be helpful.

The coding dictionaries developed for the pre-processing analysis are free text files in a flat format that form simple list of terms that should be replaced, and the tokens that are to replace them. Using the dictionary for different crime types will likely require the revision of dictionaries and the creation of different dictionary versions. Alternative research/analysis questions may require a different set of distinctions to be made between related terms, for example, distinguishing between types of injury vs treating all injuries the same, or distinguishing between different types of verbal abuse vs simply identifying that verbal abuse has occurred. Again, these different types of question would require the creation of different versions of the coding dictionary. A more interactive version

of coding dictionaries, perhaps following a tree structure, which enabled filtering to allow different aspects of MO descriptions to be investigated, at different levels of granularity, would introduce time savings and the flexibility to apply the tools to a wider range of questions.

The word-sense disambiguation of terms, that can have multiple meanings, proved to be one of the most time consuming elements of data processing (although it was discovered that the cluster analysis handled a degree of disambiguation automatically). Given that this process relies on understanding the word through the context in which a word occurred, it would be useful to produce automated rules for disambiguation (i.e. when X occurs next to Y it means  $X_i$ ). Google Translate uses probabilities rather than language rules to identify the meaning of words in text based on their context. A similar technique adopted for police data, if it could demonstrate accuracy, would provide considerable time efficiencies leaving the analyst to concentrate on the analysis and interpretation of data.

Developments in text mining, particularly those related to the analysis of social media feeds such as Twitter, are enabling the live updating of analysis, automatically adding new data as it becomes available. This approach applied to police data could allow the live monitoring of the commission of different script tracks and allow changes to be tracked over time. However, although the results of this could be automatically updated, the need to interpret and make sense of each update would remain.

## *10.6 Limitations of the Research*

The thesis aimed to explore the utility of MO descriptions in providing intelligence for the prevention of volume crime. The analysis focused on information derived from police recorded crime data. The fundamental limitation with this approach is that recorded crimes only represent a proportion of all crimes committed. Recorded crimes are unlikely to be representative of all crimes as there will be systematic biases between reported and unreported offences. There may have been distinctive scripts, tracks and permutations within unreported crimes that cannot be identified using the methods in this thesis.

The thesis analysed two crime types, theft from the person and robbery of personal property. Part of the reason for selecting these two crime types is that they are closely related and several elements of offence commission overlap. It was necessary to select a limited number of crime categories in order to ensure the research aims were feasible. However, the cut was necessarily artificial. Other similar crimes could have been inspected that also have similarities with the offence categories analysed. For example, 'other theft' (Home Office Classification 49) can include thefts from licensed premises when target property is left on a table and thefts from shopping trolleys that have been left unattended by their owners, these script tracks bear a close resemblance to the scripts identified in the analysis, illustrating that cutting the data by administrative definitions can be artificial.

Within the police crime data, MO descriptions provide the main source of information relating to the actions of the offender and the circumstances of the offences; however, other variables in the data could have been used to elaborate these descriptions. Variables such as time and day of the week, location and the demographic characteristics of victims, and where known, offenders, also describe elements of the crime situation. The creation of clusters and the identification of script tracks may have been further improved by the inclusion of these variables in the analysis. The pre-processing of data was unable to extract details of the number of actors performing different roles in each offence description. This was valuable information which could have helped to identify additional script tracks that drew on the resources of co-offender, or manipulated the situation when a victim was accompanied by others.

The fact that crime scenes, and the MO fields which describe them, contain multiple actors presented a difficulty to analysis which was not fully addressed by the use of N-gram or 'bag-of-words' approaches. The thesis initially explored the application of quantitative narrative analysis (Franzosi, 2010) but concluded that this method was far too time consuming to apply it to large datasets. N-gram analysis presents considerable efficiency savings but does not resolve questions

about the attribution of actions to actors. Resolving these questions requires further inspection of the data, for example with sequence analysis. Therefore, the results of the initial N-gram analysis cannot be regarded as 'automatic': further investigation is normally required, although the results from initial output allow this investigation to be more focused.

The preparation of coding dictionaries for the analysis was conducted by the researcher. As noted above, this is not an objective exercise, and although guided by the conceptual frameworks, there was considerable scope to vary coding decisions, such as whether to treat two words as related or maintain the distinction. To increase confidence in the robustness of the coding dictionary, a useful exercise would have been to distribute the coding, at least in part, to multiple individuals and to investigate differences in coding decisions.

### *10.7 Original Contribution*

This thesis has explored the possibility for applying, and indeed developing, new techniques to analyse police crime data, with the specific purpose of preventing volume crime. The research has focused on MO descriptions for theft from the person and robbery of personal property and successfully demonstrated that the data analysed contains relevant information that can provide intelligence relating to the causes of these crimes. Although they were time consuming to create, the research has produced both a procedure and a set of coding dictionaries, available for future research, which can be used to process MO data, and address many of the challenges to extracting and analysing this data.

The research has demonstrated that, through cluster analysis, it is possible to rapidly identify groups of offences that are committed using similar methods and distinguish between those performed differently. The resultant clusters were judged, using three separate methods, to reflect real world patterns in the commission of the two crime types. The research also developed new means of graphically representing the significant characteristics of identified scripts by modifying the, now popular, wordcloud visualisation to weight tokens based on a statistical measure of their significance to the cluster. The thesis has also developed the use of hierarchal tree diagrams that can be used to make direct comparisons between script tracks.

The use of frameworks to interpret the analysis findings has demonstrated that these conceptual tools can facilitate systematic consideration of the entire offence process, despite the fact that limitations were identified in the neat categorisation of complex interactions. The research has questioned the strictly ordered nature of script and track hierarchies which are difficult to apply and may not reflect the complexity of track permutations.

The combined application of CCO and script frameworks to interpret the data, raises the status of the script tracks, identified from the MO descriptions, from empirical to explanatory. Prior to the analysis, the raw MO data could be described 'after-the-event descriptions' (Sutton, 2012).

Following the analysis and theoretical interpretation, the identified script tracks can perform an explanatory role by casting light on crime causation and therefore, informing the development of crime prevention interventions.

The techniques demonstrated in this research have important implications for crime prevention practice. It was noted in Chapter 2 that because the situational conditions that create opportunities for crime are crime specific, crime prevention interventions should be similarly tailored towards specific crime problems. The analysis techniques utilised in this research have allowed highly specific crime scripts and tracks to be identified. The techniques have been able to take initially specific crime types (as identified by Home Office classifications) and to dissect these into yet more detailed categories. Importantly, these detailed categories have emerged 'naturally' from the data rather than by imposing the researcher's preconceived categories. This means that very specific crime categories can be identified based on dimensions that may not be immediately apparent and about which the researcher or analyst may not yet be aware. Revealing new facets of how crimes are committed provides new potential for blocking these facets through tailored crime prevention interventions. Identifying new facets in methods of crime commission also has the potential to reveal how offenders are adapting to changing situational conditions, including interventions implemented to prevent crime.

To take one example from Chapter 8, a subset of crimes were identified that targeted vehicles, within this subset of crimes a track was identified that involved robberies of vehicles that were in use on the public road network. A contrasting track was also identified which involved robberies of vehicles from homes where the cars had been parked. Labelling both of these distinct tracks simply as 'carjacking' ignores the key differences between the situational contexts in which these tracks are conducted and the methods used to complete them. To be effective, interventions designed to prevent these tracks need to be tailored to the specifics of each MO, not least because there is potential for the two tracks to be dependent – with one track becoming more popular as another is made harder to commit.

This thesis has also made an important contribution to the development of theoretical frameworks. Ekblom and Gill (2015) have called for improved links between static causal models and dynamic script frameworks. This thesis has explicitly linked the two and jointly applied them in the interpretation of data. Using CCO and scripts together has highlighted the interdependency of

offenders' choice of method with the situation. Consequently offender methods are not completely of their own choosing while situations are not static structures, being subject to manipulation by the actors that form an integral part of them.

Finally, theoretical insights drawn from the analysis of crime scripts have substantiated explanations of crime that focus on factors in the situation. However, particularly through the analysis of research literature, the thesis provides support for calls to enrich the view of the offender in these theories by better explaining offender motivations, provocations and capacities to commit crimes.

### *10.8 Future Research*

The assessment of MO descriptions as a potential source of intelligence for the prevention of volume crime, along with the development of methods to extract, analyse and interpret this intelligence, opens up several avenues for future research. The research has focused on theft from the person and robbery of personal property and there is clear potential to apply the techniques to a wider range of crime types. Analysis of MO descriptions for different crime types might run into different challenges to those identified with theft from the person and robbery, there may be more missing or incomplete fields and the terminology used may be subject to greater ambiguity. Other crime types may involve a greater number script scenes, involve longer temporal sequences and/or a greater number of actors; all issues that could present qualitatively different challenges to the techniques adopted in this thesis.

In the analysis of the crime types in this thesis, previous research literature was used to provide a context for, and for filling gaps in, the knowledge extracted from MO fields. It would be interesting to explore the success of MO analysis for crime types about which less is known, thus where wider knowledge to contextualise the findings is not available.

In addition to using the coding dictionaries with different crime types, their development could be improved by applying them in different locations to identify the impact of local variations in vocabulary use. This would also allow the exploration of local variations in scripts and the situational factors that shape them.

The analytical techniques have the potential to investigate changes in scripts over time. This research could help to spot the emergence of new techniques and the co-evolution of scripts with attempts to prevent them. This could be conducted by repeating cluster analysis over time or by saving cluster dimensions and testing whether the fit of the cluster solutions remains stable, improves or diminishes over time. To test the ability of clusters to identify change over time, it

would be useful to identify data in which a change is known to have occurred and determine whether the cluster analysis can successfully identify this change.

This thesis has concentrated on police crime data as it is recorded in police crime information systems. England and Wales police forces have adopted online reporting for a limited range of offences and incidents. These encourage members of the public to formulate crime reports which include a free text field to describe the offence. Future research could investigate whether it is possible to, and whether there is any utility in, extending the methods use in this thesis to these data.

It has been identified that previous research using data on offender MOs has had a detection focus. However, there is scope to consider whether the techniques adopted for this thesis could be extended to further improve the investigatory use of MO descriptions and fuller MO material in more detailed case files.

MO descriptions provide an account of how offenders have acted upon conditions presented by the crime situation. It was argued above that analysing MOs through criminological frameworks can transform MO data from mere description to the basis for criminological explanation. However, to better understand the reasons for a course of action taken by an offender it would be helpful to understand the situational conditions in which offenders are acting. To explore this further future research would benefit from the inclusion of data relating to the situational context of offending. Police recorded crime data includes additional variables that can shed light on these situational conditions, commonly available variables (although there is variation between different forces) include:

- Timing of the offence, date, time of day, day of week
- Location of offence including x and y co-ordinates
- Number of offenders and victims involved in the offence
- Offender and victim demographic characteristics
- Details of property stolen

In addition to variables available within police data the availability of x and y co-ordinates and time/date variables would enable the linking of MO data with other data sources. Data that have previously been used to understand offender decision making that could be applied to understanding specific scripts could include:

- Socio-economic characteristics of the offence location

- Weather conditions
- Special events occurring at the time of the offence
- Road layout and flows of pedestrians/vehicles
- Presence of specific 'crime attractors' e.g. schools, transport hubs

The inclusion of text relating to data audits was a nuisance in the extraction of data from the MO descriptions. However, the techniques adopted to analyse scripts could also be used to focus the audit process. For example, the validation of clusters in Chapter 9 identified some mismatching between the script track and the crime classification. Using analytical techniques to identify anomalies in crime classification could help to identify a sample of crimes MOs that should be audited. The approach could be adopted to assess whether crime specific flags in the data have been properly assigned. This might include tags for hate crime, alcohol related offences and domestic abuse.

Finally, this thesis has sought to adopt an applied approach and explore whether the methods adopted offer a practical approach that could be adopted by professionals analysing crime data and those using the results of analysis. It was suggested that elements of data processing and the need to interpret the analysis findings may mean that the methods remain too time consuming for use in routine practice. However, this assumption has been made without consultation with the practitioners themselves. A valuable stream of research would be to investigate how practitioners with responsibility for analysing crime data interact with data outputs such as those produced in this thesis. How do they approach the interpretation of findings? Would aide memoires built around conceptual frameworks prove helpful in these interpretations and can analytical outputs be improved to facilitate their integration into practice? It would also be useful to explore how professionals engaged in formulating crime prevention strategies interpret and react to the types of findings produced in this thesis. This could explore how professionals integrate analytical findings with local and operational knowledge. How do findings influence the identification of opportunities to block crime? How does 'crime side' intelligence mesh with 'intervention side' intelligence? (Hirschfield, 2005).

Police MO fields have been described as 'polluted' (Cox & Kusters, 2006) by errors, ambiguities and irrelevant data. While this may be a fair description, this thesis has shown that from this, far from perfect, data it is possible to identify regularities in offence commission with a degree of confidence that these regularities reflect real-life patterns. This has been achieved by conducting a thorough review of the problems in the data and then selecting appropriate methods, from text analysis disciplines, to minimise the impact of these issues. These techniques have addressed many of the

challenges of analysing natural language. However, two particularly challenging issues remain: 1) the incompleteness of MO fields, which inevitably do not provide data on preparatory stages, nor offender motivations; and 2) the inability to attribute actions to actors via automated methods.

The cluster analysis provided a model outlining alternative tracks for committing theft and robbery from the person. The thesis has shown that this model provides more than an 'excellent description of the data' (Sutton, 2012). Previous criminology relevant research using MO data has focused on crime detection and consequently has not been interpreted through crime prevention theories. Thus, while previous research successfully identified patterns, they were not linked to theoretical explanations of the crime event. With its focus on extracting intelligence for crime prevention, this latter step has been taken in this thesis with the assistance of CCO and script frameworks. The frameworks have both methodological and conceptual importance. As a method, the frameworks ensured that data analysis and interpretation was systematic. As conceptual frameworks they formed an explicit link to situational crime theories to reveal intelligence relevant to the causal factors that create situational crime opportunities. Importantly, these situational opportunities were not presented as objective or static. To the contrary, the thesis has argued that investigating situational opportunities through data on *how* offences are committed reveals the dynamics through which offenders and other actors interact with each other and with the environment to discover, shape, respond and cope with crime opportunities in the situation.

## References

- Abelson, R. P. (1981). Psychological Status of the Script Concept. *American Psychologist*, 36(7), 715–729.
- Adderley, R. W., & Musgrove, P. (2001a). Data mining case study: Modelling the behaviour of offenders who commit serious sexual assaults. *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 215–220.
- Adderley, R. W., & Musgrove, P. (2001b). Police crime recording and investigation systems - A user's view. *Policing: An International Journal of Police Strategies & Management*, 24(1), 100–114.
- Adderley, R. W., & Musgrove, P. (2003). Modus operandi modelling of group offending: a data mining case study. *International Journal of Police Science and Management*, 5(4), 265–276.
- Adderley, R. W., & Musgrove, P. B. (1999). Data mining at the West Midlands Police: A study of bogus official burglaries. *British Computer Society Specialist Group on Expert Systems, ES99*, 91–203.
- Alarid, L. F., Burton, V. S., & Hochstetler, A. L. (2009). Group and solo robberies: Do accomplices shape criminal form? *Journal of Criminal Justice*, 37(1), 1–9.
- Aldenerfer, M. S., & Blashfield, R. K. (1984). *Cluster Analysis*. London: Sage.
- Asthana, A. (2008, July 13). Students fall prey to gang initiation rites. *The Observer*. Retrieved from <http://www.theguardian.com/uk/2008/jul/13/ukcrime.highereducation>
- Audit Commission. (2007). *Police data quality 2006/07*. London.
- Bache, R., Crestani, F., Canter, D. V., & Youngs, D. (2008). A Language Modelling Approach to Linking Criminal Styles with Offender Characteristics. *Lecture Notes in Computer Science*, 5039, 257–268.
- Bache, R., Crestani, F., Canter, D. V., & Youngs, D. (2010). A Language Modelling Approach to Linking Criminal Styles with Offender Characteristics. *Data & Knowledge Engineering*, 69(3), 303–315.
- Barker, M., Geraghty, J., Webb, B., Key, T., & Britain, G. (1993). *The Prevention of Street Robbery*. Crime Prevention Unit Series, Paper No.44. London: Home Office.
- Beauregard, E., Proulx, J., Rossmo, D. K., Leclerc, B., Allaire, J.F. (2007). Script Analysis of the Hunting Process of Serial Sex Offenders. *Criminal Justice and Behavior*, 34(8), 1069–1084.
- Bennell, C., & Canter, D. V. (2002). Linking commercial burglaries by modus operandi: tests using regression and ROC analysis. *Science & Justice*, 42(3), 153–164.
- Bennell, C., & Jones, N. J. (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.
- Bennett, T., & Wright, R. T. (1984). *Burglars on Burglary: Prevention and the Offender*. Brookfield VT: Gower.

- Bernasco, W. (2009). Foraging strategies of Homo Criminalis: Lessons from Behavioral Ecology. *Crime Patterns and Analysis*, 2(1), 5–16.
- Bernasco, W. (2010). A sentimental journey to crime: Effects of residential history on crime location choice. *Criminology*, 48(2), 389-416.
- Bernasco, W., & Block, R. (2009). Where Offenders Choose To Attack: a Discrete Choice Model of Robberies in Chicago. *Criminology*, 47(1), 93–130.
- Bernasco, W., & Block, R. (2011). Robberies in Chicago: A Block-Level Analysis of the Influence of Crime Generators, Crime Attractors, and Offender Anchor Points. *Journal of Research in Crime and Delinquency*, 48(1), 33–57.
- Bernasco, W., Block, R., & Ruiters, S. (2013). Go where the money is: Modeling street robbers' location choices. *Journal of Economic Geography*, 13(1), 119–143.
- Bernasco, W., & Luykx, F. (2003). Effects of Attractiveness, Opportunity and Accessibility to Burglars on Residential Burglary Rates of Urban Neighborhoods. *Criminology*, 41(3), 981–1002.
- Bernasco, W., & Nieuwebeerta, P. (2004). How Do Residential Burglars Select Target Areas? A New Approach to the Analysis of Criminal Location Choice. *British Journal of Criminology*, 45(3), 296–315.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python*. Sebastopol, CA: O'Reilly Media.
- Bloomfield, B. P., Latham, Y., & Vurdubakis, T. (2010). Bodies, Technologies and Action Possibilities: When is an Affordance? *Sociology*, 44(3), 415–433.
- Borrion, H. (2013). Quality assurance in crime scripting. *Crime Science*, 2(1), 6. doi: 10.1186/2193-7680-2-6
- Bouchet-Valet, M. (2015). *Package "Snowball C."* Retrieved from <https://r-forge.r-project.org/projects/r-temis/>
- Bouhana, N. (2013). The reasoning criminal vs. Homer Simpson: Conceptual challenges for crime science. *Frontiers in Human Neuroscience*, 7, 682. doi: 10.3389/fnhum.2013.00682.
- Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, 11(2), 177–220.
- Bowers, K. J., & Johnson, S. D. (2004). Who Commits Near Repeats? A Test of the Boost Explanation. *Western Criminology Review*, 5(3), 12–24.
- Brantingham, P. L., & Brantingham, P. J. (1975). Residential Burglary and Urban Form. *Urban Studies*, 12(3), 273–284.
- Brantingham, P. L., & Brantingham, P. J. (1993). Environment, Routine and Situation Toward a Pattern Theory of Crime. In M. Felson (Ed.), *Routine Activity and Rational Choice, Advances in Criminological Theory, Volume 5*. (pp. 259–294). New Jersey: Transaction Publishers.

- Brayley, H., Cockbain, E., & Laycock, G. (2011). The Value of Crime Scripting: Deconstructing Internal Child Sex Trafficking. *Policing*, 5(2), 132–143.
- Burt, C. (1940). *The factors of the mind: An Introduction to Factor-analysis in Psychology*. British Journal of Educational Psychology. London: University of London Press.
- Byrne, D. (1998). *Complexity Theory and the Social Sciences*. London: Routledge.
- Byrne, D. (2009). Working within a complexity frame of reference – the potential of “integrated methods” for understanding transformation in complex social systems, Contribution towards the Communication for Social Change Consortium’s paper for UNAIDS (July), 1–6. Retrieved from [http://www.communicationforsocialchange.org/pdfs/working%20withn%20a%20complexity%20frame\\_d%20byrne.pdf](http://www.communicationforsocialchange.org/pdfs/working%20withn%20a%20complexity%20frame_d%20byrne.pdf)
- Byrne, D., & Uprichard, E. (2009). Editors’ Introduction: Cluster Analysis. In E. Uprichard & D. Byrne (Eds.), *Cluster Analysis* (pp.vii-xii). London: Sage.
- Canter, D. V, Bennell, C., Alison, L. J., & Reddy, S. (2003). Differentiating sex offences: a behaviorally based thematic classification of stranger rapes. *Behavioral Sciences & the Law*, 21(2), 157–74.
- Carroll, J., & Weaver, F. (1986). Shoplifters’ Perceptions of Crime Opportunities: A Process-Tracing Study. In R. V Clarke & D. B. Cornish (Eds.), *The Reasoning Criminal* (pp.19-38). New York City: Springer-Verlang.
- Chainey, S. (2014). *Hypothesis Testing Crime Analysis*. London: UCL Jill Dando Institute of Security and Crime Science.
- Chainey, S., & Tompson, L. (2008). *Crime Mapping Case Studies: Practice and Research*. Chichester, UK: Wiley.
- Chaplin, R., Flatley, J., & Smith, K. (2011). *Crime in England and Wales 2010/11*. London: Home Office.
- Chau, M., Xu, J. J., & Chen, H. (2002). Extracting meaningful entities from police narrative reports. In *Proceedings of the 2002 Annual National Conference on Digital Government Research*, 1–5.
- Cherbonneau, M., & Copes, H. (2005). Drive it Like You Stole It: Auto Theft and the Illusion of Normalcy. *British Journal of Criminology*, 46 (10), 193–211.
- Clare, J., Fernández, J., & Morgan, F. (2009). Formal evaluation of the impact of barriers and connectors on residential burglars’ macro-level offending location choices. *Australian & New Zealand Journal of Criminology*, 42, 139-158.
- Clarke, R. V. (1997). *Situational Crime Prevention: Successful Case Studies*. Albany, NY: Harrow and Heston.
- Clarke, R. V. (1999). *Hot Products: understanding, anticipating and reducing demand for stolen goods*. *Police Research Series*, (Vol. Paper 112). London: Police and Reducing Crime Unit, Research Development and Statistics Directorate, Home Office.

- Clarke, R. V. (2005). Seven misconceptions of situational crime prevention. In *Handbook of Crime Prevention and Community Safety* (pp. 39–70). Cullompton, Devon: Willan Publishing.
- Clarke, R. V. (2012). Opportunity makes the thief. Really? And so what? *Crime Science*, 1(1), 3. doi: 10.1186/2193-7680-1-3
- Clarke, R. V., & Cornish, D. B. (1985). Modelling offenders' decisions: A framework for research and policy. *Crime & Justice*, 6, 147–185.
- Clarke, R. V., & Eck, J. E. (2005). *Crime analysis for problem solvers*. Washington, DC.
- Clarke, R. V., & Cornish, D. B. (2004). Rational Choice. In R. Paternoster & R. Bachman (Eds.), *Explaining criminals and crime: Essays in contemporary criminological theory* (pp. 23–42). Los Angeles: Roxbury.
- Clatworthy, J., Buick, D., Hankins, M., Weinman, J., & Horne, R. (2005). The use and reporting of cluster analysis in health psychology: A review. *British Journal of Health Psychology*, 10(3), 329–358.
- Cloward, R. A., & Ohlin, L. E. (1960). *Delinquency and Opportunity: A Theory of Delinquent Gangs*. New York: Free Press. New York: Free Press.
- Cocx, T. K., & Kosters, W. K. (2006). A distance measure for determining similarity between criminal investigations. *Lecture Notes in Artificial Intelligence*, (4065), 511–525.
- Cohen, L. E., & Felson, M. (1979). Social Change and Crime Rate Trends: A Routine Activity Approach. *American Sociological Review*, 44(4), 588–608.
- Cohen, L. E., Vila B., & Machalek, R. (1995). Expropriative Crime and Crime Policy: An Evolutionary Ecological Analysis. *Studies in Crime and Crime Prevention*, 4(2), 197–219.
- Collins, B., & Mansell, R. (2004). *Cyber Trust and Crime Prevention: A Synthesis of the State-of-the-Art Science Reviews*. London: Office of Science and Technology.
- Copes, H., & Cherbonneau, M. (2006). The Key to Auto Theft: Emerging Methods of Auto Theft from the Offenders' Perspective. *British Journal of Criminology*, 46, 917–934.
- Copes, H., Hochstetler, A., & Cherbonneau, M. (2012). Getting the Upper Hand: Scripts for Managing Victim Resistance in Carjackings. *Journal of Research in Crime and Delinquency*, 49(2), 249–268.
- Copes, H., & Vieraitis, L. M. (2009a). Bounded rationality of identity thieves: Using offender-based research to inform policy. *Criminology & Public Policy*, 8(2), 237–262.
- Copes, H., & Vieraitis, L. M. (2009b). Understanding Identity Theft: Offenders' Accounts of Their Lives and Crimes. *Criminal Justice Review*, 34(3), 329–349.
- Cornish, D. B., & Clarke, R. V. (1986). *The Reasoning Criminal*. New York City: Springer-Verlag.
- Cornish, D. B., & Clarke, R. V. (1987). Understanding Crime Displacement: An Application of Rational Choice Theory. *Criminology*, 25, 933–947.
- Cornish, D. B., & Clarke, R. V. (1994). The Procedural Analysis of Offending and its Relevance for Situational Prevention. In R.V. Clarke (Ed.), *Crime Prevention Studies* (pp. 151–196). Monsey,

New York: Criminal Justice Press.

- Cornish, D. B., & Clarke, R. V. (2002). Analysing Organized Crimes. In A. R. Piquero & S. G. Tibbetts (Eds.), *Rational choice and criminal behavior: recent research and future challenges* (pp. 41–63). New York: Routledge.
- Cornish, D. B., & Clarke, R. V. (2003). Opportunities, Precipitators and Criminal Decisions: A Reply to Wortley's Critique of Situational Crime Prevention. In M. Smith J & D. Cornish B (Eds.), *Theory for Practice in Situational Crime Prevention* (pp. 41–96). New York: Criminal Justice Press.
- Cornish, D. B., & Clarke, R. V. (2008). The rational choice perspective. In R. K. Wortley & L. Mazerolle (Eds.), *Environmental Criminology and Crime Analysis* (pp. 21–44). Cullompton, Devon: Willan Publishing.
- Coupe, R. T., & Blake, L. (2005). The effects of patrol workloads and response strength on arrests at burglary emergencies. *Journal of Criminal Justice*, 33(3), 239–255.
- Coupe, R. T., & Blake, L. (2006). Daylight and darkness targeting strategies and the risks of being seen at residential burglaries. *Criminology*, 44(2), 431–464.
- Cozens, P. M., Saville, G., & Hillier, D. (2005). Crime prevention through environmental design (CPTED): a review and modern bibliography. *Property Management*, 23(5), 328–356.
- Cromwell, P. F., & Olson, J. N. (2004). *Breaking and Entering: Burglars on Burglary*. Belmont, CA: Wadsworth.
- Crowston, K., Allen, E. E., & Heckman, R. (2012). Using natural language processing for qualitative data analysis. *International Journal of Social Research Methodology*, 15(6), 523-543.
- D'Alessio, S., & Stolzenberg, L. (1990). A crime of convenience: The Environment and Convenience Store Robbery. *Environment and Behaviour*, 22(2), 255–271.
- De Haan, W., & Vos, J. (2003). A crying shame: The over-rationalized conception of man in the rational choice perspective. *Theoretical Criminology*, 7(1), 29–54.
- Deakin, J., Smithson, H., Spencer, J., & Medina-Ariza, J. (2007). Taxing on the Streets: Understanding the Methods and Process of Street Robbery. *Crime Prevention and Community Safety*, 9(1), 52–67.
- Di Gregorio, S. (2000). Using NVIVO for your literature review. In *Strategies in Qualitative Research: Issues and Results from Analysis Using QSR NVivo and NUD\* IST Conference at the Institute of Education, London* (pp. 29-30).
- Donkin, S., & Wellsmith, M. (2006). Cars Stolen in Burglaries: The Sandwell Experience. *Security Journal*, 19(1), 22–32.
- Dunne, W. (2012). *The Dramatic Writer's Companion* (Vol. XXXIII). Chicago: The University of Chicago Press.
- Eck, J. E. (1993). The Threat of Crime Displacement. *Criminal Justice Abstracts*, 25(3), 527–546.
- Eck, J. E. (1994). *Drug markets and drug places: A case-control study of the spatial structure of illicit*

- drug dealing*. Ph.D dissertation, University of Maryland.
- Eck, J., & Weisburd, D. (1995). Crime places in crime theory. In J. E. Eck & D. Weisburd (Eds.), *Crime and place, Crime Prevention Studies* (pp. 1–34). Monsey, New York: Criminal Justice Press.
- Ekblom, P. (2001). The Conjunction of Criminal Opportunity: For Crime Reduction Toolkits. *Home Office*, 1–22. Retrieved from <http://ualresearchonline.arts.ac.uk/1104/>
- Ekblom, P. (2002). Future Imperfect: Preparing for the Crimes to Come. *Criminal Justice Matters*, 46(Winter 2001/02), 38–40.
- Ekblom, P. (2003). Organised Crime and the Conjunction of Criminal Opportunity Framework. In A. Edwards & P. Gill (Eds.), *Transnational Organised Crime: Perspectives on Global Security* (pp. 241–263). London: Routledge.
- Ekblom, P. (2007). Making Offenders Richer. In G. Farrell, K. J. Bowers, & S. D. Johnson (Eds.), *Imagination for Crime Prevention: Essays in Honour of Ken Pease*. Devon, UK:Willan Publishing.
- Ekblom, P. (2010). The Conjunction of Criminal Opportunity Theory. In B. S. Fisher & S. P. Lab (Eds.), *Encyclopedia of Victimology and Crime Prevention*. Sage:London.  
doi: <http://dx.doi.org/10.4135/9781412979993.n53>
- Ekblom, P. (2011). *Crime Prevention, Security and Community Safety Using the 5Is Framework* (First). Basingstoke: Palgrave:Macmillan.
- Ekblom, P. (2012a). Conceptual and Methodological Explorations in Affordance and Counter Terrorism. In M. Taylor & P. Currie (Eds.), *Terrorism and Affordance*. London: Continuum.
- Ekblom, P. (2012b). Happy returns: ideas brought back from situational crime prevention’s exploration of design against crime. In N. Tilley & G. Farrell (Eds.), *The Reasoning Criminologist: Essays in Honour of Ronald V. Clarke* (pp.38-49). Abingdon, Oxon: Routledge.
- Ekblom, P., & Gill, M. (2015). Rewriting the Script Cross-Disciplinary Exploration and Conceptual Consolidation of the Procedural Analysis of Crime. *European Journal of Criminal Policy and Research*. Advance online publication. Doi: 10.1007/s10610-015-9291-9.
- Ekblom, P., & Sidebottom, A. (2008). What Do You Mean, “Is It Secure?” Redesigning Language to be Fit for the Task of Assessing the Security of Domestic and Personal Electronic Goods. *European Journal on Criminal Policy and Research*, 14(1), 61–87.
- Ekblom, P., & Tilley, N. (2000). Going Equipped, Criminology, Situational Crime Prevention and the Resourceful Offender. *British Journal of Criminology*, 40, 376–398.
- Estivill-castro, V., & Lee, I. (2001). Data Mining Techniques for Autonomous Exploration of Large Volumes of Geo-referenced Crime Data. In *Proceedings of the 6th International Conference on Geocomputation*.
- Everitt, B. S., & Dunn, G. (2001). *Applied Multivariate Data Analysis* (Second Edition). London: Arnold.
- Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis*. Chichester, UK: John Wiley &

Sons, Ltd.

- Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996). *Advances in knowledge discovery and data mining*. Cambridge, MA:MIT Press.
- Feeney, F. (1986). Robbers as decision-makers. In D. B. Cornish & R. V. Clarke (Eds.), *The Reasoning Criminal: Rational Choice Perspectives on Offending* (pp. 53–71) New York:Springer-Verlag.
- Feinerer, I., Hornik, K., & Wallace, M. (2015). *Package "Wordnet."* Retrieved from <https://cran.r-project.org/web/packages/wordnet/wordnet.pdf>
- Felson, M. (1986). Linking Criminal Choices, Routine Activities, Informal Control and Criminal Outcomes. In D. B. Cornish & R. V. Clarke (Eds.), *The Reasoning Criminal. Rational Choice Perspectives on Offending* (pp. 119–128). New York: Springer-Verlag.
- Felson, M. (1988). *Crime and Everyday Life*. Thousand Oaks, CA: Pine Forge Press.
- Felson, M. (1995). Those Who Discourage Crime. In J. E. Eck & D. Weisburd (Eds.), *Crime and Place, Crime Prevention Studies*, (Vol.4 pp. 53–66). Monsey, NY: Criminal Justice Press.
- Felson, M. (2008). Routine Activity Theory Approaches. In R. K. Wortley & L. Mazerolle (Eds.), *Environmental Criminology and Crime Analysis* (pp.70-77). Abingdon, Oxon: Routledge.
- Felson, M. (2009). *Crime and Nature*. Thousand Oaks: Sage.
- Felson, M., & Clarke, R. V. (1998). *Opportunity Makes the Thief: Practical Theory for Crime Prevention. Police Research Series*. London: Home Office.
- Felson, M., & Cohen, L. E. (1980). Human ecology and crime: A routine activity approach. *Human Ecology*, 8(4), 389–406.
- Felson, R. B., Baumer, E. P., & Messner, S. F. (2000). Acquaintance Robbery. *Journal of Research in Crime and Delinquency*, 37(3), 284–305.
- Francis, L., & Flynn, M. (2010). Text Mining Handbook. *Casualty Actuarial Society*, E-Forum, Spring, 1–61.
- Frantzi, K., Ananiadou, S., & Mima, H. (2000). Automatic recognition of multi-word terms. *International Journal of Digital Libraries*, 3(3), 117–123.
- Franzosi, R. (2010). *Quantitative Narrative Analysis*. London: Sage.
- Franzosi, R., Doyle, S., McClelland, L. E., Putnam Rankin, C., & Vicari, S. (2013). Quantitative narrative analysis software options compared: PC-ACE and CAQDAS (ATLAS.ti, MAXqda, and NVivo). *Quality & Quantity*, 47(6), 3219–3247.
- Gamman, L. (2001). *In the Bag, CD ROM Design Resource*. London: Central Saint Martins College of Art & Design. Retrieved from <http://www.designagainstcrime.com/projects/in-the-bag/>
- Garland, D. (2000). The Culture of High Crime Societies. *British Journal of Criminology*, 40(3), 347–375.
- Garson, D. (2014). *Cluster Analysis*. Asheboro, NC: Statistical Associates Publishers.

- Garwood, J. (2009). A quasi-experimental investigation of self-reported offending and perception of criminal opportunity in undergraduate students. *Security Journal*, *06*(1), 1–15.
- Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.
- Gill, M. (2000). *Commercial Robbery*. London: Blackstone.
- Glymour, C., Madigan, D., Pregibon, D., & Smyth, P. (1997). Statistical themes and lessons for data mining. *Data Mining and Knowledge Discovery*, *1*, 11–28.
- Goodhew, G. W., Cammock, P. A., & Hamilton, R. T. (2008). The management of poor performance by front-line managers. *Journal of Management Development*, *27*(9), 951–962.
- Gottfredson, M., & Hirschi, T. (1990). *The General Theory of Crime*. Stanford, CA: Stanford University Press.
- Graesser, A. C., Gordon, S. E., & Sawyer, J. D. (1979). Recognition memory for typical and atypical actions in scripted activities: Tests of a script pointer and tag hypothesis. *Journal of Verbal Learning and Verbal Behavior*, *18*(3), 319–332.
- Grant, M.J. & Booth, A. (2008). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, *26*, 91–108.
- Greene, J. O. (1989). The stability of nonverbal behaviour: An action-production approach to problems of cross-situational consistency and discriminativeness. *Journal of Language and Social Psychology*, *8*(3-4), 193–220.
- Groff, E., & Mcewen, T. (2008). *Modeling the Dynamics of Street Robberies*. Alexandria, Virginia: Institute for Law and Justice.
- Hakim, S., Rengert, G. F., & Shachmurove, Y. (2001). Target search of burglars: A revised economic approach. *Papers in Regional Science*, *80*, 121–137.
- Hallsworth, S. (2005). *Street Crime*. Cullompton, Devon: Willan Publishing.
- Hamilton-Smith, N. (2002). Anticipated Consequences: Developing a Strategy for the Targeted Measurement Of Displacement And Diffusion Of Benefits. In N. Tilley (Ed.), *Evaluation for Crime Prevention* (Vol. 14, pp. 11–52). Cullompton, Devon: Willan Publishers.
- Haracopos, A., & Hough, M. (2005). *Drug Dealing in Open-Air Markets. Policing: An International Journal of Police Strategies & Management*. US Department of Justice.
- Harper, D. W. (2006). The Tourist and His Criminal: Patterns in Street Robbery. In Y. Mansfield & A. Pizam (Eds.), *Tourism, Security and Safety: From Theory to Practice* (pp. 125-137). New York: Routledge.
- Harper, D. W., Khey, D. N., & Nolan, G. M. (2013). Spatial Patterns of Robbery at Tourism Sites : A Case Study of the Vieux Carré in New Orleans. *American Journal of Criminal Justice*, *38*, 589–601.
- Harrington, V., & Mayhew, P. (2001). *Mobile phone theft*. Home Office research studies. Home Office:London.

- Harvey, C. D. L. (2009). Complexity and Case. In D. Byrne & C. C. Ragin (Eds.), *The SAGE Handbook of Case-Based Methods* (pp. 15–39). London: Sage.
- Hauck, R. V., Atabakhsb, H., Ongvasith, P., Gupta, H., & Chen, H. (2002). Using COPLINK to analyze criminal justice data. *Computer*, 35(3), 30–37.
- Hayward, K. (2007). Situational Crime Prevention and its Discontents: Rational Choice Theory vs the “Culture of Now.” *Social Policy and Administration*, 41(3), 232–250.
- Hepenstal, S., & Johnson, S. D. (2010). The concentration of cash-in-transit robbery. *Crime Prevention and Community Safety*, 12(4), 263–282.
- Hirschfield, A. F. (2005). Analysis for Intervention. In N. Tilley (Ed.), *The Handbook of Crime Prevention and Community Safety*. Cullompton, Devon: Willan Publishing.
- Hirschfield, A. F., & Bowers, K. (2003). *Mapping and analysing crime data: Lessons from research and practice*. London: Taylor Francis.
- Hirschfield, A. F., & Bowers, K. J. (1997). The effect of social cohesion on levels of recorded crime in disadvantaged areas. *Urban Studies*, 34(8), 1275–1295.
- Hirschi, T. (1969). *Causes of Delinquency*. Berkeley: University of California Press.
- Hirschi, T., & Gottfredson, S. D. (1995). A General theory and the life-course perspective. In *Studies in Crime and Crime Prevention* (Vol. 4, pp. 131–142).
- HMIC. (2012). *The Crime Scene: A Review of Police Crime and Incident Reports*. London.
- HMIC. (2014). *Crime recording : A matter of fact. An interim report of the inspection of crime data integrity in police forces in England and Wales*. London.
- Hochstetler, A. (2001). Opportunities and Decisions: Interactional Dynamics in Robbery and Burglary Groups. *Criminology*, 39(3), 737–764.
- Hochstetler, A., Copes, H., & DeLisi, M. (2002). Differential association in group and solo offending. *Journal of Criminal Justice*, 30(6), 559–566.
- Hollis-Peel, M. E., Reynald, D. M., Bravel, M., Ellfers, H., & Welsh, B. C. (2011). Guardianship for crime prevention: a critical review of the literature. *Crime Law and Social Change*, 56, 53–70.
- Holt, T., & Spencer, J. (2005). A Little Yellow Box: The Targeting of Automatic Teller Machines as a Strategy in Reducing Street Robbery. *Crime Prevention and Community Safety: An International Journal*, 7(2), 15–28.
- Home Office. (2013). *Counting Rules for Robbery*. London.
- Home Office. (2015a). *Counting Rules for Theft*. London.
- Home Office. (2015b). *Crime Recording General Rules*. Retrieved from [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/452294/count-general-august-2015.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/452294/count-general-august-2015.pdf)
- Home Office. (2015c). National Crime Recording Standard. Retrieved from

<https://www.gov.uk/government/publications/counting-rules-for-recorded-crime>

- Hopkins Burke, R. (2009). *An Introduction to Criminological Theory* (3rd ed.). Cullumpton, Devon: Willan Publishing.
- Hornik, K., Feinerer, I., Hornik, K., & Meyer, D. (2008). Text Mining Infrastructure in R. *Journal of Statistical Software*, 25(5), 1–54.
- Irvine, J., Miles, I., & Evans, J. (1979). Introduction. In J. Irvine, I. Miles, & J. Evans (Eds.), *Demystifying Social Statistics* (pp. 1–9). London: Pluto Press.
- Jacobs, B. A. (2010). Serendipity in Robbery Target Selection. *British Journal of Criminology*, 50, 514–539.
- Jacobs, B. A. (2012). Carjacking and Copresence. *Journal of Research in Crime and Delinquency*, 49, 471–488.
- Jacobs, B. A., Topalli, V., & Wright, R. T. (2000). Managing Retaliation: Drug Robbery and Informal Sanction Threats. *Criminology*, 38(1), 171–198.
- Jacobs, B. A., & Wright, R. T. (2008). Moralistic Street Robbery. *Crime & Delinquency*, 54(4), 511–531.
- Jacobs, B. A., & Wright, R. T. (2010). Bounded rationality, retaliation, and the spread of urban violence. *Journal of Interpersonal Violence*, 25(10), 1739–66.
- Jacobs, B. A., Wright, R. T., & Topalli, V. (2003). Carjacking, Street Life and Offender Motivation. *British Journal of Criminology*, 43, 673–688.
- Jeffrey, C. R. (1971). *Crime Prevention through Environmental Design*. Beverley Hills, CA: Sage.
- Johnson, S. D., & Bowers, K. J. (2004). The burglary as clue to the future the beginnings of prospective hot-spotting. *European Journal of Criminology*, 1(2), 237–255.
- Johnson, S. D., & Bowers, K. J. (2013). How guardianship dynamics may vary across the street network : A case study of residential burglary. In S. Ruiter, W. Bernasco, W. Huisman, & G. Bruinsma (Eds.), *Liber amicorum voor Henk Elffers*. Amsterdam: NCSR.
- Johnson, S. D., Bowers, K. J., Gamman, L., Mamerow, L., & Warne, A. (2010). *Theft of Customers' Personal Property in Cafes and Bars*. Washington: US Department of Justice.
- Johnson, S. D., & Summers, L. (2015). Testing Ecological Theories of Offender Spatial Decision Making Using a Discrete Choice Model. *Crime & Delinquency*, 61(3), 454–480.
- Jurafsky, D., & Martin, J. H. (2008). *Speech and Language Processing*. New Jersey: Pearson.
- Katz, J. (1988). *Seductions of Crime*. New York: Basic Books.
- Kershaw, C., & Tseloni, A. (2005). Predicting Crime Rates, Fear and Disorder Based on Area Information: Evidence from the 2000 British Crime Survey. *International Review of Victimology*, 12(3), 293–311.
- Kleck, G., & DeLone, M. (1993). Victim resistance and offender weapon effects in robbery. *Journal of Quantitative Criminology*, 9(1), 55–81.

- Krahé, B., Bieneck, S., & Scheinberger-Olwig, R. (2007). Adolescents' sexual scripts: schematic representations of consensual and nonconsensual heterosexual interactions. *Journal of Sex Research, 44*(4), 316–327.
- Kriven, S., & Ziersch, E. (2007). New Car Security and Shifting Vehicle Theft Patterns in Australia. *Security Journal, 20*(2), 111–122.
- Lancia, F. (2012). *The Logic of the T-Lab Tools Explained*. Retrieved from [www.tlab.it](http://www.tlab.it)
- Lang, D. T. (2005). Aspell package for R. Retrieved from <http://aspell.sourceforge.net>
- Langer, E. J., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of“ placebo” information in interpersonal interaction. *Journal of Personality and Social Psychology, 36*(6), 635–642.
- Laycock, G. (2005). Deciding What to Do. In *Handbook of Crime Prevention and Community Safety*. Cullumpton, Devon: Willan Publishing.
- Leclerc, B., & Tremblay, P. (2007). Strategic behavior in adolescent sexual offenses against children: Linking modus operandi to sexual behaviors). *Sexual Abuse-A Journal of Research and Treatment, 19*(3), 331–332.
- Leclerc, B., & Wortley, R. (2014). *Cognition and Crime: Offender Decision-Making and Script Analyses*. Abingdon, Oxon: Routledge.
- Leclerc, B., Wortley, R. K., & Smallbone, S. (2011). Getting into the script of adult child sex offenders and mapping out situational prevention measures, *Journal of Research in Crime and Delinquency, 48*,209-23.
- Leddo, J., & Abelson, R. P. (1986). The Nature of Explanations. In J. A. Galambos, R. P. Abelson, & J. B. Black (Eds.), *Knowledge Structures* (pp. 103–122). New Jersey: Lawrence Erlbaum Associates.
- Leetaru, K. H. (2012). *Data Mining Methods for the Content Analyst: An Introduction to the Computational Analysis of Content*. New York: Routledge.
- Leigh, T., & Rethans, A. (1983). Experiences with script elicitation within consumer decision making contexts. *Advances in Consumer Research, 10*, 667–672.
- Liddy, E. D., Jorgensen, C. L., Sibert, E. E., & Yu, E. S. (1993). A sublanguage approach to Natural Language Processing for an Expert System. *Information Processing & Management, 29*(5), 633–645.
- Lin, S., & Brown, D. E. (2006). An outlier-based data association method for linking criminal incidents. *Decision Support Systems, 41*(3), 604–615.
- Luckenbill, D. F. (1981). Generating compliance: The case of robbery. *Urban Life, 10*(1), 25–46.
- Ma, L., Chen, Y., & Huang, H. (2010). AK-Modes: A weighted clustering algorithm for finding similar case subsets. *2010 IEEE International Conference on Intelligent Systems and Knowledge Engineering*.
- Maechler, M., Struyf, A., Hubert, M., Hornik, K., Studer, M., & Roudier, P. (2015). *Package “Cluster.”*

Retrieved from <https://cran.r-project.org/web/packages/cluster/cluster.pdf>

- Marcum, C. D., Higgins, G. E., & Ricketts, M. L. (2010). Potential Factors of Online Victimization of Youth: An Examination of Adolescent Online Behaviors Utilizing Routine Activity Theory. *Deviant Behavior, 31*(5), 381–410.
- Marshall, B., & Townsley, M. (2006). *Needles or needless? The application of data mining to the academic criminological community*. London: UCL Jill Dando Institute of Crime Science.
- Mativat, F., & Tremblay, P. (1997). Counterfeiting credit cards - Displacement effects, suitable offenders and crime wave patterns. *British Journal of Criminology, 37*(2), 165–183.
- Michael, S. E., Hull, R. B., & Zahm, D. L. (2001). Environmental Factors Influencing Auto Burglary: A Case Study. *Environment and Behavior, 33*(3), 368–388.
- Mischel, W. (1999). Personality coherence and dispositions in a cognitive-affective personality system (CAPS) approach. In D. Cervone & Y. Shoda (Eds.), *The coherence of personality: Social-cognitive bases of consistency, variability, and organization*. New York: The Guildford Press.
- Monteiro, R., Carrillo, J., & Santiago, A. (2010). Teacher scripts in science teaching. *Teaching and Teacher Education, 26*(6), 1269–1279.
- Morr Serewicz, M. C., & Gale, E. (2007). First-Date Scripts: Gender Roles, Context, and Relationship. *Sex Roles, 58*(3-4), 149–164.
- Nee, C., & Meenaghan, A. (2006). Expert Decision Making in Burglars. *British Journal of Criminology, 46*(5), 935–949.
- Nee, C., & Taylor, M. (2000). Examining burglars' target selection: Interview, experiment or ethnomethodology? *Psychology, Crime & Law, 6*(1), 45–59.
- Newman, G. R., & Clarke, R. V. (2003). *Superhighway Robbery: Preventing e-commerce crime*. Abingdon, Oxon: Willan Publishing.
- Newman, O. (1973). *Defensible Space*. London: Architectural Press.
- Newton, A. (2014). Activity nodes and licensed premises: Risky mixes and risky facilities. In *The International Symposium on Environmental Criminology and Crime Analysis (ECCA)*. Kerkrade, Holland.
- Norrie, A. (1986). Practical Reasoning and Criminal Responsibility: A Jurisprudential Approach. In D. Cornish & R. Clarke V (Eds.), *The Reasoning Criminal* (pp.217-228). Berlin-New York: Springer-Verlag.
- Oatley, G. C., & Ewart, B. W. (2003). Crime analysis software: "pins in maps", clustering and Bayes net prediction. *Expert Systems with Applications, 25*(4), 569–588.
- Oatley, G. C., Ewart, B. W., & Zeleznikow, J. (2006). Decision Support Systems For Police : Lessons From The Application of Data Mining Techniques to 'Soft ' Forensic Evidence. *Artificial Intelligence and Law, 14*, 35-100.
- Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.

- Pease, K. (1997). Predicting the future: The roles of routine activity and rational choice theory. In G. R. Newman, R. V. Clarke, & S. G. Shoham (Eds.), *Rational Choice and Situational Crime Prevention: Theoretical Foundations* (pp. 233–245). Dartmouth: Ashgate.
- Petrosino, A., & Brensilber, D. (2003). The motives, methods and decision making of convenience store robbers: interviews with 28 incarcerated offenders in Massachusetts. *Crime Prevention Studies*, 16, 237–263.
- Porter, M. F. (2001). Snowball: A language for stemming algorithms. Retrieved from <http://snowball.tartarus.org/texts/introduction.html>
- Poyner, B., & Webb, B. (1997). Reducing Theft from Shopping Bags in City Center Markets. In R. V. Clarke (Ed.), *Situational Crime Prevention: Successful Case Studies* (Second Ed.). New York: Harrow and Heston.
- Ratcliffe, J. H. (2008). *Intelligence-led Policing*. Cullompton, Devon: Willan Publishing.
- Rogerson, M., Smithsons, H., & Hirschfield, A. F. (2008). *Exploring Strategies for Targeting Student Robbery*. (Unpublished Report).
- Schank, C. R., & Abelson, R. P. (1977). Scripts, plans and knowledge. In P. N. Johnson-Laird & P. Cathcart Wason (Eds.), *Thinking: readings in cognitive science* (pp. 421–434). Cambridge: Cambridge University Press.
- Seifert, C. M., Hammond, K. J., Johnson, H. M., Converse, T. M., McDougal, T. F., & Vanderstoep, S. W. (1994). Case-based learning: Predictive features in indexing. *Machine Learning*, 16(1), 37–56.
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(July & October), 379–423 & 623–656.
- Shover, N., & Copes, H. (2010). Decision Making by Persistent Thieves and Crime Control Policy. In H. Barlow & S. H. Decker (Eds.), *Criminology and Public Policy: Putting Theory to Work* (pp. 128–149). Philadelphia: Temple University Press.
- Shover, N., & Honaker, D. (1992). The Socially Bounded Decision Making of Persistent Property Offenders. *Howard Journal of Criminal Justice*, 31(4), 276–293.
- Smith, C., Bowers, K. J., & Johnson, S. D. (2006). Understanding bag theft within licensed premises in Westminster. Identifying initial steps towards prevention. *Security Journal*, 19(1), 3–21.
- Smith, J. (2003). *The nature of personal robbery*. London: Home Office.
- Smith, M. J. (2003). Exploring Target Attractiveness in Vandalism: An Experimental Approach. *Crime Prevention Studies*, 16, 197–236.
- Smith, M.J. and Clarke, R. (2000). Crime and Public Transport. *Crime and Justice — A Review of Research*, 27, 169–233.
- Smyth, P. (2001). Breaking out of the Black-Box: Research Challenges in Data Mining. In *Data Mining and Knowledge Discovery Workshop (DMKD)*. Santa Barbara.

- Steinbach, M., Karypis, G., & Kumar, V. (2000). A comparison of document clustering techniques. *KDD Workshop on Text Mining*, 1–20.
- Steinbach, R., Perkins, C., Tompson, L., Johnson, S., Armstrong, B., Green, J., Grundy, C., Wilkinson, P., & Edwards, P. (2015). The effect of reduced street lighting on road casualties and crime in England and Wales : controlled interrupted time series analysis. *Journal of Epidemiology and Community Health*, 69, 1118-1124.
- Sutton, M. (2011). Opportunity is a very bad explanation for crime: A critique of Crime Science [https://www.bestthinking.com/articles/science/applied\\_science/opportunity-is-a-very-bad-explanation-for-crime-a-critique-of-crime-science](https://www.bestthinking.com/articles/science/applied_science/opportunity-is-a-very-bad-explanation-for-crime-a-critique-of-crime-science)
- Sutton, M. (2012). Opportunity Does Not Make a Thief Busting the Myth that Opportunity is a Cause of Crime. Retrieved from [https://www.bestthinking.com/articles/science/social\\_sciences/sociology/opportunity-does-not-make-the-thief-busting-the-myth-that-opportunity-is-a-cause-of-crime](https://www.bestthinking.com/articles/science/social_sciences/sociology/opportunity-does-not-make-the-thief-busting-the-myth-that-opportunity-is-a-cause-of-crime)
- Tabangin, D. R., Flores, J. C., & Emperador, N. F. (2008). Observing the correlation of crime and urban physical environment: an input to urban renewal and revitalization. *North Luzon Journal*, 2, 29–42.
- Tarlow, P. E. (2006). Crime and Tourism. In J. Wilks, D. Pendergast, & P. Leggat (Eds.), *Tourism in Turbulent Times* (pp.93-106). New York: Routledge.
- Tesch, R. (1990). *Qualitative Research: Analysis Types & Software Tools*. Press. Britsol, PA: Falmer Press.
- Tilley, N., & Laycock, G. (2002). *Working Out What to Do: Evidence-Based Crime Reduction, Crime Reduction Research Series Paper 11*. London: Home Office.
- Tompson, L. A, & Bowers, K. J. (2015). Testing time-sensitive influences of weather on street robbery. *Crime Science*, 4(8). Retrieved from <http://www.crimesciencejournal.com/content/4/1/8>
- Tompson, L.A, & Chainey, S. (2011). Profiling Illegal Waste Activity: Using Crime Scripts as a Data Collection and Analytical Strategy. *European Journal on Criminal Policy and Research*, 17(3).
- Topalli, V. (2005). Criminal expertise and offender decision-making: An experimental analysis of how offenders and non-offenders differentially perceive social stimuli. *British Journal of Criminology*, 45(3), 269–295.
- Topalli, V., Jacques, S., & Wright, R. (2015). “It takes skills to take a car”: Perceptual and procedural expertise in carjacking. *Aggression and Violent Behavior*, 20(Jan), 19–25.
- Topalli, V., Wright, R., & Fornango, R. (2002). Drug Dealers, Robbery and Retaliation: Vulnerability, Deterrence and the Contagion of Violence. *British Journal of Criminology*, 42(2), 337–351.
- Topalli, V., & Wright, R. T. (2004). Dubs and Dees, Bats ad Rims: Carjackers and Urban Violence. In D. Dabney (ed) *Crime Types: A Text Reader*, 149–69. Belmont, CA: Wadsworth.
- Tremblay, P. (1986). Designing Crime - the Short Life Expectancy and the Workings of a Recent Wave

- of Credit Card Bank Frauds. *British Journal of Criminology*, 26(3), 234–253.
- Tremblay, P., Cusson, M., & Morselli, C. (1998). Market Offenses and limits to growth. *Crime, Law and Social Change*, 29, 311–330.
- Tremblay, P., Talon, B., & Hurley, D. (2001). Body Switching and Related Adaptations In the Resale of Stolen Vehicles. Elaborations and Aggregate Crime Learning Curves. *British Journal of Criminology*, 41(4), 561–579.
- Tukey, J. W. (1980). We Need Both Exploratory and Confirmatory. *The American Statistician*, 34(1), 23–25.
- Tunnell, K. D. (1990). Choosing crime: Close your eyes and take your chances. *Justice Quarterly*, 7(4), 673–690.
- Von Goldammer, E., Paul, J., & Newbury, J. (2003). Heterarchy–Hierarchy: Two Complementary Category of Description. *Vordenker*, August(I), 1–27.
- Walsh, D. (1986). Victim Selection Procedures Among Economic Criminals: The Rational Choice Perspective. In D. Cornish & R. V Clarke (Eds.), *The Reasoning Criminal. Rational Choice Perspectives on Offending*, (pp 39–52). New York: Springer-Verlag.
- Wang, J. Z. (2002). Bank Robberies by an Asian Gang: An Assessment of the Routine Activities Theory. *International Journal of Offender Therapy and Comparative Criminology*, 46(5), 555–568.
- Weaver, F. M., & Carroll, J. S. (2008). Crime Perceptions in a Natural Setting by Expert and Novice Shoplifters, 48(4), 349–359.
- Weber, R. P. (1990). *Basic Content Analysis*. London: Sage.
- Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the Process of Sensemaking. *Organization Science*, 16(4), 409–421.
- Wikström, P.O. H. (2006a). In search of causes and explanations of crime. In R. King & E. Wincup (Eds.), *Doing Research On Crime and Justice* (pp. 117–140). Oxford: Oxford University Press.
- Wikström, P.O. H. (2006b). Individuals, settings, and acts of crime: situational mechanisms and the explanation of crime. In P.-O. H. Wikström & R. J. Sampson (Eds.), *The Explanation of crime: Context, mechanisms and development* (pp. 61–107). Cambridge: Cambridge University Press.
- Wikström, P.O. H. (2009). *Situational Action Theory*. Oxford: Oxford University Press.
- Wikström, P.O. H., & Treiber, K. H. (2009). Violence as Situational Action. *International Journal of Conflict and Violence*, 3(1), 75–96.
- Wiles, P., & Costello, A. (2000). *The “Road to Nowhere”: The Evidence for Travelling Criminals. Home Office Research Study 207*, London:Home Office.
- Willet, P. (2006). The Porter stemming algorithm: then and now. *Program: Electronic Library and Information Systems*, 40(3).
- Woodhams, J., Hollin, C. R., & Bull, R. (2007). The psychology of linking crimes: A review of the

- evidence. *Legal and Criminological Psychology*, 12, 233–249.
- 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(1), 59–85.
- Wortley, R. K. (1998). A Two-Stage Model of Situational Crime Prevention. *Studies in Crime and Crime Prevention*, 7, 173–188.
- Wortley, R. K. (2001). A Classification of Techniques for Controlling Situational Precipitators of Crime. *Security Journal*, 14(4), 63–82.
- Wortley, R. K. (2008). Situational precipitators of crime. In R. K. Wortley & B. Leclerc (Eds.), *Environmental Criminology and Crime Analysis* (pp.48-69). Cullompton, Devon: Willan Publishing.
- Wortley, R. K. (2011). *Psychological Criminology An Integrative Approach* (First). Abingdon, Oxon: Routledge.
- Wortley, R. K. (2014). *Psychological Criminology An integrative approach*. Abingdon, Oxon: Routledge.
- Wright, R. T., Brookman, F., & Bennett, T. (2006). The Foreground Dynamics of Street Robbery in Britain. *British Journal of Criminology*, 46(1), 1–15.
- Wright, R. T., & Decker, S. H. (1994). *Burglars on the Job, Streetlife and Residential Break-ins*. Boston, MA: Northeastern University Press.
- Wright, R. T., & Decker, S. H. (1997). *Armed Robbers in Action: Stickups and Street Culture*. Boston, MA: University Press of New England.
- Wright, R. T., Logie, R. H., & Decker, S.H.(1995). Criminal Expertise and Offender Decision Making: An Experimental Study of the Target Selection Process in Residential Burglary. *Journal of Research in Crime and Delinquency*, 32(1), 39–53.
- Yokota, K., & Watanbe, S. (2002). Computer-based retrieval of suspects using similarity of modus operandi. *International Journal of Police Science and Management*, 4, 5.
- Yu, G., Shao, S., & Luo, B. (2008). Mining Crime Data by Using New Similarity Measure. *2008 Second International Conference on Genetic and Evolutionary Computing*, 389–392.
- Yun, M., & Roth, M. (2008). Terrorist Hostage-Taking: Using Script Theory to Predict the Fate of the Hostage. *Studies in Conflict and Terrorism*, 31(8), 736–748.

## APPENDICES

## *Appendix 1. Example of an MO field coded according to Franzosi's semantic triplet*

*Raw MO text:*

THE IP WHO IS A 14 YEAR OLD BOY WAS STANDING WITHIN THE GROUND FLOOR AREA OF [RESIDENTIAL FLATS] [TOWN NAME] WHEN A UNKNOWN MALE APPROACHED HIM AND WAS HIT IN THE CHEST AND GRABBED IP'S MOBILE PHONE FROM HIS HAND.

*Coding with Franzosi's Triplet*

1. First each clause of the text must be split into a skeletal narrative:

**Narrative 1:** The IP was holding his mobile phone in his hand

**Narrative 2:** An unknown offender, who was male, approached the IP

**Narrative 3:** The offender hit the IP in the chest

**Narrative 4:** The IP grabbed the IP's mobile phone

2. Secondly, each skeletal narrative should be coded according to the triplet (participant, object, triplet) and characteristics identified for each element.

### **Semantic Triplet Narrative 1**

**Participant** (Actor: IP) Characteristics: (Number: one) (Gender: Male) (Age: 14)

**Process** (Verb: hold) Circumstances (Time: As Specified in Data) Space (Town:X)(Location type: Residential Flats Foyer)

**Object** (Physical object: Mobile Phone)

### **Semantic Triplet Narrative 2**

**Participant** (Actor: Offender) Characteristics: (Number: one) (Gender: Male) (Age: unknown) (Identity: unknown to victim?)

**Process** (Verb: approach)

**Object** (Human participant: victim)

### **Semantic Triplet Narrative 3**

**Participant** (Actor: Offender)

**Process** (Verb: hit) Circumstances (location of punch: Victim's Chest: Date: ) Space (Town: X)(Location type: Residential Flats Foyer)

**Object** (Human participant: Victim) (Part of Victim: Chest)

#### **Semantic Triplet Narrative 4**

**Participant** (Actor: Offender)

**Process** (Verb: grab)

**Object** (Physical object: Victim's mobile phone)

(Franzosi, 2010)

*Appendix 2: Mapping the script tracks identified in the MO descriptions.*

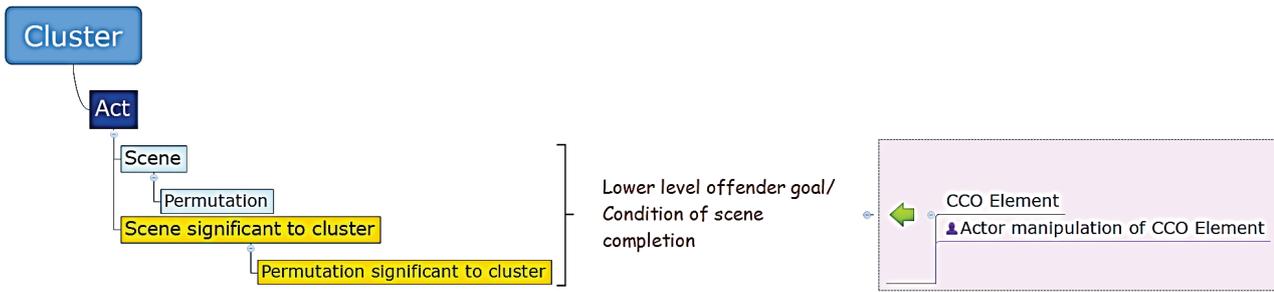
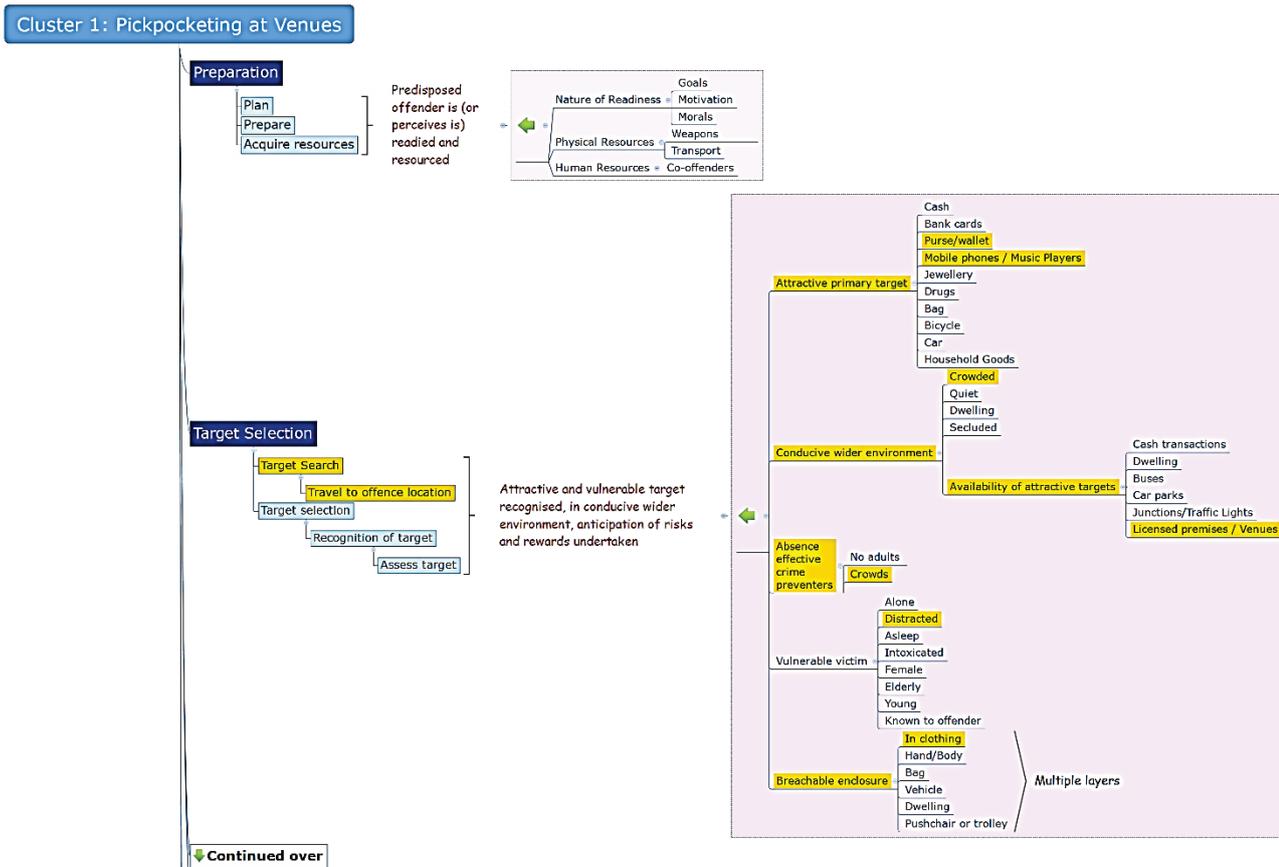
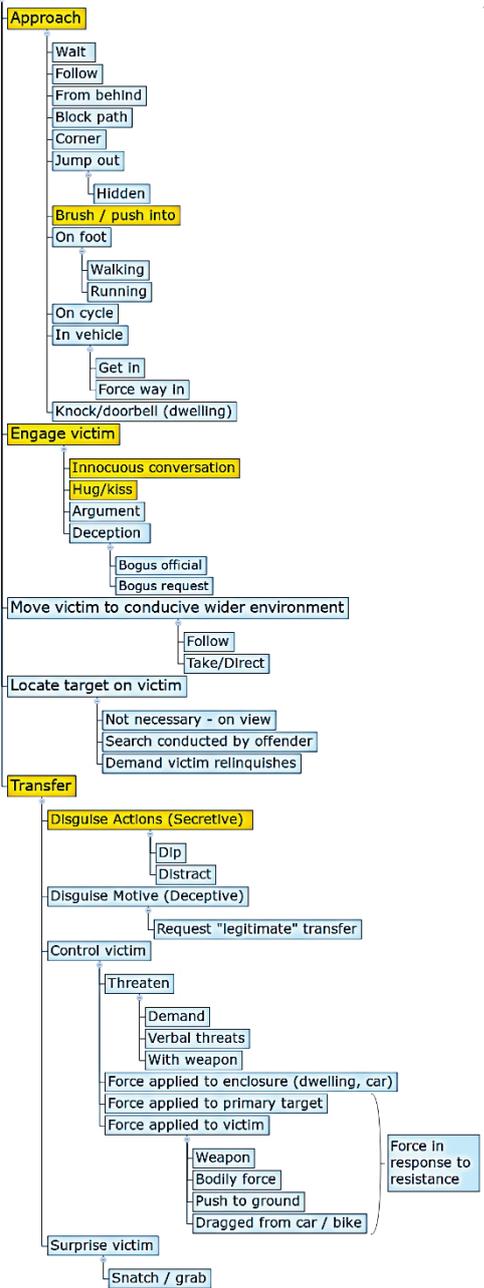


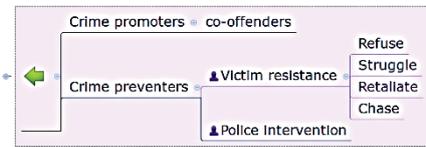
Figure A1. Mapping Cluster 1: Pickpocketing at Venues



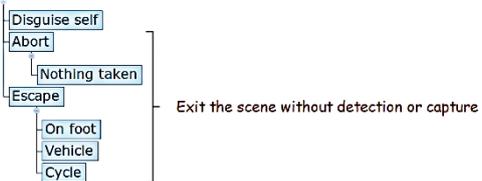
## Transfer of property



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



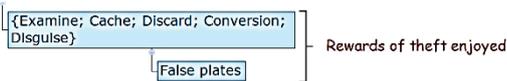
## Escape



Exit the scene without detection or capture



## Disposal



Rewards of theft enjoyed

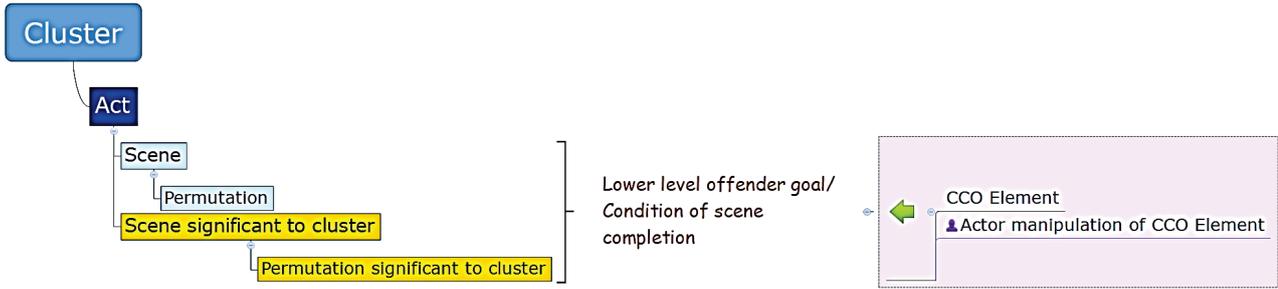
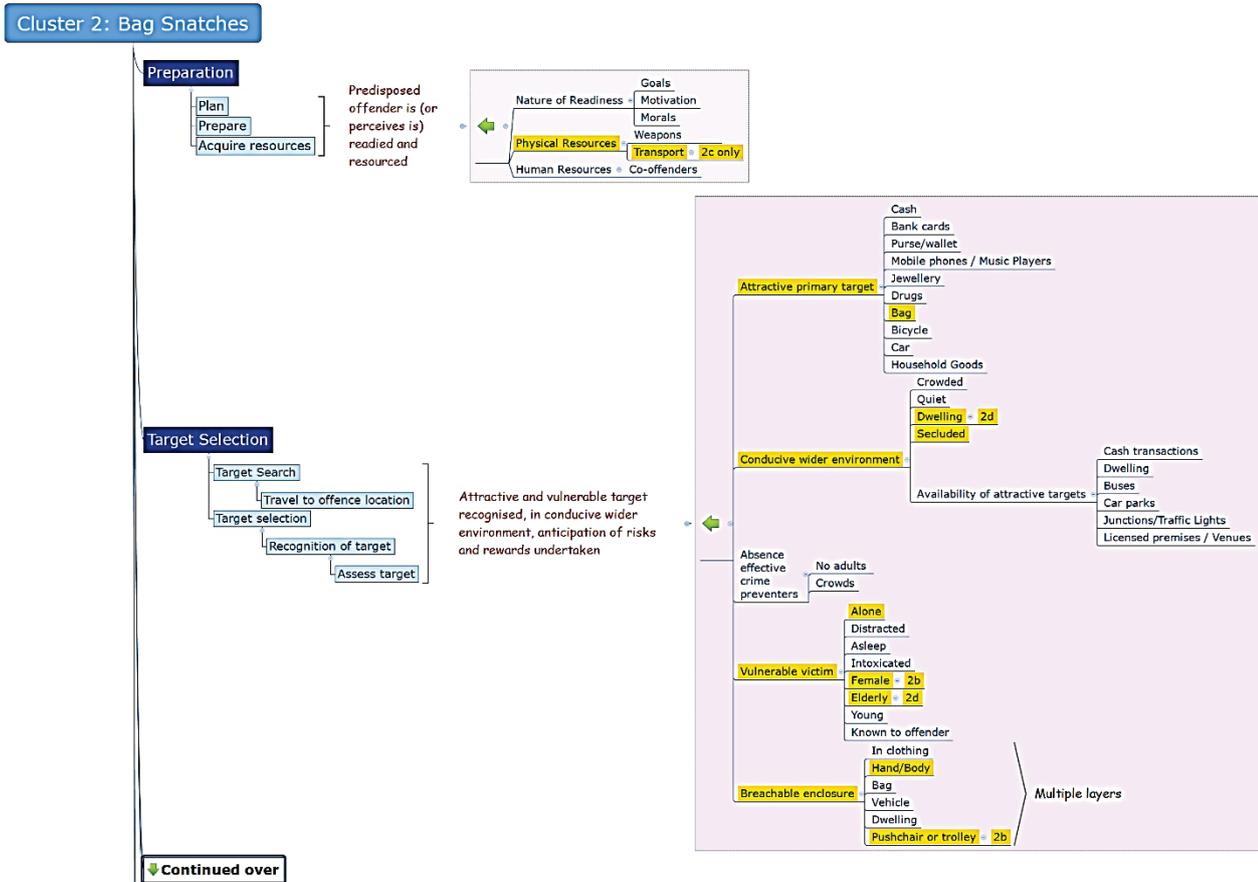
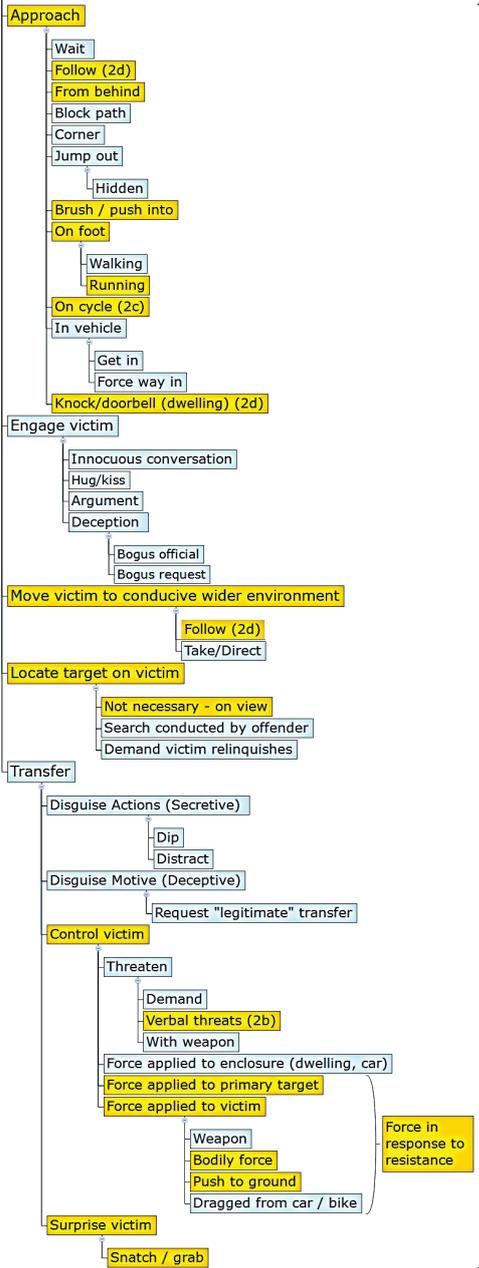


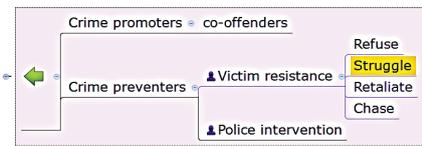
Figure A2. Mapping Cluster 2: Bag Snatches



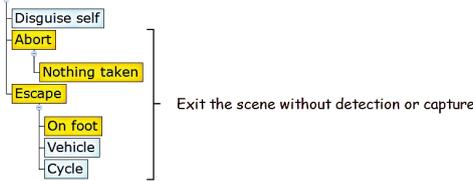
**Transfer of property**



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



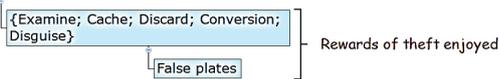
**Escape**



Exit the scene without detection or capture



**Disposal**



Rewards of theft enjoyed

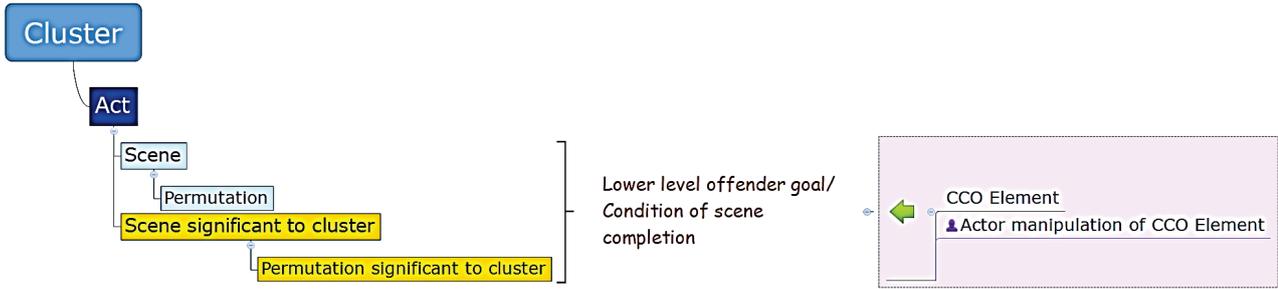
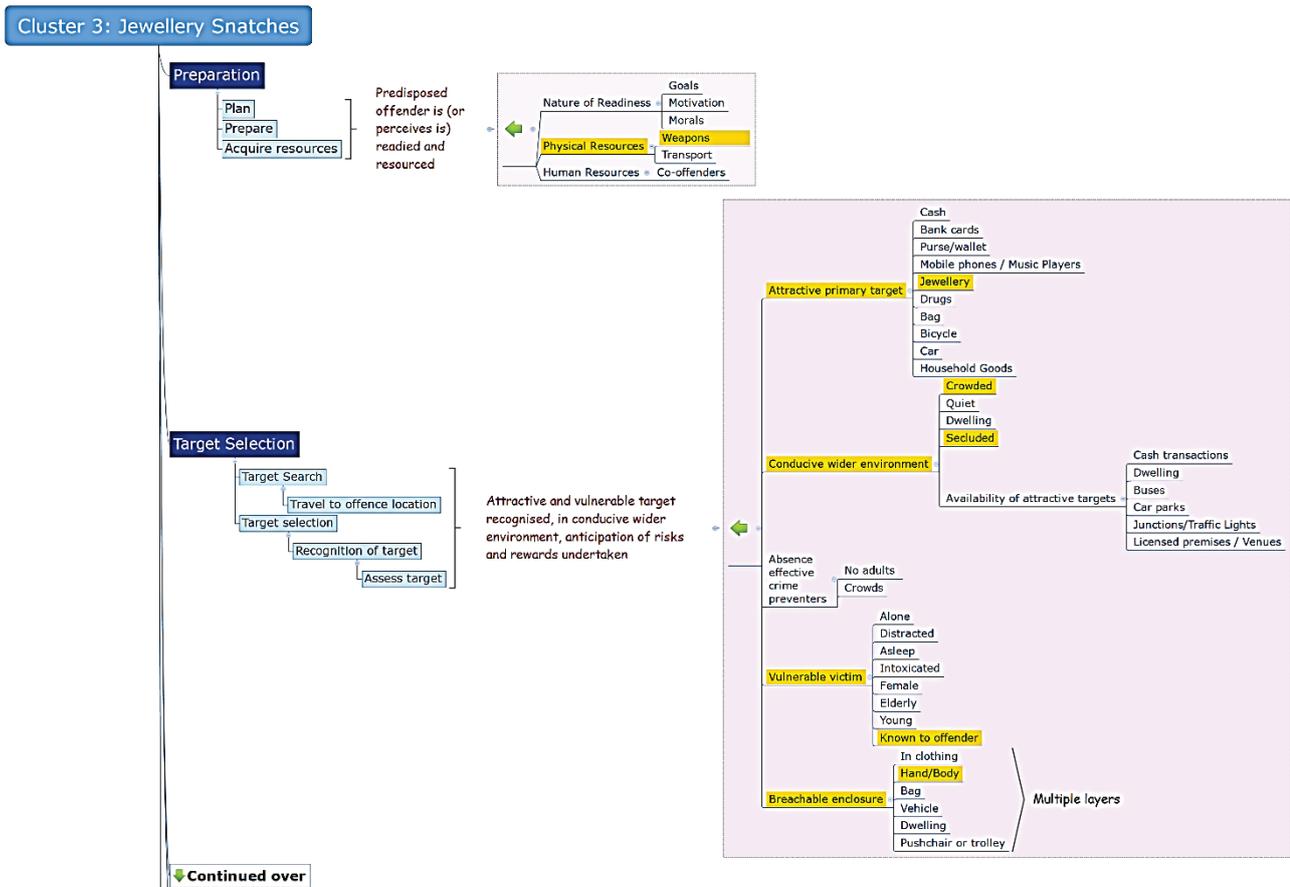
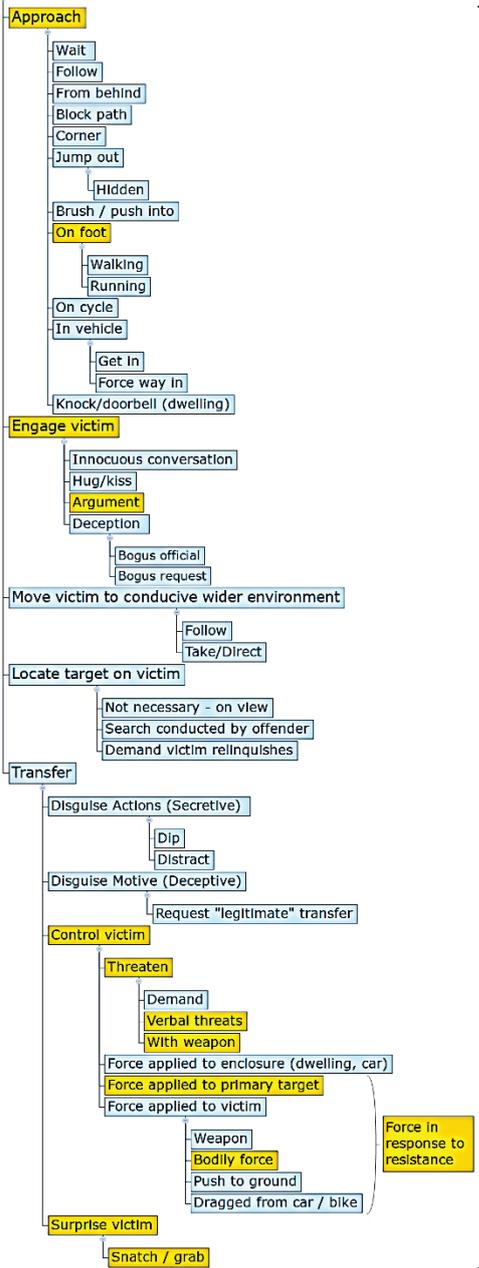


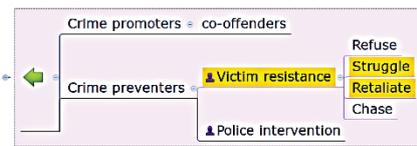
Figure A3. Mapping Cluster 3: Jewellery Snatches



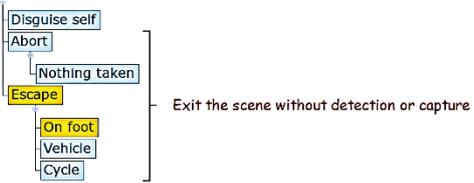
**Transfer of property**



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



**Escape**



**Disposal**



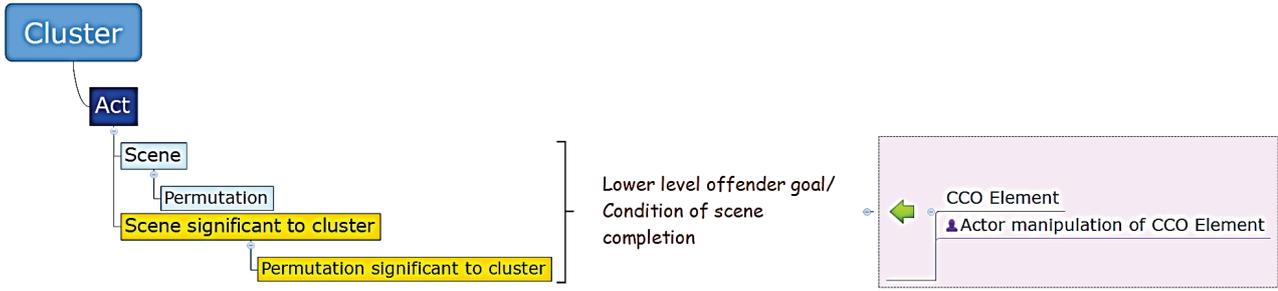
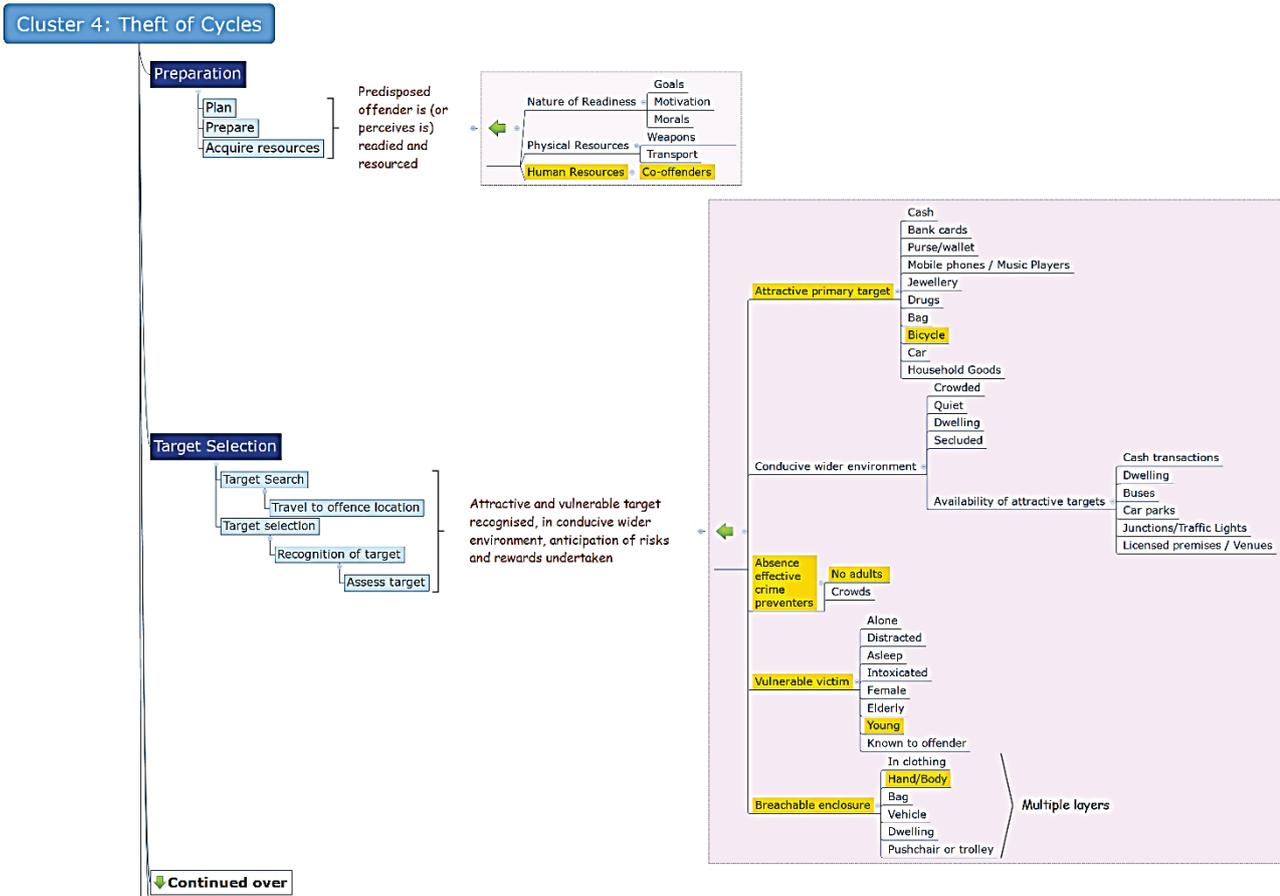
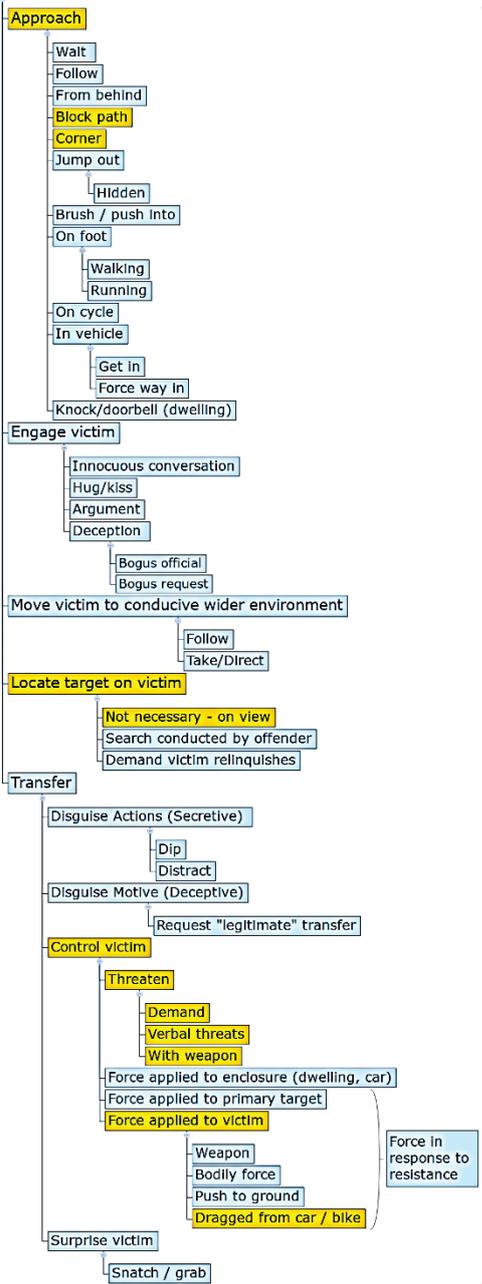


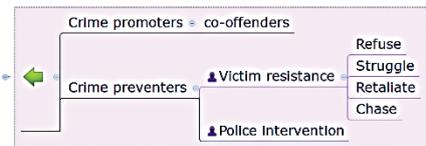
Figure A4. Mapping Cluster 4: Theft of Cycles



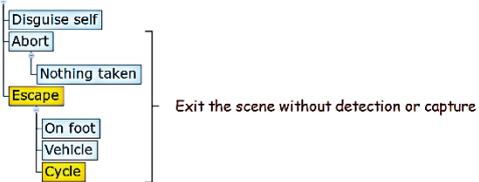
**Transfer of property**



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



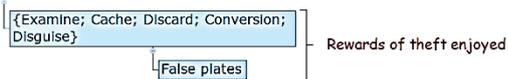
**Escape**



Exit the scene without detection or capture



**Disposal**



KEY

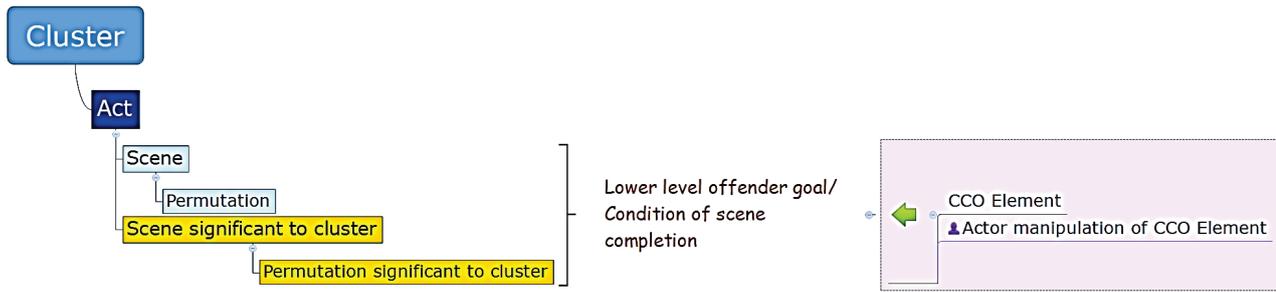
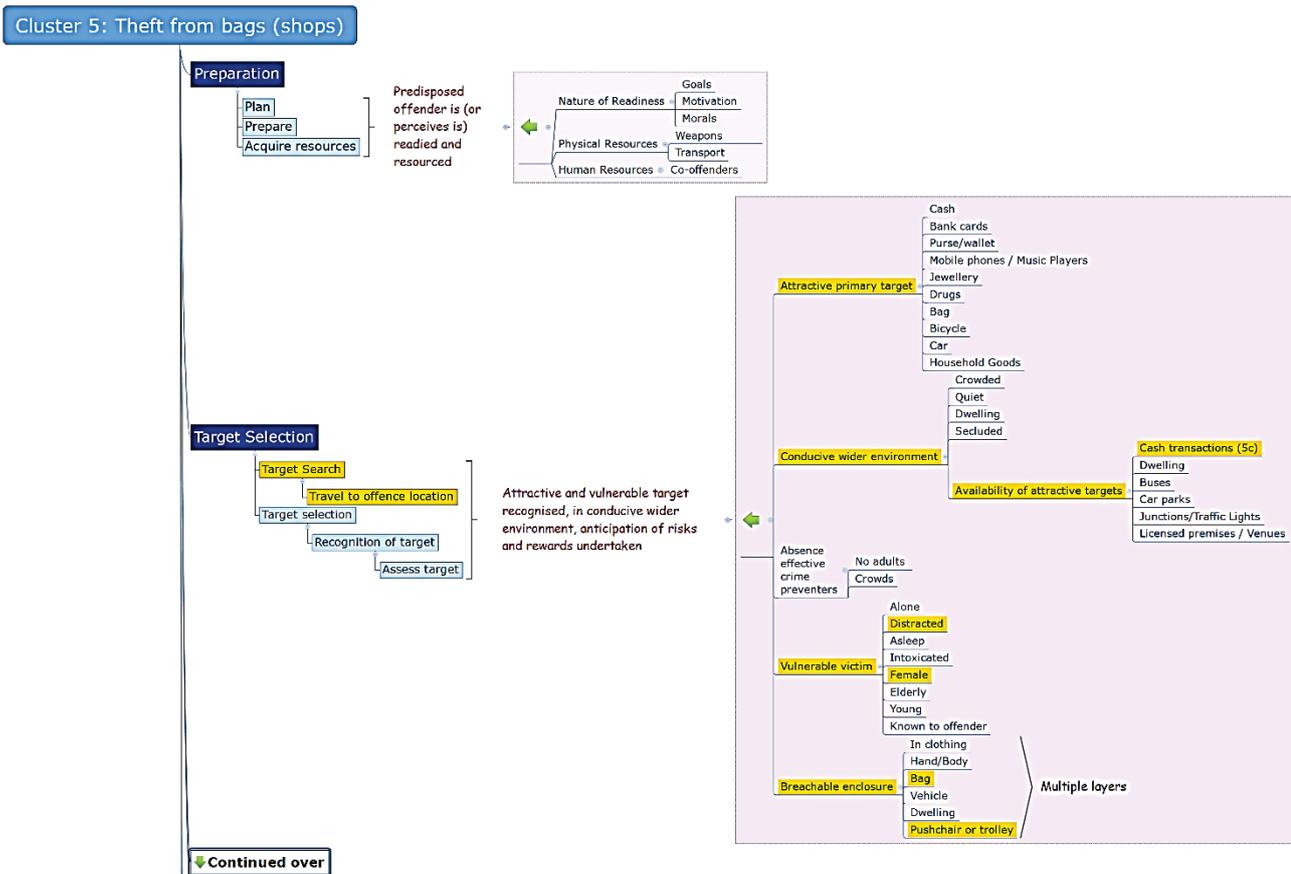
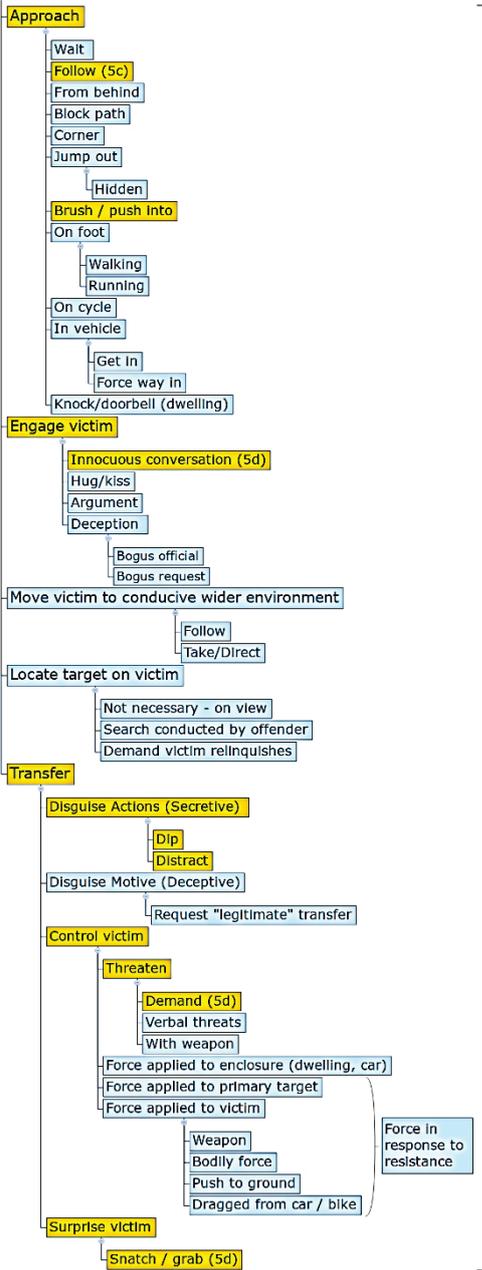


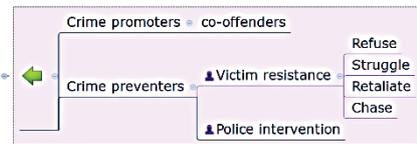
Figure A5. Mapping Cluster 5: Theft from Bags (Shops)



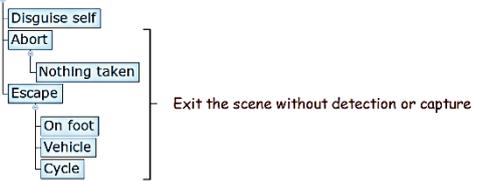
**Transfer of property**



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



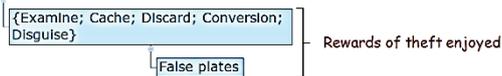
**Escape**



Exit the scene without detection or capture



**Disposal**



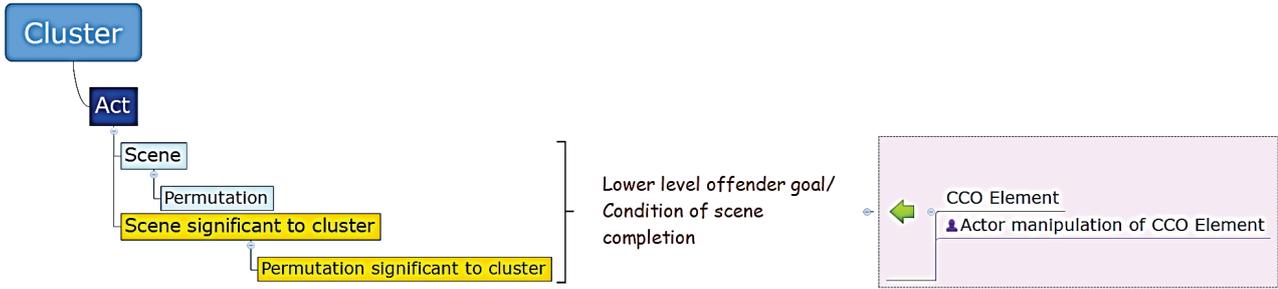
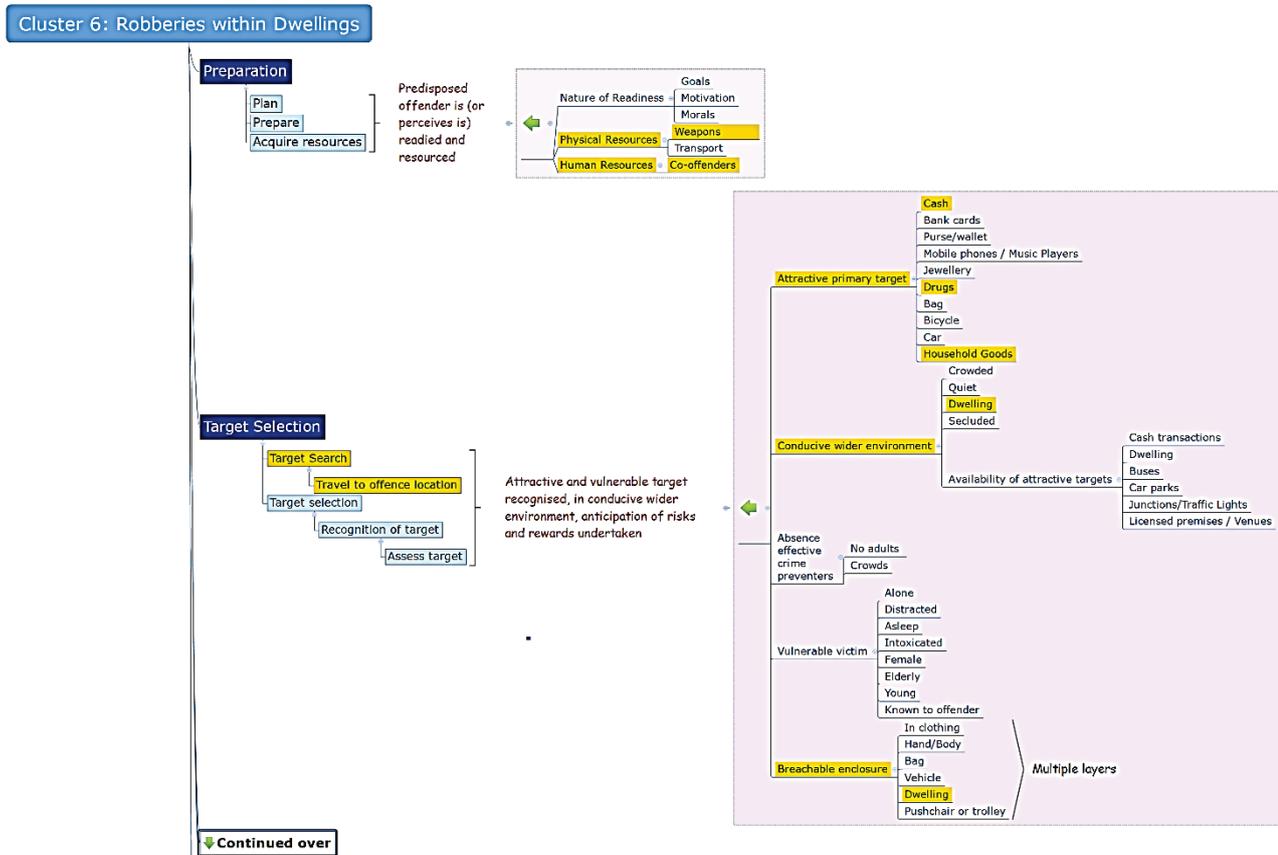
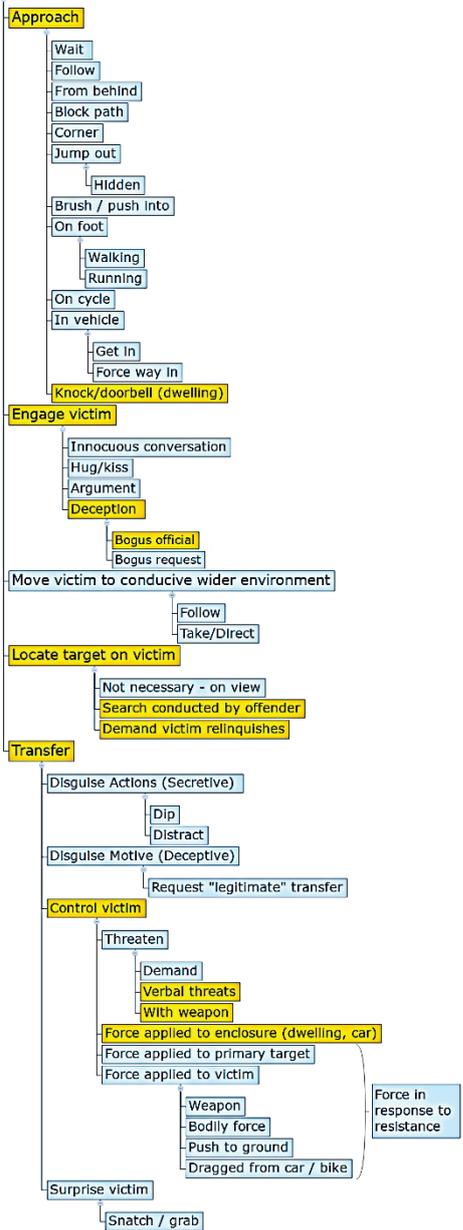


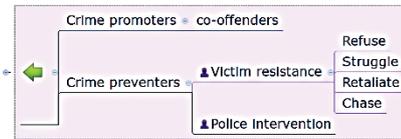
Figure A6. Mapping Cluster 6: Robberies within Dwellings



## Transfer of property



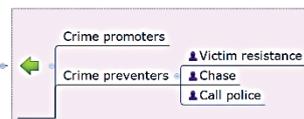
Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



## Escape



Exit the scene without detection or capture



## Disposal



Rewards of theft enjoyed

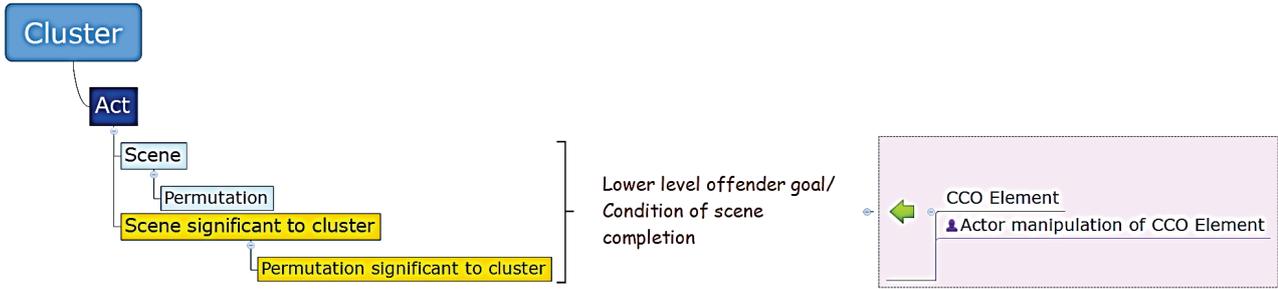
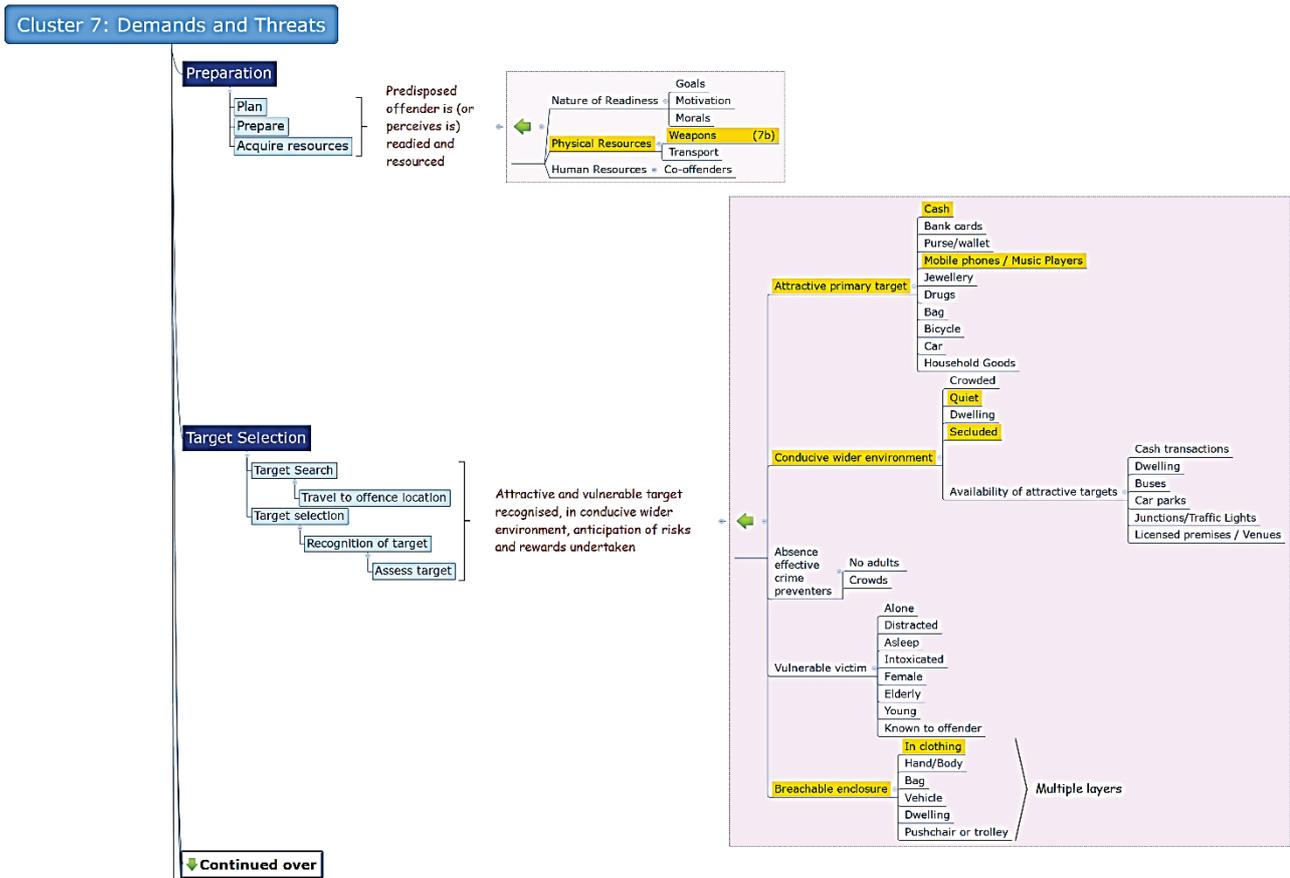
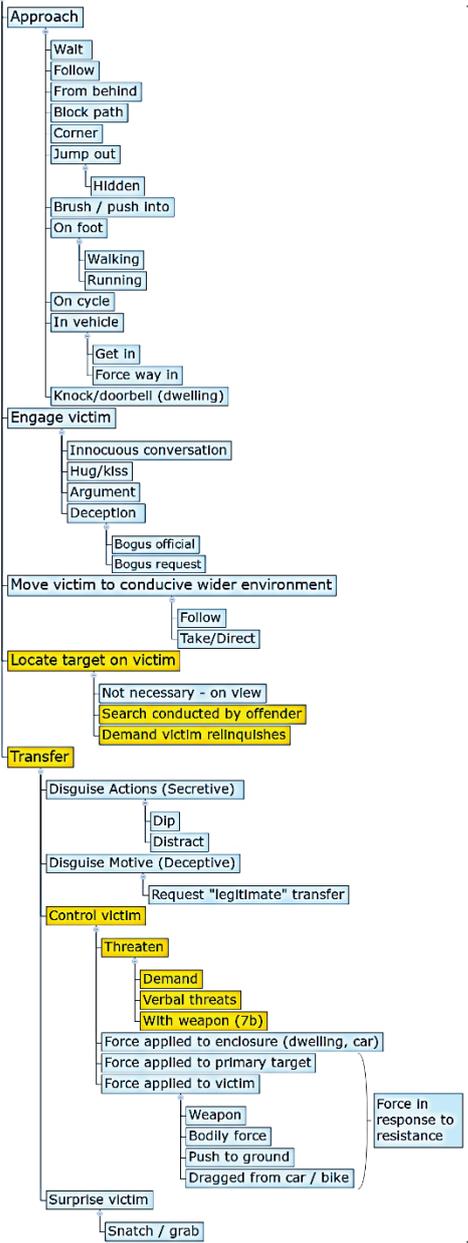


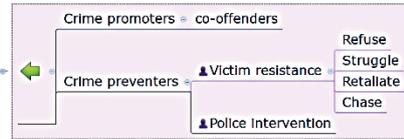
Figure A7. Mapping Cluster 7: Demands and Threats



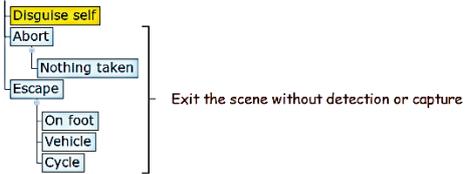
**Transfer of property**



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



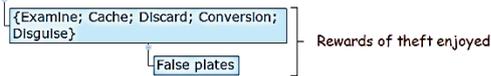
**Escape**



Exit the scene without detection or capture



**Disposal**



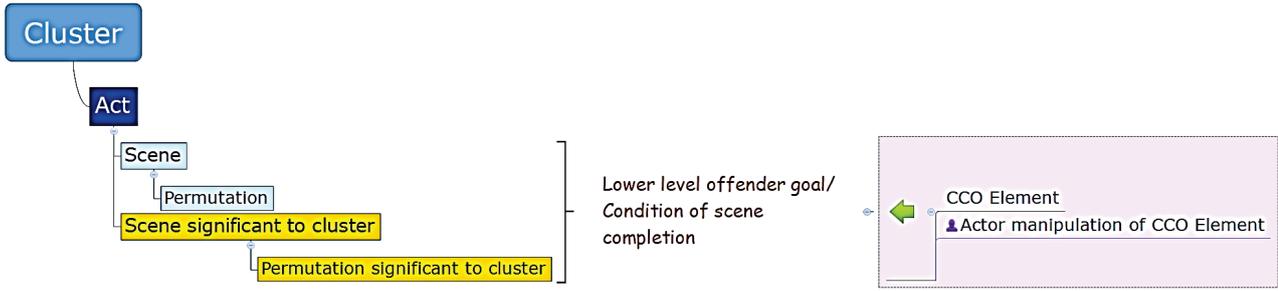
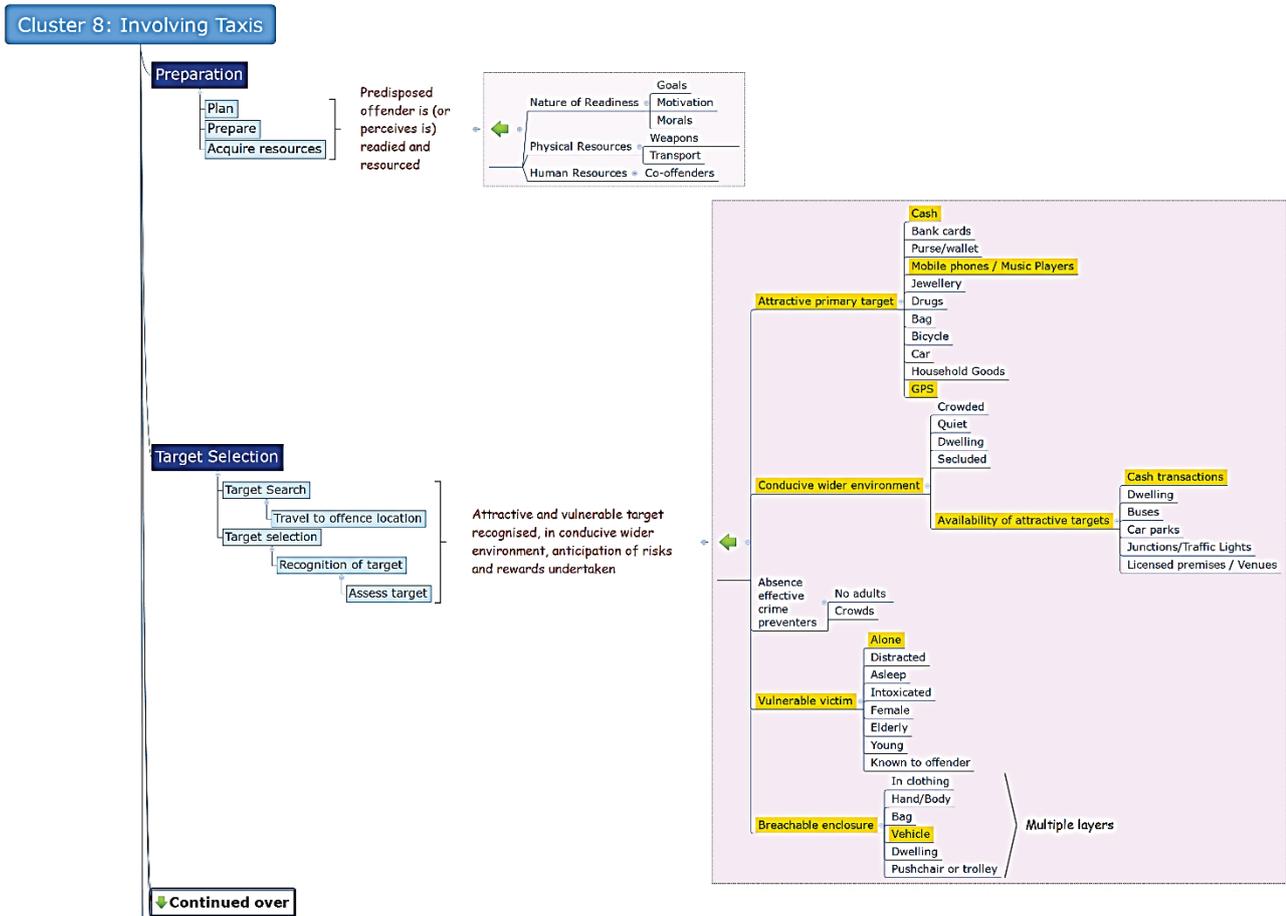
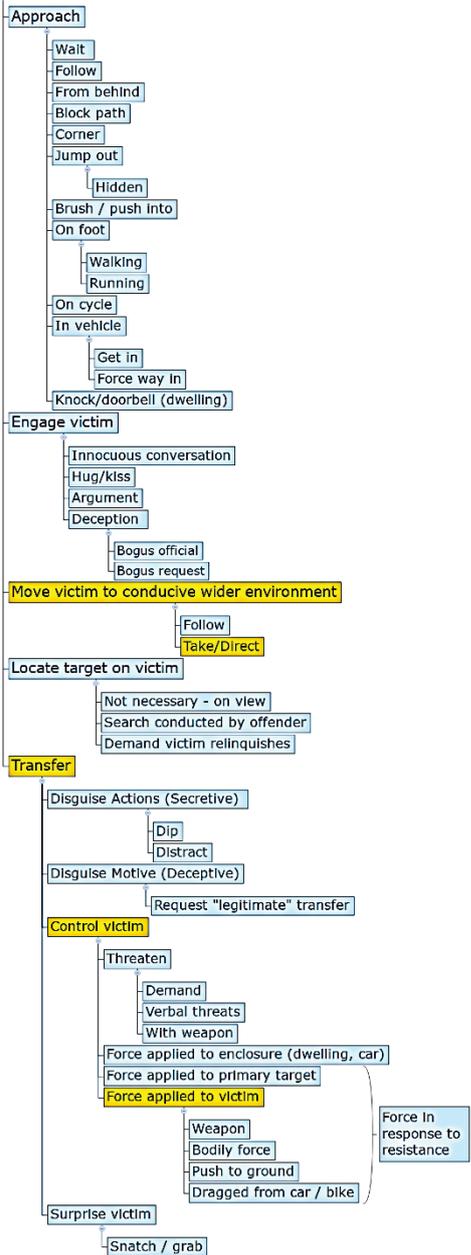


Figure A8. Mapping Cluster 8: Involving Taxis



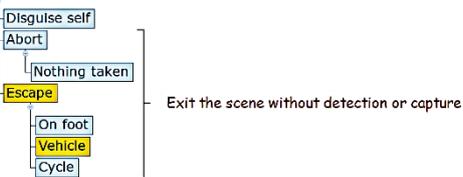
## Transfer of property



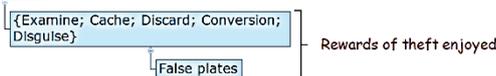
Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



## Escape



## Disposal



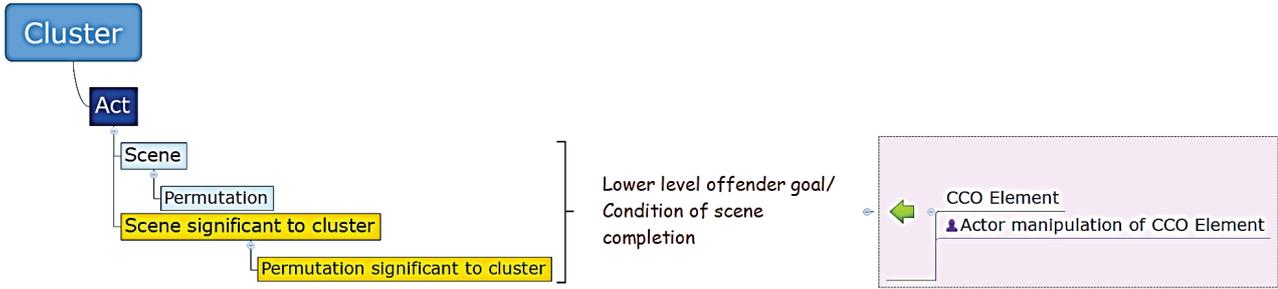
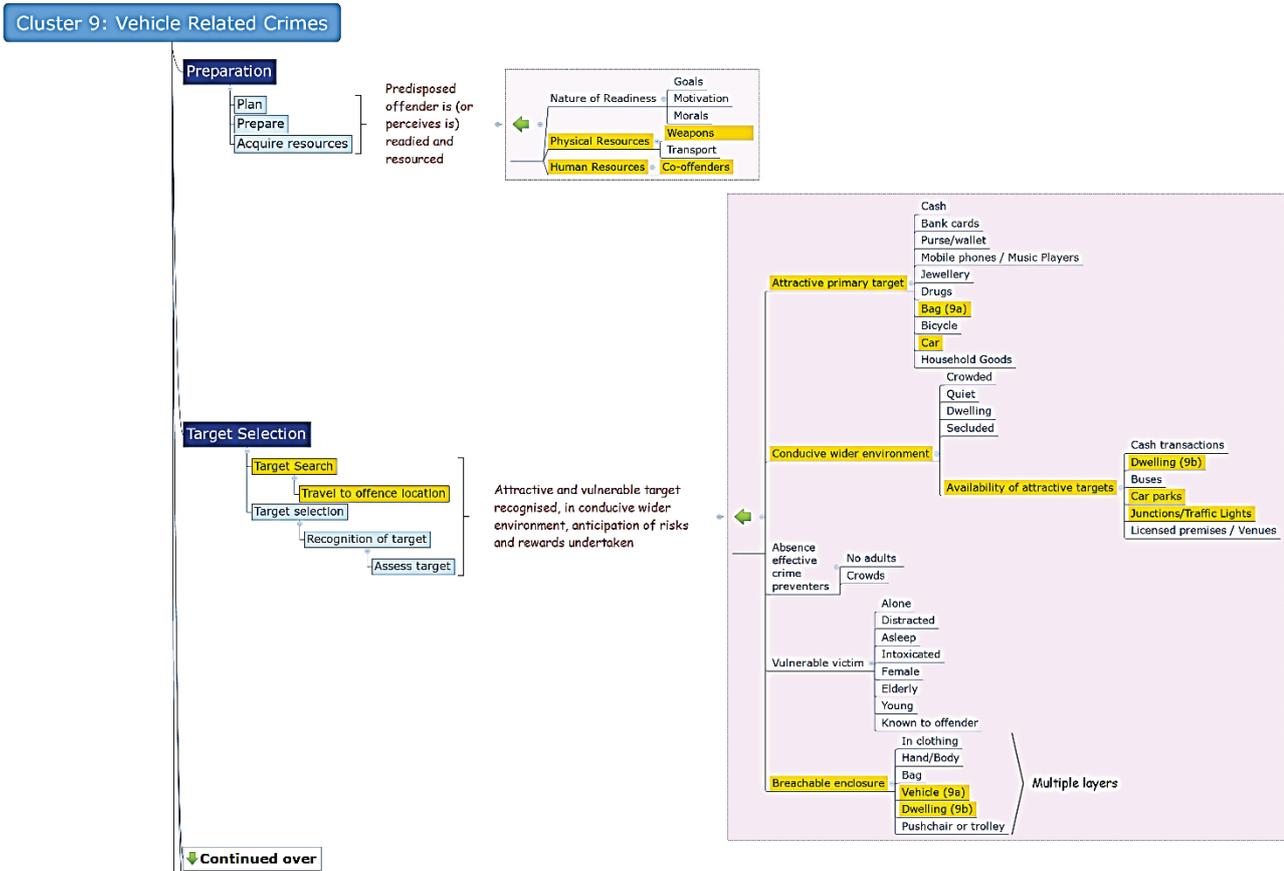
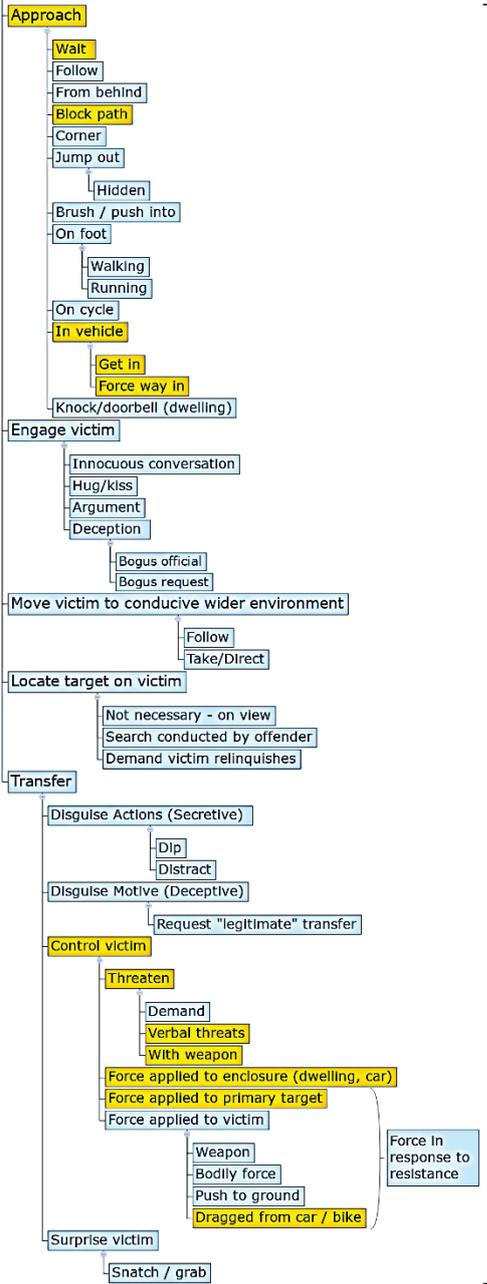


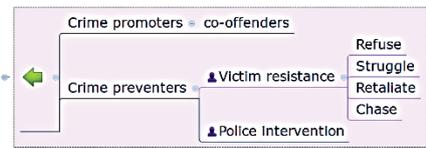
Figure A9. Mapping Cluster 9: Vehicle Related Crimes



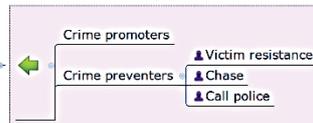
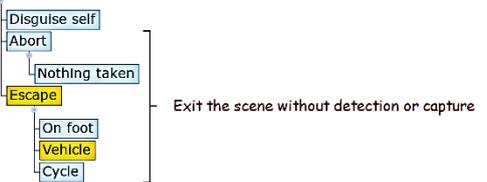
**Transfer of property**



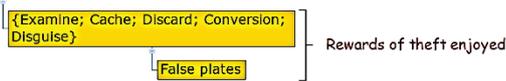
Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



**Escape**



**Disposal**



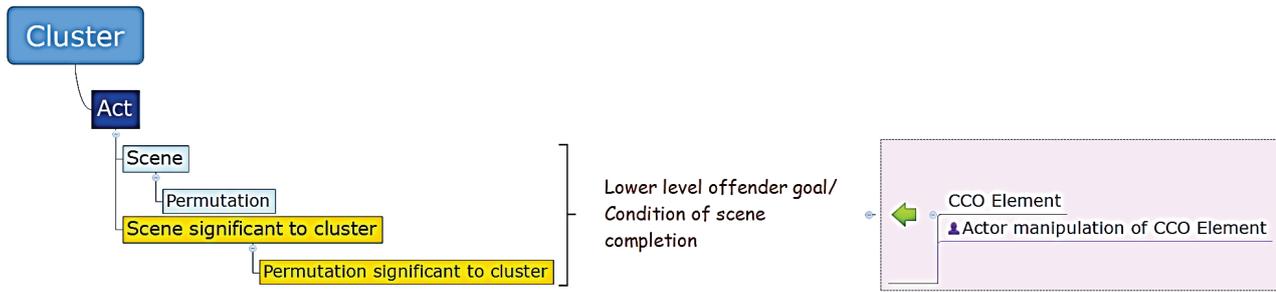
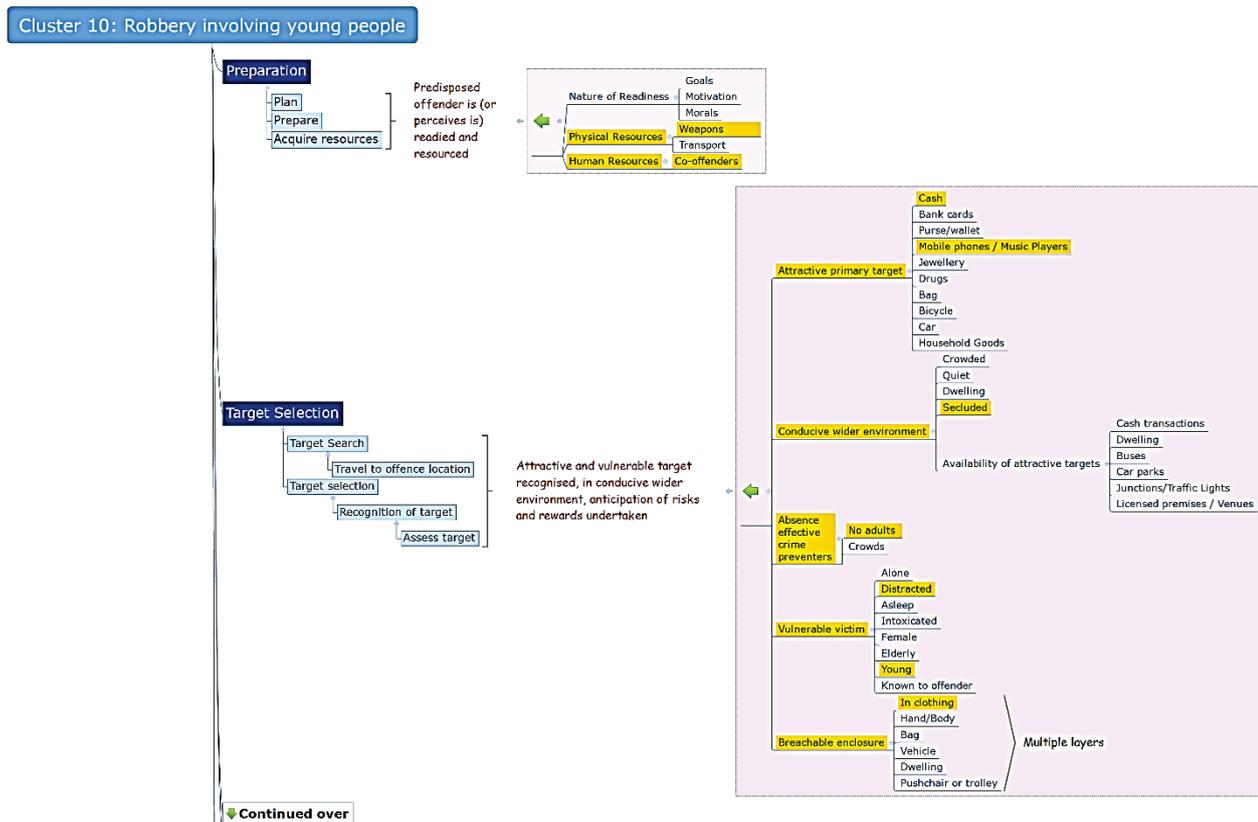
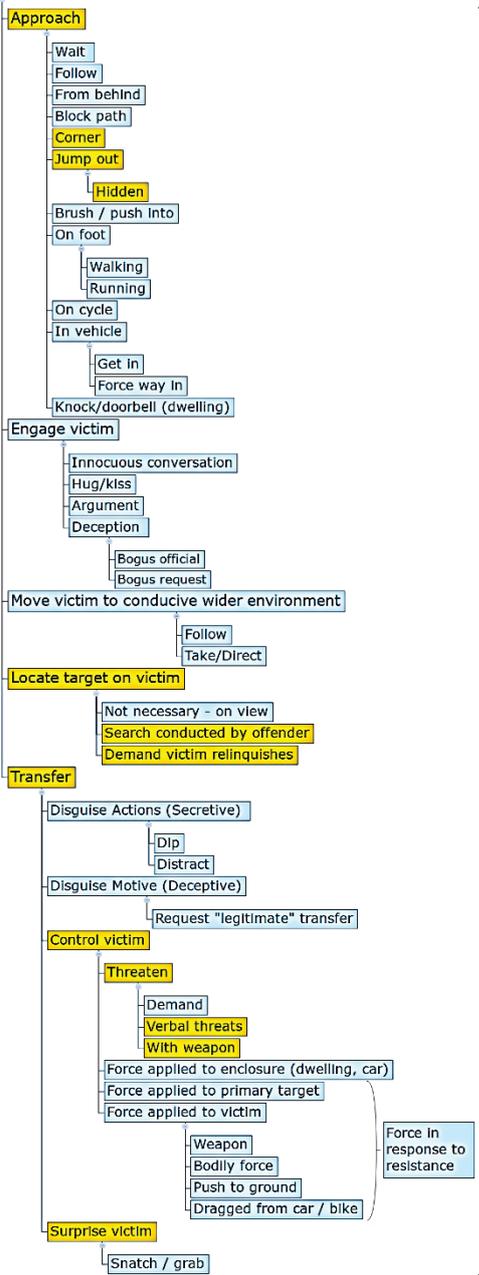


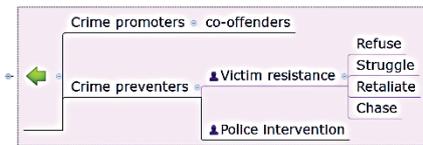
Figure A10. Mapping Cluster 10: Robbery Involving Young People



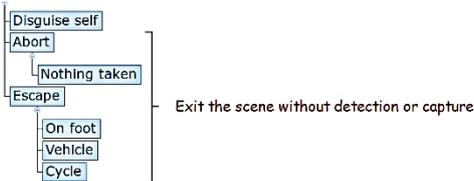
## Transfer of property



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



## Escape



## Disposal



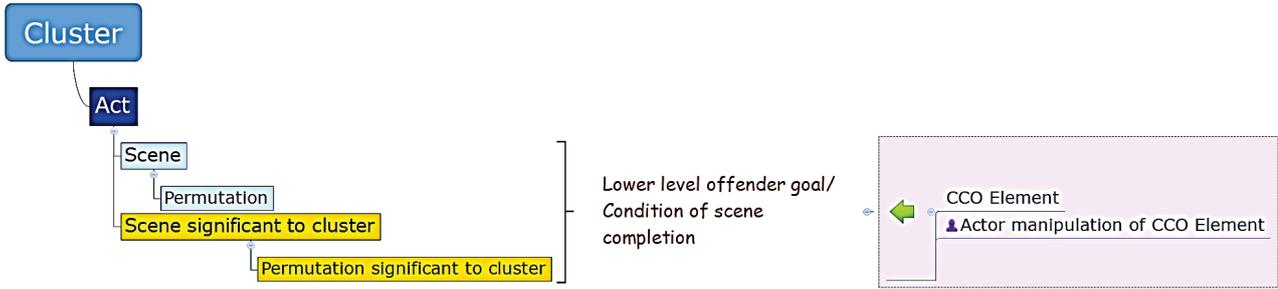
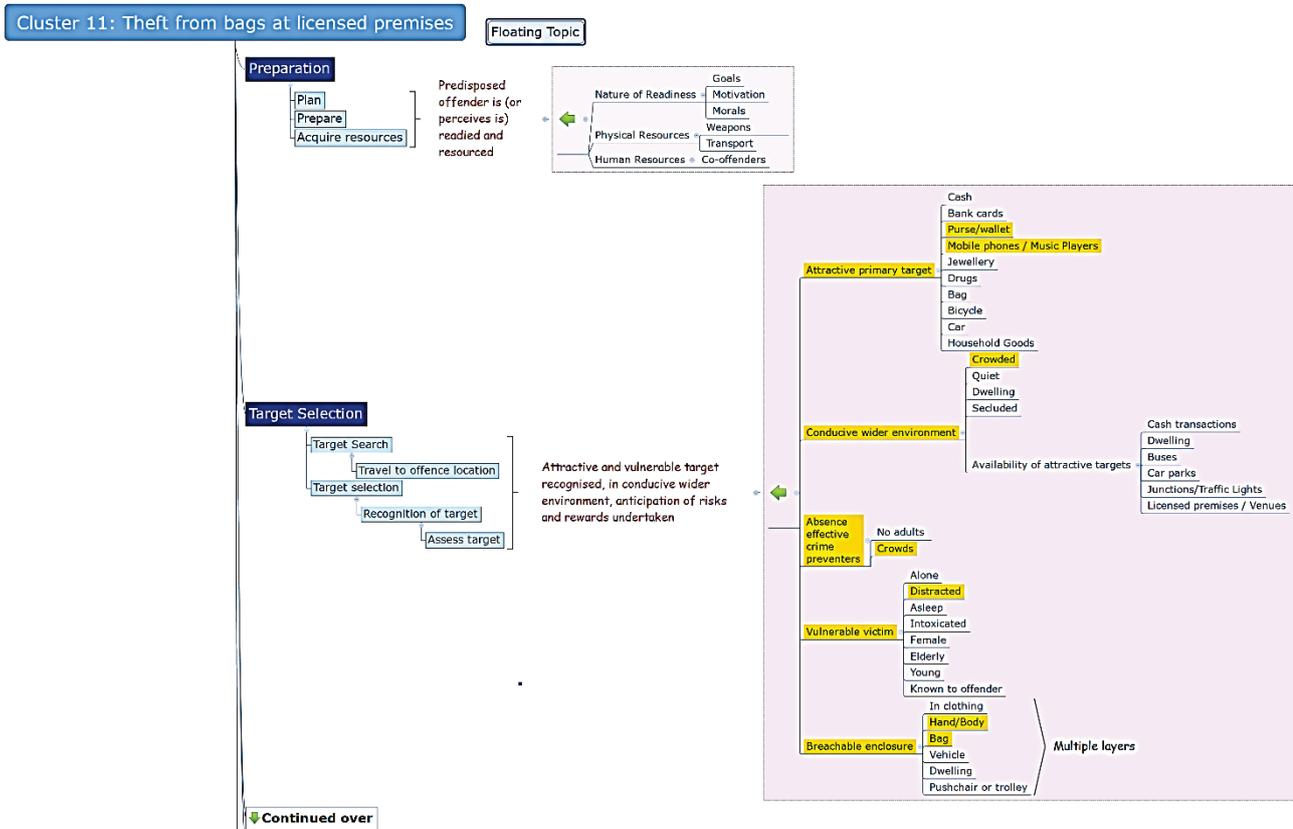
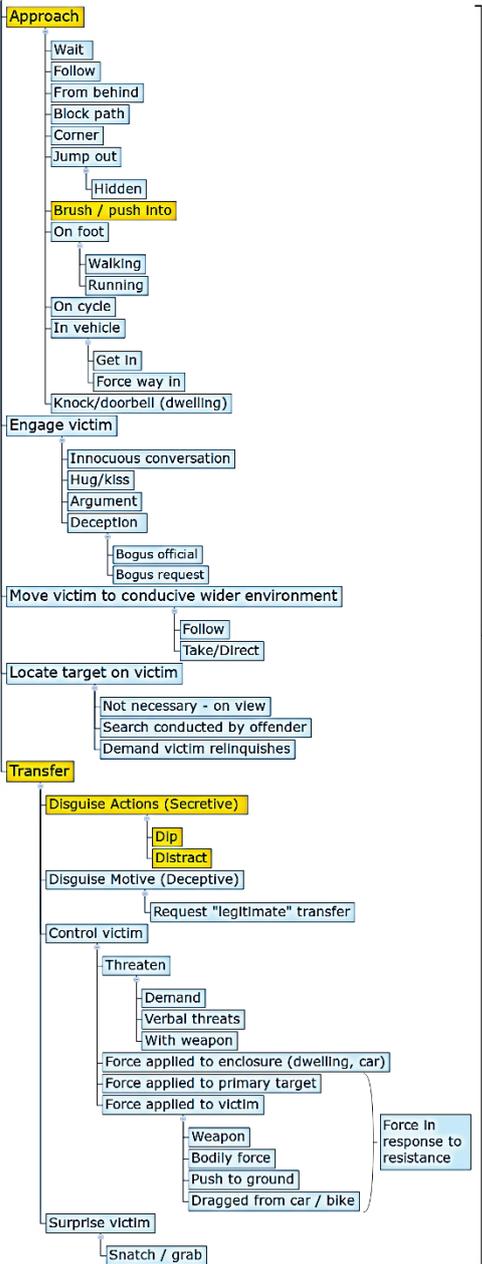


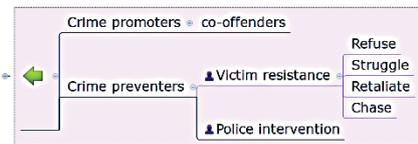
Figure A11. Mapping Cluster 11: Theft from Bags at Licensed Premises



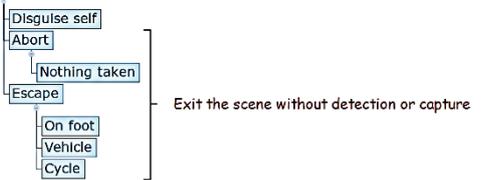
**Transfer of property**



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



**Escape**



Exit the scene without detection or capture



**Disposal**



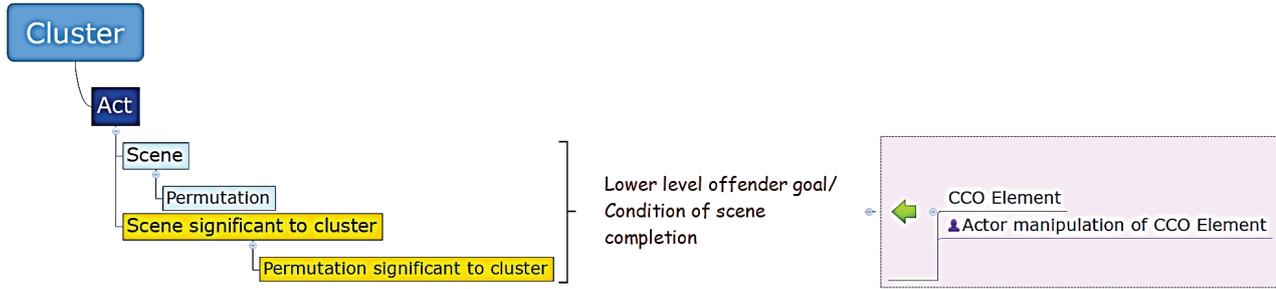
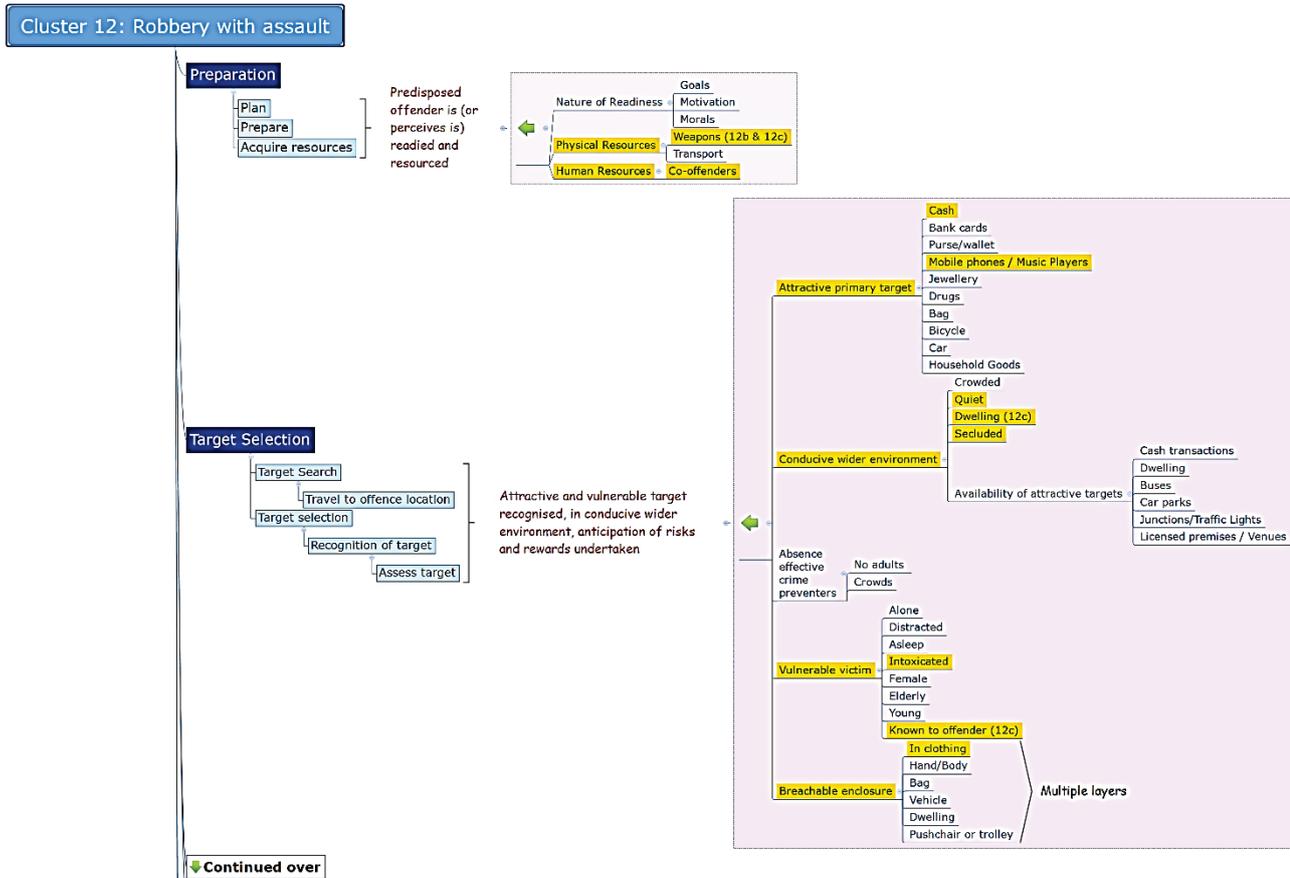
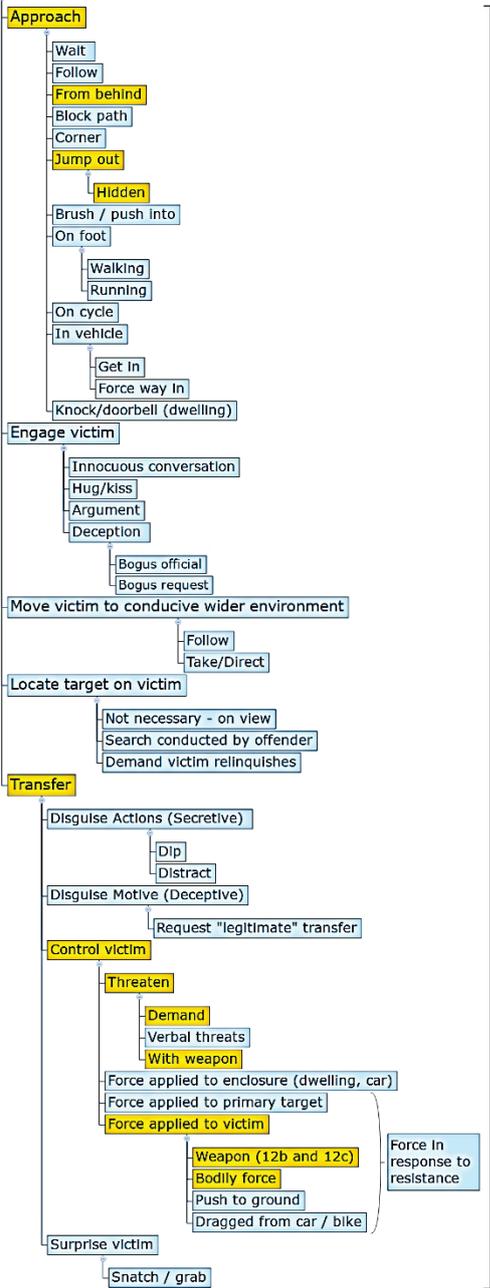


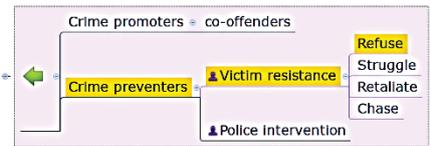
Figure A12. Mapping Cluster 12: Robbery with Assault



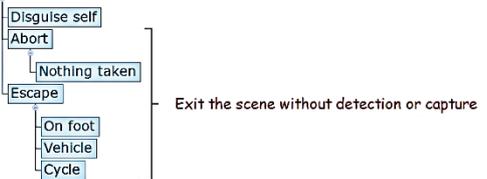
**Transfer of property**



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



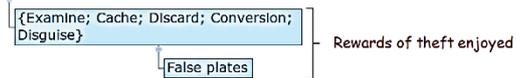
**Escape**



Exit the scene without detection or capture



**Disposal**



Rewards of theft enjoyed

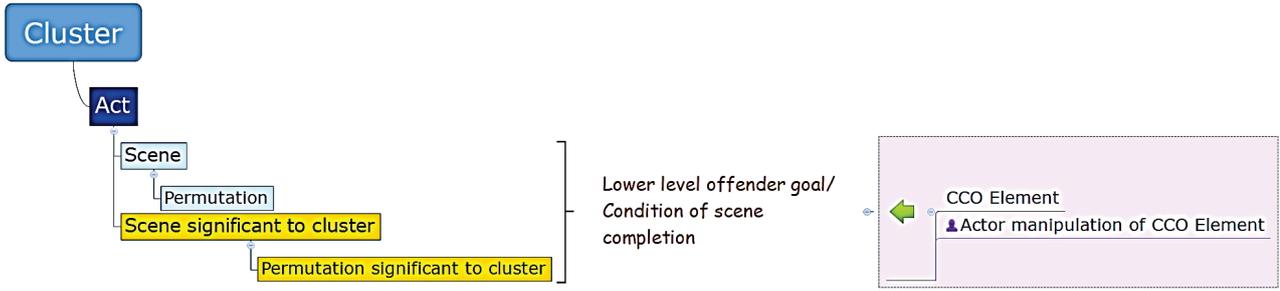
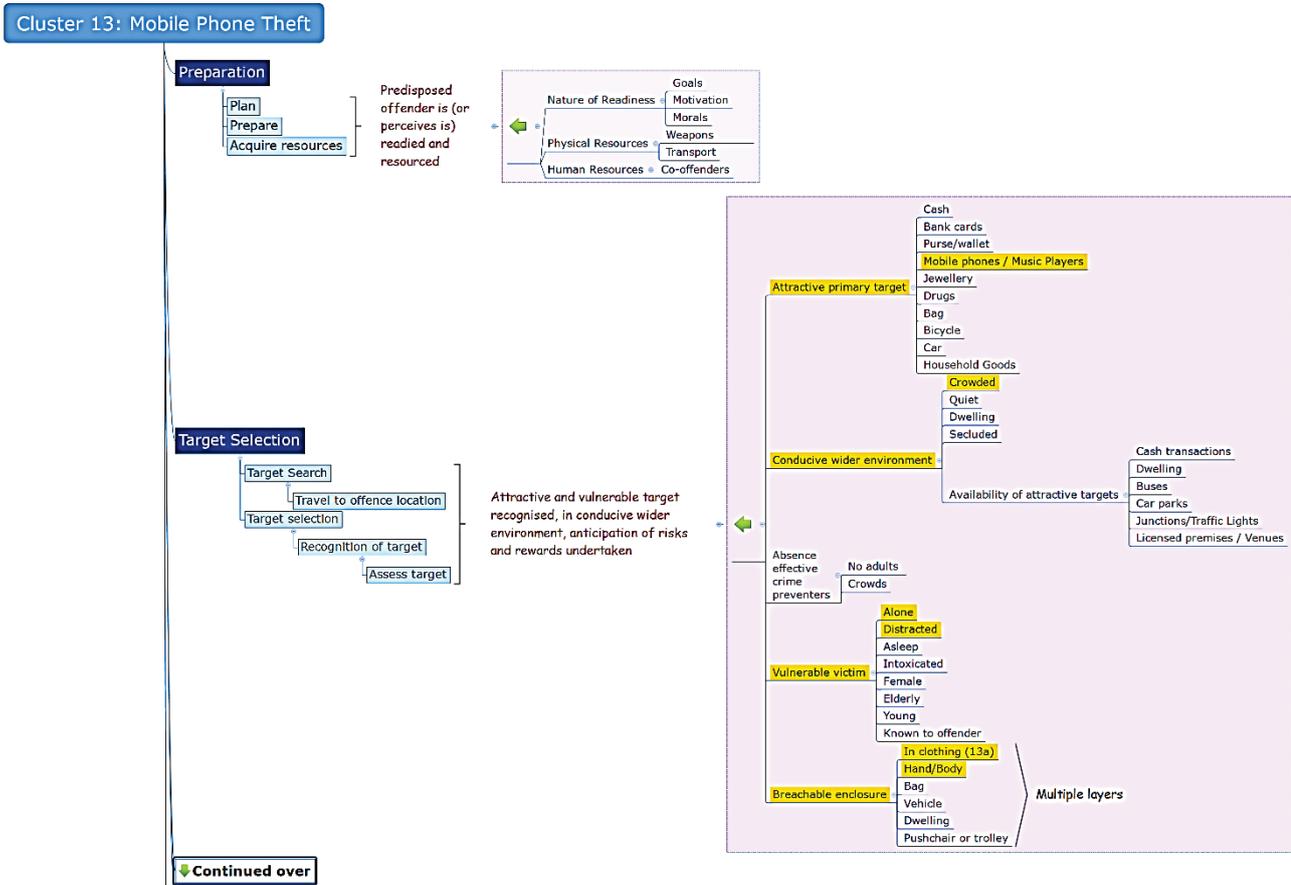
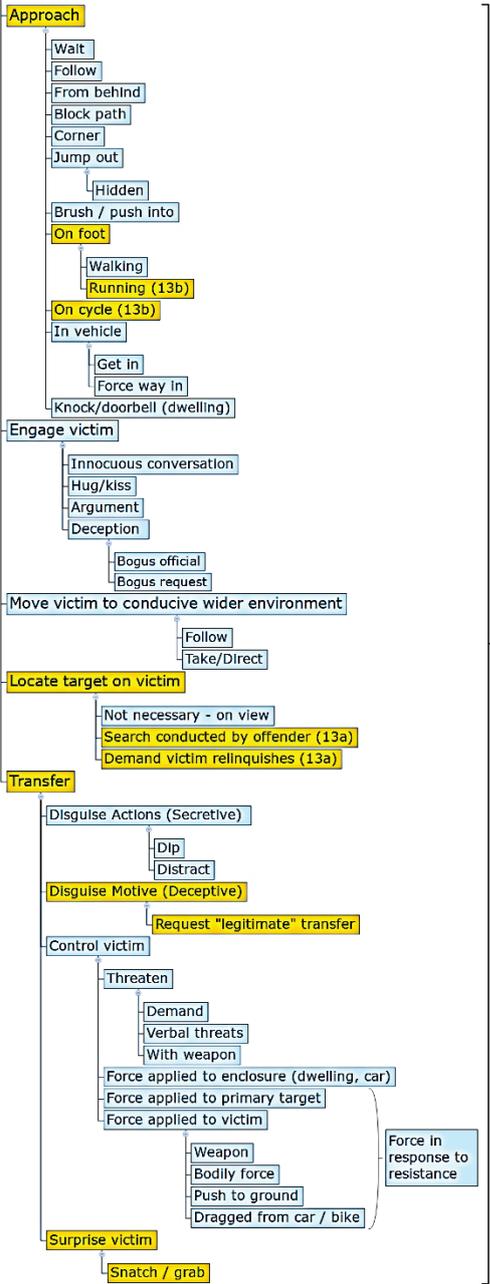


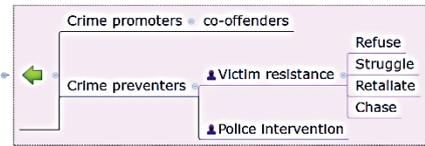
Figure A13. Mapping Cluster 13: Mobile Phone Theft



**Transfer of property**



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



**Escape**



Exit the scene without detection or capture



**Disposal**



Rewards of theft enjoyed

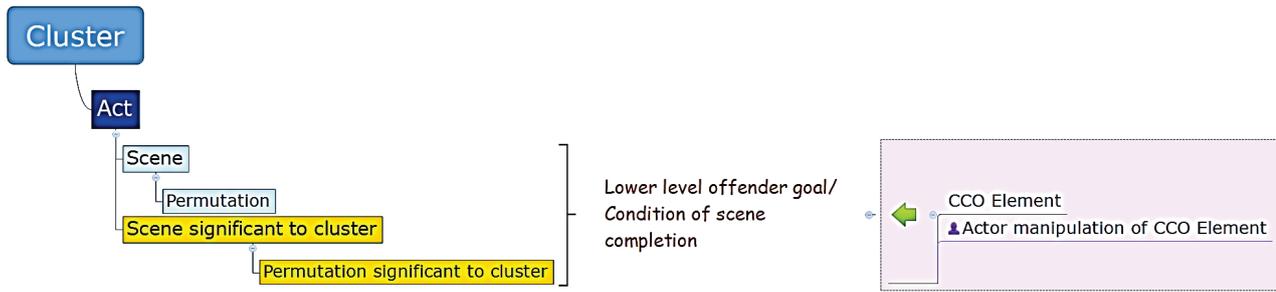
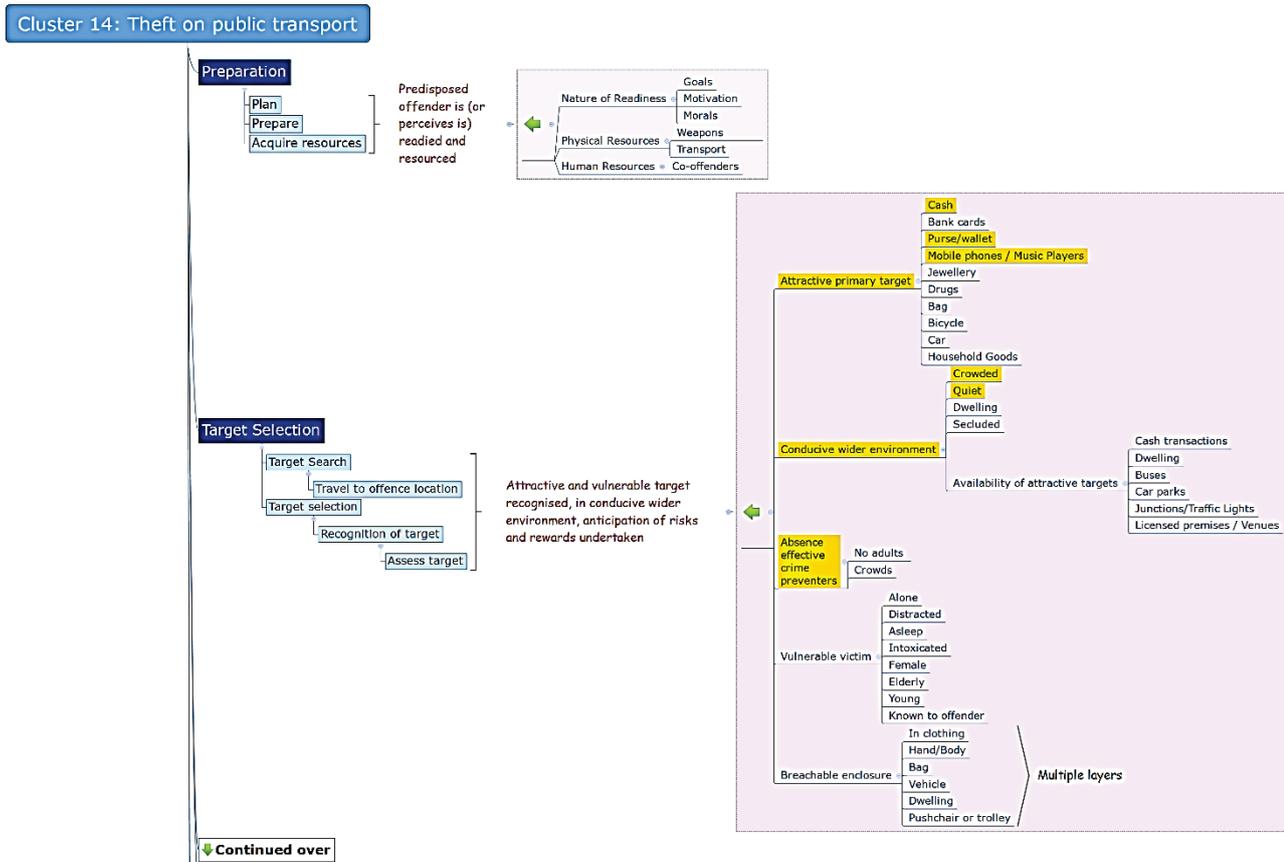
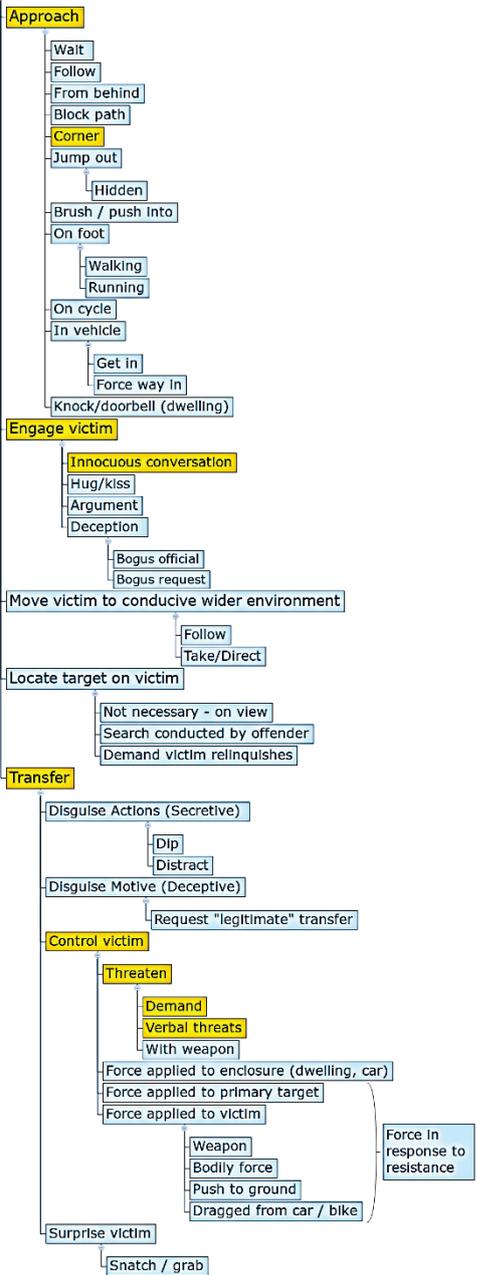


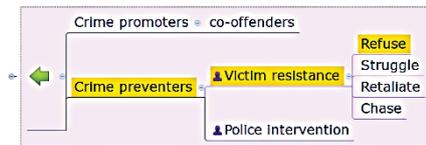
Figure A14. Mapping Cluster 14: Theft on Public Transport



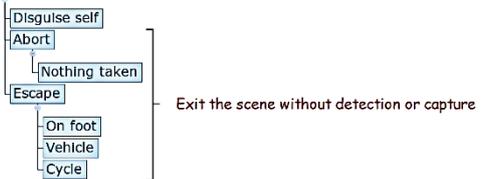
**Transfer of property**



Offender is co-present with victim, target enclosures breached, crime preventers countered, possession of property transferred



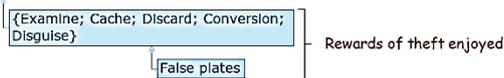
**Escape**



Exit the scene without detection or capture



**Disposal**



Rewards of theft enjoyed