The London School of Economics and Political Science

OFFENDING RISK FACTORS AND AREA: AN INVESTIGATION USING STRUCTURAL EQUATION MODELLING

Jack Cunliffe

A thesis submitted to the Department of Social Policy of the London School of Economics for the degree of Doctor of Philosophy, London, September 2015

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent.

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 93,805 words.

I can confirm that my thesis was copy edited for conventions of language, spelling and grammar by Lisa Findley.

Jack Cunliffe

Abstract

This thesis has two main aims. The first is substantive: to investigate whether and how an individual's perceptions of their area act as risk factors for offending. The second is methodological: to demonstrate that theoretically-informed structural equation modelling can make best use of existing and often under-utilised datasets, particularly cross-national studies such as those typically conducted by large-scale organisations or governments.

Using the United Kingdom Offending, Crime and Justice Survey (OCJS) conducted between 2003 and 2006, and taking a range of questions on individual perceptions, family circumstance, self-reported offending and variables relating to the area in which the respondent lives, the work reviews previous criminological measurement constructs of well-known risk factors (from both an analytical and theoretical perspective) and once these are defined moves on to examine self-report offending using structural equation modelling.

Findings are predominantly consistent with previous work and show that individual criminogenic propensities matter most, but also that a complex interrelationship of area perceptions operate in conflicting directions. Once this is accounted for, living in an area with higher disorder seems to increase self-report offending, with part of the relationship explained by perceptions of lower collective efficacy. However, this relationship seems to operate only at one time point and when looking longitudinally it appears that it is the family situation that takes precedence.

This leads in turn to mixed policy implications. In the short-term, it appears that interventions to address perception of area would be most successful to combat offending behaviour but over the longer term addressing the family situation would be more appropriate. Implications for data collection processes and analytical approaches to existing data are centred on the simple analytical framework that pays equal attention to the set of questions: 1) What can be measured? 2) Can these measures be structured? 3) What are the results?

Acknowledgements

First thanks must go to my primary supervisor Professor Sir John Hills, without whom this work not have even been started. His support through the process has been simply amazing and it's been an inspiration to work closely with him. Special thanks also to Professor Tim Newburn whose incisiveness, criminological knowledge and attention to detail have been absolutely vital to the completion of this work. His constant reminders to be 'less pointy headed' is a lesson not to be forgotten.

Further acknowledgements must go to the Economics and Social Research Council for their funding and to the whole of the Centre for Analysis of Social Exclusion (CASE) within the LSE for the resources that have been available to me. Although I owe a debt of gratitude to all involved with the Centre special thanks must go Jane Dickson for making my life easier, Nic Warner and Jo Joannes for dealing with my computing demands, Rod Hick, Ben Baumberg, Kenia Parsons, Ludovica Gambaro, Marigen Narea-Biscupovich, Kenzo Asahi, Rikki Dean, Kerris Cooper and Elena Mariani.

Beyond CASE I owe thanks to the wider LSE, Ben Wilson, Sally Stares, Else Knudsen, Jouni Kuha, Mike Shiner, Reem Abou-El-Fadl, Seb Falk, Dan Ferguson, Tom Pollak (the last four for getting me papers I could not find elsewhere) and the participants in the Offending, Crime and Justice Survey.

Finally I must thank Adina Oros for being so understanding of the demands this work has placed on my time, and to the soon to be with us Mia Floarea Oros Cunliffe for making a hard deadline all that much harder.

Table of Contents

Declaration 2

Abstract 3

Acknowledgements 4

Table of Contents 5

List of Figures 11

List of Tables 14

Glossary 17

Chapter 1 Introduction 19

1.1 Motivation 19

1.2 Overview of the Thesis 21

Chapter 2 Background 25

2.1 Introduction 25

2.2 Developmental and Life Course Criminology and Risk Factors 25

2.3 The Chicago School, Social Disorganisation and Social Ecology 37

2.3.1 The Old Days 37

2.3.2 Modern Directions 39

2.4 The Development of Collective Efficacy Models 41

2.5 A Note on Social Capital 44

2.6 Situational Action Theory 45

2.7 Summary 49

Chapter 3 Research Specification 51

3.1 Overview 51

3.2 Specific Research Questions 53

3.2.1 Methodological Questions 53

3.2.2 Substantive Questions 55

3.3 Summary 56

Chapter 4 An Introduction to the Offending, Crime and Justice Survey 57

4.1 Possible Data Sources 57

4.2 Introduction to the Offending, Crime and Justice Survey 60

4.3 Nothing is Perfect: OCJS Limitations 63

4.4 OCJS 2004 to 2006 Dataset Basics and Weights 65

4.4.1 Basics 65

4.4.2 Weights 67

4.5 Question Structure and Format 69

4.6 A Discussion on Self-report Offending 72

4.7 Self-report Offending, Basic Descriptives 77

4.8 Data round-up 82

Chapter 5 A Short Overview of Structural Equation Modelling 84

5.1 Advantages and Limits of Secondary Data Analysis 84

5.2 Introduction to Structural Equation Modelling 86

5.3 Graphical Displays of SEM 90

5.4 Ontology and Causality within SEM 94

5.4.1 Ontology 94

5.4.2 Causality 97

5.5 Fitting a SEM Measurement Model 100

5.6 Moving to a Structural Model 105

5.7 Calculation of Total and Indirect Effects 109

5.8 Interpretation of Model Results with a Dichotomous (or Categorical) Dependent 111

5.9 Concluding Remarks 114

Chapter 6 Measurement of Self-report Risk Factors 116

6.1 Process and Aims 116

6.2 Defining Predictors at the Level of the Individual 122

6.2.1 The Available Data 122

6.2.2 Morals 124

6.2.3 Empathy, Self-control and Contentment 126

6.2.4 Individual Level Factor Theoretical Consideration 130

6.2.5 Bringing it together to Individual Attitude 134

6.2.6 Missing Individual Domains 136

6.3 Defining Predictors at the Family Level 139

6.3.1 The Available Data 139

6.3.2 Understanding the Family Question Routing and Low Response Questions 141

6.3.3 Family Level Latent Variables 144

6.3.4 Family Level Theoretical Considerations and Missing Domains 149

6.4 Defining Predictors at the Area Level 150

6.4.1 What is meant by 'Area'? 150

6.4.2 The Available Data 152

6.4.3 Defining a Collective Efficacy Dimension 155

6.4.4 Adding Neighbourhood Disorder and Local Fear of Crime 163

6.5 Latent Variable Measurement Conclusions 167

Chapter 7 Building a Psychological Risk Factor Map of Interrelations 169

7.1 Process and Aims 169

7.2 The Centrality and Stability of Individual Attitude 172

7.3 Covariance, Moderation or Mediation? 174

7.4 Starting Steps: Testing Each Domain 178

7.4.1 Individual Attitude 178

7.4.2 Family Situation 180

7.4.3 Area Character 182

7.5 Combining Two Domains 190

7.5.1 Family Level with Controls for Individual Attitude 191

7.5.2 Area Variables with Controls for Individual Attitude 193

7.6 Modelling All the Domains Together 197

7.7 Structural Conclusions 203

Chapter 8 Attempting to Model Different Crime Types 206

8.1 Aims: A Deepening of Understanding? 206

8.2 Offending Definitions and Statistical Power 208

8.3 Modelling Different Versions of All Offending 213

8.4 Modelling Alternative Offending Types 216

8.4.1 Violent Offending 216

8.4.2 Property Offending 217

8.4.3 Serious Offending 218

8.4.4 Frequent Offending 218

8.5 Summary So Far 219

8.6 Testing Conclusions with Further Analysis 221

8.6.1 Focusing on the Dependents 221

8.6.2 Focusing on the Mediation and Sample Size 222

8.7 Implications and Discussion....224

Chapter 9 Expanding the Model to Account for Structural Characteristics....227

9.1 Overview of the Process....227

9.2 Control Variables for Inclusion....229

9.2.1 Individual Level Controls....229

9.2.2 Family Level Controls....231

9.2.3 Higher Level Controls....233

9.2.4 Review of controls for inclusion and a note on missing values....237

9.3 Control Variable Results....239

9.4 Control Discussion....244

Chapter 10 Towards a Longitudinal Model....247

10.1 Scope of the Chapter....247

10.2 An Overview of Longitudinal Modelling with SEM....248

10.2.1 Measurement Invariance....248

10.2.2 Sample Size and Power....251

10.2.3 Longitudinal Structure....252

10.2.4 Longitudinal Mediation Analysis....258

10.3 Approach to Be Taken: A Lagged Model....261

10.4 Lagged Model Results....262

10.4.2 Individual Attitude-only Results....263

10.4.3 Model Results with Controls Added....265

10.5 Discussion....270

Chapter 11 Discussion and Conclusion....273

11.1 Introduction....273

11.2 Recap of Background and Research Specification....273

11.3 Data and Method 277

11.4 Summary of Results 279

11.4.1 Measurement of Risk Factors 279

11.4.2 Structuring Relationships between Psychological Risk Factors 281

11.4.3 Modelling Different Types of Offending 286

11.4.4 Adding Non-psychological Constructs 287

11.4.5 Incorporating Longitudinal Data 290

11.4.6 Overview of Results 292

11.5 Limitations of the Data Available 293

11.6 Implications and Wider Limitations 295

11.6.1 Policy Implications 296

11.6.2 Methodological Implications for Existing Data 298

11.6.3 Wider Limitations and Implications for Data Collection and Survey Design 300

11.7 Future Work 302

Appendix A 305

Bibliography 309

List of Figures

Figure 4.1 OCJS achieved sample, by year of entry to study 63

Figure 4.2 Number of respondents in 2004 to 2006 longitudinal dataset, by gender and age (in 2004) 66

Figure 5.1 Prevalence of structural equation modelling cited as methodology in the social sciences on Web of Science, 1990 to 2012 89

Figure 5.2 A fictional general structural equation model 90

Figure 5.3 A simplified presentation of a full structural model 93

Figure 5.4 Measurement models present in Figure 5.1 94

Figure 5.5 Two empirically equivalent measurement models 97

Figure 5.6 Standardised structural path model for plant recovery after fire, adapted from Grace and Bollen (2005) 107

Figure 5.7 SEM with a single mediation effect 111

Figure 5.8 SEM with a single mediation effect and categorical dependent 112

Figure 6.1 Moral latent variable loading diagram 126

Figure 6.2 Empathy and self-control possible model specifications 127

Figure 6.3 Final individual level latent variable construct 138

Figure 6.4 Parental reinforcement and discipline model diagram and loadings 146

Figure 6.5 Possible collective efficacy model definitions 160

Figure 6.6 Scatter of f-scores from both definitions of collective efficacy, 16-year-olds 162

Figure 6.7 Final area level model diagram and loadings 166

Figure 7.1 Theoretical schematic for the psychological map of interrelations 170

Figure 7.2 Covariance, mediation and moderation SEM path diagrams 175

Figure 7.3 Alternative representation of a moderation effect 177

Figure 7.4 Offending and individual attitude structural model 179

Figure 7.5 Offending and family reinforcement/discipline structural model 181

Figure 7.6 Alternative causal paths for neighbourhood level relationships.....183

Figure 7.7 Revised neighbourhood level relationships 184

Figure 7.8 Offending regressed on area perceptions 185

Figure 7.9 Reduced neighbourhood level relationships.....186

Figure 7.10 Adjusted neighbourhood level relationships 189

Figure 7.11 Individual attitude mediated by family reinforcement and discipline 192

Figure 7.12 Individual attitude mediated by collective efficacy and neighbourhood disorder 194

Figure 7.13 Individual attitude mediated by revised collective efficacy and neighbourhood disorder specification.....196

Figure 7.14 Individual attitude modelled by perceptions of both area and family situation.....198

Figure 7.15 Predicted probability of offending in final model, at differing standard deviation levels of characteristic.....199

Figure 8.1 Empirical model specification when altering the dependent variable 207

Figure 9.1 Model specification with control variables.....228

Figure 10.1 Autoregressive and cross-lagged simplex model of family level constructs 254

Figure 10.2 Autoregressive and cross-lagged simplex model of area level constructs..255

Figure 10.3 Modified autoregressive and cross-lagged simplex model of area level constructs 256

Figure 10.4 Modified autoregressive simplex model of area level constructs with no cross-lag.....257

Figure 10.5 Autoregressive simplex model of individual attitude 258

Figure 10.6 Schematic for longitudinal mediation 259

Figure 10.7 General lagged model path diagram 263

Figure 10.8 Self-report offending by age at interview for each wave 269

Figure 11.1 Empirical specification of the relationship between perceived latent constructs 284

List of Tables

Table 2.1 Sources of knowledge in DLC theory 35

Table 4.1 Self-report offending, 2004 to 2006, longitudinal dataset (all ages) 77

Table 4.2 Self-report repeat offending, longitudinal dataset 78

Table 4.3 Pair-wise Spearman's correlation between number of offences per year 79

Table 4.4 Offence categories grouping and original question reporting rate for 2004 79

Table 4.5 Percentage reporting one or more offence, by offence group and year, with Spearman's correlation between years 81

Table 5.1 Summary table of fit criteria under WLSMV estimator 104

Table 6.1 Individual level questions 123

Table 6.2 Exploratory two factor rotated loadings for items i9 to i14 127

Table 6.3 Empathy and Self-control/contentment model possibility results 128

Table 6.4 Family level questions 139

Table 6.5 Derived parental living situation 141

Table 6.6 Derived parental criminality variable frequencies 143

Table 6.7 Family level final and derived variables 143

Table 6.8 Additional items with parental reinforcement and discipline models 148

Table 6.9 Area level variables 152

Table 6.10 Collective efficacy variables 159

Table 6.11 Collective efficacy model fit and loadings 161

Table 7.1 Predicted probabilities of offending at various standard deviations from mean score of individual attitude 180

Table 7.2 Predicted probability for changes in parental reinforcement or discipline 181

Table 7.3 Predicted probability of offending for changes in neighbourhood disorder and collective efficacy (modelled separately) 185

Table 7.4 Predicted probability of offending for changes in neighbourhood disorder and collective efficacy (modelled separately) 190

Table 7.5 Predicted probability of offending for changes of individual attitude, controlling for family discipline and reinforcement 192

Table 7.6 Predicted probability of offending for changes of individual attitude, collective efficacy or neighbourhood disorder 195

Table 7.7 Predicted probability of offending for changes of individual attitude, collective efficacy or neighbourhood disorder with revised causal structure 196

Table 7.8 Proportion of mediation effect sizes in final model 200

Table 8.1 Grouped offences by base survey offence types, 2004 sweep of longitudinal dataset, 16 and under 210

Table 8.2 Number of self-report offences in last year by offence type, 2004 sweep of longitudinal dataset, 16 and under 212

Table 8.3 Proportion of mediation for various measures of all offending 215

Table 8.4 Proportion of mediation for various measures of violent offending 217

Table 8.5 Proportion of mediation for various measures of property offending 217

Table 8.6 Proportion of mediation for various measures of serious offending 218

Table 8.7 Proportion of mediation for various high frequency offenders 219

Table 8.8 Conditional correlations for the 2004 offending dependents 221

Table 8.9 Proportion of mediation for binary measures of different offending types on alternative datasets 223

Table 9.1 Percentage of OCJS sample in each socioeconomic classification, and percentage of each status by household income 233

Table 9.2 Control variables for inclusion 238

Table 9.3 Model results 240

Table 10.1 Level of factorial invariance 248

Table 10.2 Testing strong factorial invariance for family and area latent variables 250

Table 10.3 Testing strong factorial invariance for individual attitude 251

Table 10.4 Total and indirect effects of time point 1 collective efficacy on time point 2 neighbourhood disorder 256

Table 10.5 Lagged model results with individual attitude only 264

Table 10.6 Lagged model results with controls 265

Table 11.1 Model results for subdomains of the index of multiple deprivation, from section 9.3 305

Glossary

ALSPAC	Avon Longitudinal Study of Parents and Children
AP	Antisocial potential
BME	Black and Minority Ethnic
BMRB	British Market Research Bureau Limited
CAPI	Computer-assisted personal interviewing
CASI	Computer-assisted self-interviewing
CDF	Cumulative distribution function
CFI	Comparative fit index
CJS	Criminal justice system
CSEW	The Crime Survey for England and Wales
DCLG	Department for Communities and Local Government
DLC	Developmental and life course
EFA	Exploratory factor analysis
ESRC	Economic and Social Research Council
ESTYC	Edinburgh Study of Youth Transitions and Crime
ICAP	Integrated cognitive antisocial potential
IDACI	Income deprivation affecting children index
IMD	Index of multiple deprivation
ISRD	International self-report delinquency study
LSOA	Lower layer super output areas
LSYPE	Longitudinal Study of Young People
MI	Modification indices
NatCen	National Centre for Social Research
NCDS	National Child Development Study
NSSEC	National Statistics Socioeconomic classification
OCJS	Offending, Crime and Justice Survey
ODPM	Office of the Deputy Prime Minister

OLS	Ordinary least squares
ONS	Office for National Statistics
PADS+	Peterborough Adolescent and Young Adult Development Study
PHDCN	Project of Human Development in Chicago Neighbourhoods
PNC	Police National Computer
PPO	Prolific and priority offender
RMSEA	Root mean square error of approximation
SAT	Situational Action Theory
SEM	Structural Equation Modelling
SES	Socioeconomic status
SOA	Super output areas
TLI	Tucker-Lewis Index
WLSMV	Weighted least squares with missing values
WRMR	Weighted root mean square residual

Chapter 1

Introduction

1.1 Motivation

"Does the area in which you live affect your offending?" is an attractive criminological question for investigation, not least due to the familiar idiom 'growing up on the wrong side of the tracks' (Young, 2003), but the answer to this question has been surprisingly elusive. Many of the risk factors associated with adolescent offending are well known (see, for example, Farrington and Welsh, 2007). Individual factors such as low self-control, empathy or impulsivity are well-researched and correlate with increased offending behaviour. Similarly, family factors such as parental discipline, parent/child relations, and parental conflict have also received much attention. Less, however, is known about 'higher' level processes such as school, peer and, particularly, the character of the area in which the child lives. This despite neighbourhood being consistently linked to a number of poor educational, emotional and health outcomes (Leventhal and Brooks-Gunn, 2003).

Early work in the University of Chicago around the middle of the 20th century led to some influential theories as to how area and space affected crime, known under the umbrella terms of 'social disorganisation' and 'social ecology'(Shaw and McKay, 1942) but a combination of criticisms, both substantive and methodological, led to this school of thought falling out of favour and criminology to focus more on individual pathology, leaving the wider articulation of components of social setting as 'a road not taken' (Messner, 1988). Partly attributable to this individual focus and, also, for pragmatic data collection reasons, data allowing area level analysis have been of scarce supply (Farrington, 2005). Whilst 'risk factors' have thus been thoroughly investigated and understood at the individual and family level, there remains a lack of concrete findings at the community level. In part because of the popular significance of the 'broken windows' theory of policing (Wilson and Kelling, 1982) and work by Sampson and

colleagues (Sampson and Groves, 1989) including the extremely influential field of collective efficacy (Sampson, Raudenbush and Earls, 1997; Sampson, 2012), recent developments have made steps to understanding and theorising at the neighbourhood level again. At present the evidence is limited, restricted to either specific crime types (McVie and Norris, 2006b; Gibson, 2012), certain age groups (Odgers et al., 2009), or within restricted settings (Sampson, 2012; Wikström, 2012) and this limit has often been caused by limited data sources (Sampson, Morenoff and Gannon-Rowley, 2002).

This thesis has two aims: the first is substantive and is to investigate whether and how the area in which an individual lives has an effect on their offending, with the second to try to answer this question from an existing and publicly available dataset, to try to prove that this substantive question can be answered using data that has previously been deemed lacking the required detail. Secondary data analysis is an important tool in the social researcher's armoury but is perhaps not as prevalent as it could be, so much so that the Economic and Social Research Council (ESRC) set up a specific secondary data analysis initiative that "aims to deliver high-quality high-impact research through the deeper exploitation of major data resources created by the ESRC and other agencies"[1]. More specifically, this work will demonstrate that theoretically-informed, structural equation modelling can make best use of existing and often under-utilised datasets, particularly national studies such as those typically conducted by large scale organisations or governments. Structural equation modelling (SEM) is an ideal technique as it allows for detailed measurement of concepts to be undertaken, and for the relationships between the measured concepts to be linked, whilst offending behaviour is an ideal test bed for such a task as there is a long tradition of empirically tested theoretical arguments. These theoretical arguments are vital for a SEM approach as they can be used to inform relationships and avoid the potential pit-falls of data-mining.

[1] http://www.esrc.ac.uk/research/skills-training-development/sdai/ [accessed May 2015]

1.2 Overview of the Thesis

This thesis is structured as one coherent and complete piece of work with each chapter or section building upon the previous part. Chapter 2 gives some more background for the work and fleshes out some of the statements made in the motivation section of this introduction. It reviews the contribution to knowledge about offending behaviour made through developmental and life course (DLC) theories of offending and focuses on some of the known risk factors, highlighting the gap in findings about the area. It also reviews where criminological work has focused on the geographical in the fields of social disorganisation and collective efficacy, before moving on to more recent developments such as situational action theory. In doing this it places the current work as an attempt to work within the risk factor paradigm and highlights a gap in the evidence base about how area affects the behaviour of those in young adolescence. Chapter 3 takes this realm of investigation and turns it into a specific research agenda, setting out in detail the questions to answer. These split the work into methodological and substantive goals though acknowledges that this is something of a false dichotomy meant only to aid to the structuring of the analysis.

Chapters 4 and 5, respectively, define the data to be used in the work and give an overview of structural equation modelling. The data are drawn from the Offending, Crime and Justice Survey (OCJS) that was conducted on behalf of the Home Office between 2003 and 2006 and is the only nationally representative panel study with self-report offending in the UK. It was created in response to a known gap in data on offending in the general population but has been used in relatively few publications. Chapter 4 reviews the self-report methodology and a few of the preliminary problems with the dataset before defining the specific dataset that will be used for the remaining thesis. Chapter 5 gives an overview of structural equation modelling (SEM) its historical context and outlines how it shall be used in this work. This review is not intended to be a comprehensive explanation of the method and more detail is worked into later chapters where appropriate but gives enough detail for an understanding of the processes involved in fitting measurement models, creating structure from these and

some guidance on interpretation. It also contains a section on what SEM approaches mean for the ontology of the work and any associated causality.

Following this, Chapter 6 is the first empirical chapter proper and starts with creating measurement models of offending risk factors based on the OCJS and, in keeping with the different level as talked about in Developmental and Life Course theory, splits these into three levels, the individual, the family and higher level factors, of which area is the focus. The work progresses in that order creating first a criminogenic propensity measure (which is labelled 'individual attitude'), then a set of measures looking at perceptions of the family environment and, finally, measures of collective efficacy, neighbourhood disorder and fear of local crime. Each of these sets of measures is approached analytically and from a theoretical viewpoint to ensure that the measurements of these concepts match to previously used operationalisation and that their measurement strengths/weaknesses are known.

Chapter 7 starts by defining a theoretical schematic into which to fit these measured risk factors and discusses at length the role that criminogenic propensity has played in previous theoretical discussions around developmental and life course criminology. Finding a unanimous story of early development and then relative stability, it uses this to define the interaction of the family and area level measured concepts and crucially makes a distinction between perceived factors and those that are measured externally, or with less scope for attitudinal variation. Through this argument it is shown that mediation analysis is the most appropriate way to fit these items together. Once this schematic is defined the analytical works starts with the simplest structures before building up to what becomes the main operationalised model for the remainder of the thesis. Some initial substantive results show the centrality of the individual attitude component in the model and indicate that this is mediated by changes in perceptions of neighbourhood disorder, with part of that relationship operating through collective efficacy (which in itself has no direct effect on self-report offending). Interestingly, family level measures appear to have no effect on reported offending and this becomes an important debate later in the work.

The next two chapters go on to try to extend this basic cross-sectional model to look at different crime types available within the OCJS and to incorporate risk factors that go beyond those reported according to the individual's perceptions. The first part, Chapter 8, finds mixed results with initial analysis showing that the mediation of the area level variables seems to be strongest for property offending and reduces for serious and further for frequent offending, but further analysis provides some doubt as to the robustness of these results. Chapter 9 produces more concrete findings. Specifically, it finds that the mediation effects by perceptions of neighbourhood do not disappear once structural characteristics of the area are included in the model and that reported family contact with the criminal justice system seems to increase self-reported offending. Age and gender also play an important role.

The final analytical chapter, Chapter 10, attempts to bring in the longitudinal possibilities of the data and produces a number of interesting and important findings. After reviewing the stability of the measurement models over the period available (2004 to 2006) it finds high levels of year on year stability in the concepts included in the work, especially in the reported individual criminogenic propensity. The chapter moves on to outline an approach to the longitudinal modelling that sidesteps some of the problems with creating the most complete mediated longitudinal model in the purest sense. This lagged model produces findings that indicate that although it is the area that operates as a mediating effect at any one point in time it is the family environment that seems to be operating over the longer term.

Chapter 11 draws together the entire thesis, presenting a detailed review of the analysis and results before going on to draw some implications. These fall into three categories of policy implication, those for secondary data analysis within criminology and more generally and some about the design and data collection procedures for the future. The work draws to a close with avenues for further work, both on the OCJS dataset and in the wider field of criminology.

This thesis is important because of three reasons. First, it demonstrates that with appropriate attention to theoretical structures SEM can be used to get results from existing datasets that have previously been thought impossible. Second, it finds

evidence that a combination of both perceptions of neighbourhood disorder and perceptions of local collective efficacy have a role in shaping adolescents' offending behaviour. Finally, it also uncovers the temporal nature of the neighbourhood in the determining offending behaviour, with the family environment ultimately taking the most prominent role.

Chapter 2

Background

2.1 Introduction

This chapter is intended to provide a picture of the background upon which this work is set. It will focus on risk factors that have been found from the developmental and life course approaches to criminology and highlights the lack of knowledge at the area level, whilst setting this against the well-developed fields of social disorganisation/collective efficacy theory. In doing so it highlights the gap that is to be addressed, namely how area (as classified by its collective efficacy and other well-established cross-cutting neighbourhood level mechanisms) acts as a risk factor for adolescent offending. The chapter is not exhaustive of all studies that have been undertaken, which would probably form a full thesis in and of its self and be something to rival Ellis, Beaver and Wright's (2009) *Handbook of Crime Correlates.* Nor is it comprehensive of all the variant theories that have been proposed to underpin deviant behaviour (reviewed in textbooks such as Cullen and Agnew, 2013) but provides an overview of the main arguments. More technical details relating to specific parts of these works, and others of relevance, will be reviewed in the chapters later in the thesis.

2.2 Developmental and Life Course Criminology and Risk Factors

From Lombroso's *Criminal Man* (1876, republished in 2006) and his focus on the (pseudo)science of phrenology, the roots of criminology most probably lie in the late 18th and early 19th century discourse on "moral insanity and uncontrollable, remorseless criminal behaviour" (Rafter, 2004, p979) and it has been a constant of criminology to attempt to understand and classify the criminal, the reasons for their behaviours and ways to control and alter these. Since the middle of the 20th century and the seminal work of Glueck and Glueck (1950), the life course perspective on criminal behaviour has come to the fore and has been extremely influential in criminology. The

Gluecks' study *Unravelling Juvenile Delinquency* examined, before the aid of powerful computing technology, the life course of two sets of offenders, one of which had been convicted of crimes whilst the others remained, by and large, free of convictions. Their analysis revealed multiple risk factors which, they claimed, propelled some people towards offending. The work established a precedent for the application of scientific methods to try to unpick the offending trajectories of individuals.

The earlier works that shed light on this idea of individual trajectories or criminal careers were, counter-intuitively, cross-sectional. The Philadelphia cohort study, which captured all boys born in the city in 1945 and aged 10 to 18 at the time of the data collection, discovered for the first time that a relatively small number of offenders were responsible for the majority of crimes (Wolfgang, Figlio and Sellin, 1972) and comparable results were found based on data collected for similar cohorts from Racine, Wisconsin (Shannon, 1982) and later from the UK (Farrington and West, 1990). Together these studies highlighted the existence of what was referred to as a 'career criminal' or more precisely individuals whose offending remained relatively frequent and continuous over a significant period of their life course. Another key finding that propelled this movement is the age/crime curve (McAra and McVie, 2012). This is found, ubiquitously, when the age of the offender is plotted against criminal activity rates and shows a sharp rise in offending during the teenage years followed by a slow decline across adulthood. The exact peak and shape of the curve depends on the source of the crime data (official versus self-report), the gender of those in the study, where the data originates and the type of offending that is being examined but "this distribution… represents one of the brute facts of criminology" (Hirschi and Gottfredson, 1983, p552). Whether the patterns applies to the individual has been the subject of much contested debate, and it has often been found that the individual trajectory differs from the aggregate, with the latter often explained by changes in individual time-varying engagement in offending behaviour rather than offending frequency (Farrington, 1986).

Piquero, Farrington and Blumstein (2003) identify a 1986 publication by the National Academy of Sciences that reported on criminal careers and career criminals (Blumstein et al., 1986) that marked a particular turning point in the work of life-course

criminology and raised a number of theoretical issues that needed to be understood for life course criminology to progress adequately. Together these theories form the background to developmental and life course theories of offending which are concerned with three main issues: the development of offending and antisocial behaviour; risk and protective factors at different ages; and the effects of life events on the course of individual development (Farrington, 2005).

This centrality of theoretical argument within this field is well recognised, with Farrington, Snyder and Finnegan (1988) opening their work that looks at offending specialisation by stating that "in investigation of the causes of delinquency, it is important to establish the theoretical construct(s) underlying delinquent behaviour" (p461). They move on to outline two possible theoretical positions, the first where the delinquent behaviour is caused by a single underlying factor such as 'delinquent tendency' or 'antisocial tendency'. The alternative position they sketch is where the behaviour reflects a number of different dimensions. They see versatile offending (i.e. offending across a variety of different crime types) as the result of there being just one underlying theoretical construct that drives offending, whilst specialized offending would be the result of multiple constructs on which different people occupy different positions, given relative stability on those dimensions. Their work highlights a previously uncovered amount of specialisation that contrasts the highly regarded 'cafeteria style' versatile offending previously defined by Klein (1984). Although subsequent work that has delved into the issues of specialisation has found that the picture is not entirely clear and is often a symptom of the crime classifications, counting rules and methods of classifying multiple offending episodes (Fisher and Ross, 2006), the distinction between the two types of theories, single or multiple causes, is seen throughout life-course criminology. The former, known as *general offending pattern theories*, are mainly characterised by an identification of low 'self-control' and are reviewed extensively by Gottfredson and Hirschi (1990), whilst the more developmental approaches, as identified and studied by Loeber and Leblanc (1990), are characterised by different causal factors that may themselves change throughout life.

As an example, Gottfredson and Hirschi's general theory of offending hypothesises that an offender's self-control will be the principal factor determining an individual's involvement in crime. 'Self-control' in their terminology incorporates elements such as self-centredness, impulsivity, perseverance, and an inclination to thrilling or risky activities. Variation in the level of self-control can be primarily attributed to parenting practices, such as lax supervision, ambiguous affectional interaction and inconsistent discipline. They fundamentally connect a lack of self-control with involvement in criminal behaviour, and that the lower the onset age of offending, then the lower the self-control and the higher the versatility in resultant offending, and that this self-control is invariant. As such, offending artefacts such as the age-crime curve are attributable to biological processes associated with aging. This self-control is, predominantly, based on individual factors and early life experiences, beyond the influence of the physical setting in which the person finds themselves. Farrington (2005) sees this in essence as 'anti-developmental' in that the cause of offending, self-control, is fixed in childhood (dependent on the socialisation process), leaving no space for situational change.

Tittle (1995) recasts control theory into a control *balance* where deviant acts (a behaviour that is likely to be against the normative opinions of the majority and typically leads to some form of negative societal response) can be the consequence of being controlled *by* others or exhibiting control *over* others. A particular deviant behaviour thus becomes a balance of deviant motivation and constraints, which result in the probability of a range of different offence types that either express hostility and contempt towards a controlling force (individual, group or norm) that shows a control deficit, or are a selfish behaviour that acts autonomously with little regard for others (again individuals, groups or norms) representing a control surplus. Empirical tests of the theory have generally found limited and mixed support (Delisi and Hochstetler, 2003; Piquero and Hickman, 1999) and often for only certain types of offence. This perhaps isn't surprising as in the original work Tittle (1995) is sceptical about the ability to test the theory with existing data due to the nuances and complexity at the very heart of the control balance.

Sampson and Laub's (1993) age-graded theory of offending focuses more on the social control side of control theory and sought to explain changes in criminal behaviour as influenced by changes in the presence, strength and quality of bonds to societal informal social controls. Their goal was to bring together research into both childhood and adulthood, and the continuity of offending through the life course. During adolescence they identify social bonds as primarily with the school and family but also with their wider neighbourhood and peers. In adulthood these social bonds alter to employment, marriage and other long lasting relationships, and these bonds can alter early-age criminal propensities. What is important to the theory is criminal involvement is reduced not just by the presence of the bonds but also their quality and how these mediate the context of social structural influences. The theory is tested against data gained from a 50-year follow-up of the Gluecks' Boston cohort and as such has a strong life course analysis empirical foundation. Their highly readable follow-up *Shared Beginnings, Divergent Lives* (Laub and Sampson, 2003) deepens this understanding by identifying 'turning points' away from an offending lifestyle such as moving home, good marriages, time in the army and enjoyable, trusting employment.

Moffitt (1993) provides a more developmental theory. Whilst she identifies a small percentage (between approximately 5% and 10% of the offending population) as 'life course persistent offenders' who are tied into offending from very early ages, triggered by genetic disposition and maternal failure such as poor nutrition and parental substance abuse, the majority of offenders are "adolescence-limited". This second group are generally well-socialised and adjusted, and their delinquency is adaptive to reinforce certain stages of the maturing process. She identifies maturity gaps, problematic transition through puberty, and desires for adult status and privileges as key factors that drive these offenders' behaviour and as these problems subside their delinquent behaviour will alter or diminish. She states offenders in the second group, the "adolescence-limited" offenders, will partake in crime to address what are the most pertinent inadequacy factors of the time, such as public order offences, substance abuse and 'status' crimes such as running away and theft. Moffitt directly confronts the environmental aspect of childhood upbringing as one of the many variables in her multidimensional causation model - it is engaged in a constant process of reciprocal

interaction: "It is now widely acknowledged that personality and behaviour are shaped in large measure by interactions between the person and the environment" (Moffitt, 1993, p9). These environments are most clearly articulated as the family condition (financial, behavioural and cognitive) with attention to the area in which the child grows up added as an unconsidered adjunct: "Vulnerable children are often subject to adverse homes and neighbourhoods because their parents are vulnerable to problems too" (p9). The adolescent limited group, in trying to address maturity gaps, may interact with peers in specific physical spaces, and there is limited reference to the protective effects on this group growing up with limited criminal opportunity (such as in rural settings) or under particularly strict supervision arrangements. This debate is, however, framed in terms of the individual development. The area effect, although touched upon and incorporated as one of many situational factors that can mediate, is not fully explored.

Farrington's Integrated Cognitive Antisocial Potential (ICAP) theory (Farrington, 2005b) developed out of his years of work formulating predictive risk factors of offending, and was in part a reaction to criticism that the nature of the criminal career model that was prominent in the 1980s was atheoretical (Piquero et al., 2003). His resultant ICAP model tries to integrate theories such as strain, control, learning, labelling and rational choice, with a key underlying construct of antisocial potential (AP) and that the transfer of this to antisocial behaviour is dependent on both opportunity and victims (Farrington, 2005b). AP, the potential to commit antisocial acts, can manifest itself over the short and long term – short-term variations depend on motivating and situational factors, whilst long-term variations depend on impulsiveness, strain, modelling/socialisation processes and on life events. The model, laid out in diagrammatic form (p78), raises the role of neighbourhood both in terms of the development of AP and in relation to the opportunity to commit crime, and is seen as an important aspect: "criminologists should carry out more research on situational influences on offending" (p85).

Although this chapter has covered a number of competing different theories within the DLC paradigm it is not the goal to cover all of their intricacies. These will be reviewed

later in the thesis where greater detail is called for (notably in Chapter 7). What is important is that many of these theories have led to empirical testing, and that empirical testing has led to more theories (Lilly, Cullen and Ball, 2014), creating something of a virtuous circle (though Lilly, Cullen and Ball are slightly scathing about some aspects of this[2]). As Farrington notes: "The main reason why DLC paradigms became important during the 1990s was because of the enormous volume and significance of longitudinal research on offending that was published during that decade" (2005, p3). Piquero et al. (2003) highlight "a relatively small number of classic and contemporary studies [that] are especially important" (p363). Table 2.1 at the end of this section presents a summary of the main features of those listed, plus some others of relevance. The risk factors generally defined from these sources, and common to both DLC theory and the previous criminal career paradigm as far back as the Gluecks' study, are well-established and collated results are widely reported (Farrington and Welsh, 2008; Rutter, Giller and Hagell, 1998; Jolliffe and Farrington, 2010; Blackburn, 1993) and shan't be rehashed here. They broadly split into one of three categories:

- Individual factors such as low intelligence and attainment, low empathy, impulsivity, 'poor' temperament (which although hard to define is discussed in terms of 'inhibition' (Kagan et al., 1989) or 'under controlled' and a lack of social cognitive skills (Caspi, 2000)).
- Family factors such as large family size, criminal families, poor parental supervision, childhood physical punishment, low parental reinforcement, erratic discipline and inconsistent rewards, low parental warmth, child abuse and neglect, and parental conflict. For this last factor the exact processes are complicated: when controlling for 'lower level' factors, Kolvin (1988) found that separation from a biological parent before the age of 5 doubled the conviction risk to age 32, whilst Henry et al. (1996) found that boys from a single parent family were more likely to be convicted. Fergusson (1992) found that parental

[2] To quote, "One (but not the only) reason that theories flourish is that they are able to provide scholars with opportunities to conduct research and gain publications—the very accomplishment that allows for tenure and career advancement" from p109, chapter 6 of *Criminological Theory* by Lilly, Cullen and Ball (2014).

changes (i.e., to step-parents or to single parent families) had no effect on conviction rates. Capaldi and Patterson (1991), however, found a parental transfer effect, but traced it though antisocial mothers that caused these transitions, which in turn caused antisocial behaviour. Haas et al. (2004) and Juby and Farrington (2001) both see the post-disruption trajectory as problematic rather than the separation event itself, adding weight to the life course model rather than trauma or selection theories.

- 'Higher' level factors such as socioeconomic status (SES), peer, school and community factors – evidence here is more mixed. SES is classically associated with higher offending (Cohen, 1955), although Thornberry and Farnworth (1982) review the literature and find a mixed message, often caused through a confluence of the individual and familial factors occurring in conjunction with a poor economic position. Peer and school influences find similarly mixed results, although where effects are seen the causal pathways are not clear. Similarly with community level factors, other than residing in a high crime neighbourhood, "less is known about community or neighbourhood risk factors, or about the development of offending in different neighbourhood contexts" (Farrington, 2005a, p80).

Most of the studies listed in Table 2.1 for predominantly pragmatic reasons, were conducted in small geographical areas and, with the notable exceptions of the Project on Human Development and the Thousand Family Study (which were specifically designed to be able to tap neighbourhood level variation), are mainly unsuited to geographical community or neighbourhood analysis. This lack of geographically suitable datasets is one of the reasons for the lack of understanding of higher level risks. From a policy perspective too, the classification of the individual offender makes policies that are easier to enact at the macro level (Dowden and Andrews, 2004). Messner (1988) sees this as a consequence of the two analytically distinct theoretical arguments within Merton's 1938 essay *Social Structure and Anomie*, the first a theory of social organisation and the second a theory of deviant motivation. The former is an articulation of components of social systems and has been neglected (Messner, 1988), whereas the latter, with its focus on "man's imperious biological drives which are not

adequately restrained by social control" (Merton, 1938), has received significant attention.

There are exceptions to the rule that higher level community factors have been unanalysed, one example being Wikström and Loeber's (2000) work using the Pittsburgh Youth Study. They found that those with high criminogenic risk factors (defined as those with hyperactivity-impulsivity-attention problems) had no perceivable difference based on their neighbourhood in regards to their serious juvenile offending, yet for those with lower risk factors, living in disadvantaged areas significantly increased the probability of offending. In the most disadvantaged areas, those with a majority of public housing, the neighbourhood effect was so strong as to swamp the individual risk factors. Separating out the effects of the neighbourhood from a general cohort effect proved beyond the data source and an alternative reading could be that the poorer areas had people with a high combination of risk scores, therefore despite low singular risk factors they offended at a higher rate, which drowned out those very risk scores. McVie and Norris (2006b), using the Edinburgh Study of Youth Transitions and Crime[3] data and looking at adolescents aged 12 to 17, found that perceptions of areas that were poorly controlled by adults led to increased property

[3] The Edinburgh study deserves special attention as the impressive data source has resulted in some analytically significant findings for the current work. The reason this discussion is confined to a footnote is that the empirically grounded theories generated from the data are about the systemic interactions of the individual with the social structures around them and therefore have a different focus to the risk factor paradigm approach being applied here. Based on a sample of 4,300 young people who started school in Edinburgh in 1998, the project captured six sweeps of quantitative interviews with the adolescents, a series of qualitative semi-structured interviews with around 55 cohort members, teacher and parent questionnaires and administrative data on social work, police and school records. McAra and McVie's (2012) negotiated order theory is something of a broader variant of labelling theory and explores how different types of regulation, both formal (schools, the police, social workers, etc.) and informal (family, peers, street culture) interact with the individual's identity and form a complex web of pathways that need to be negotiated in order for the young person to retain a sense of self integrity (or 'ego continuity'). Exclusion by formal agencies is seen as being "profoundly exclusionary for certain types of young people… [and]… appears to limit the capacity to negotiate actively, leading to a downward spiral of increased marginalisation" (McAra and McVie, 2012b, p545). They highlight "four key facts about youth crime which any system of youth justice 'ought to fit'" (McAra and McVie, 2010, p212) that include persistent serious offending being closely linked to victimisation, that early identification of at-risk children is not fool proof, that critical moments in development over adolescence can have a lasting impact and that diversionary strategies can aid desistance.

offending though the influence seemed to decrease as the respondents grew older. Other work has either focused on wholesale changes to the area in which people lived (such as Osborn, 1980, who found that moving out of London decreased offending or Elliot et al., 1989, who looked at urban versus rural offenders), classified areas by their social housing rates (Baldwin et al., 1976; Bottoms and Wiles, 2002), or focused on specific crime type (notably violence, for example Beyers et al., 2001).

Table 2.1 Sources of knowledge in DLC theory

Study	Sample and detail	Location	Measure
Glueck and Glueck Unravelling Juvenile Delinquency	500 delinquents matched to 500 non-delinquents. White males age 10 to 17 beginning in the 1940s. Followed up at 14, 25 and 32. Further follow-up at 70 with death record and criminal history searches, 52 contacted for qualitative life course history interviews (Sampson and Laub, 2003).	Boston, Massachusetts. Delinquent group drawn from 2 reform schools	Official and self, teacher, parent-report records
McCord's Cambridge-Somerville Youth Study	506 boys identified by local police, shop keepers and social workers matched on range of characteristics and split into control and treatment groups	Cambridge-Somerville, Massachusetts	Official records
1945 Philadelphia Birth Cohort	9,945 boys born in 1945, followed to age 17. 35% found to have had contact with the police at least once, 6% were 'chronic offenders'	Philadelphia	Official records
1958 Philadelphia Birth Cohort	27,160 boys and girls born in 1958, followed to 17, then extended to 26 (Tracey and Kempf-Leonard, 1996). Similar results to 1945 cohort, plus findings that career continuity more likely than discontinuity.	Philadelphia	Official records
1942, 1949 and 1959 Racine, Wisconsin	1942, n = 633, 1949, n = 1,297, 1955, n = 2,149 with a similar design to the Philadelphia study	Racine, Wisconsin	Official records
Cambridge Study in Delinquent Development	411 males born between September 1952 and August 1954. Interviews at 8, 10, 14, 16, 18, 21, 25, 32 and 46. Parental interview annually between 8 and 15. Criminal record search to age 40, records of immediate family also retrieved.	South London, UK	Official records and self-report
National Youth Survey	1,725 male and female adolescents aged 11-17. First wave in 1977, then 9 subsequent waves until 1993.	US national representative survey	Self-report
Montreal Adjudicated Youths	505 male and 150 female delinquents sentenced to probation or a correctional institution at the Montreal Juvenile Court. Started in 1992 with an average age of 15.	Montreal	Official records and self-report
Montreal Adolescent Sample	3,070 adolescents age 14 in 1974, followed up to age 40. Subsample of 458 interviewed at 16, 30 and 40	Montreal	Self-report records
Project Metropolitan	15,117 males and females born 1953 and collecting data from police recorded crime data	Stockholm	Official records

Study	Sample and detail	Location	Measure
Individual Development and Environment	1,027 8- and 9–year-olds in 1965, followed from 10 to 30, collecting administrative data on criminal offences, mental health, employment, education, alcohol abuse	Orebro, Sweden	Official records
Pittsburgh	1,517 boys in public schools followed at age 7,10 and 13, part of the *Causes and Correlation study* ran by the Office of Juvenile and Delinquency Prevention in the US Dept of Justice and as such has comparability with the other studies included. Data collected through interview.	Pittsburgh	Official records and self-report
Denver	1,527 youths from 'high risk' Denver neighbourhoods, part of the *Causes and Correlation study* ran by the Office of Juvenile and Delinquency Prevention in the US Dept of Justice	Denver	Official records and self-report
Rochester	1,000 children from Rochester, aged 12, part of the *Causes and Correlation study* ran by the Office of Juvenile and Delinquency Prevention in the US Dept of Justice	Rochester, New York	Official records and self-report
Dunedin Multidisciplinary	1,037 children (or 91% of eligible births) born 1972-73, predominantly to study health outcomes but also crossing over to criminal/antisocial behaviour	Dunedin, New Zealand	Official records and self-report
Project on Human Development	6,500 individuals in 80 communities in Chicago. This multidisciplinary study includes an accelerated longitudinal design following 9 different age groups (spaced by 3 year intervals) as well as social observations of neighbourhood characteristics.	Chicago	Official records and self-report
Nottingham Study	Class stratified sample of 700 children, with parental (at younger than 7) and self (at 11 and 16) reported offending questions and other characteristics to the age of 16 (Lewis, Newson and Newson, 1982)	Nottingham, UK	Self-report, parental report
Thousand Family Survey	1,142 infants born between 1st May and 30th June 1947, collecting interview data through childhood and administrative data through adulthood	Newcastle-Upon-Tyne, UK	Official records
Edinburgh Study of Youth Transitions and Crime	4,300 young people starting secondary school in 1998. Six annual sweeps of self-report surveys from cohort members (aged 12-17) as well as official records from police, social work, children's hearings, schools and criminal convictions records; surveys of parents and teachers; a community survey; and compilation of a Geographic Information System incorporating census and police-recorded crime data.	Edinburgh, Scotland	Official records and self-report

2.3 The Chicago School, Social Disorganisation and Social Ecology

2.3.1 The Old Days

What seems clear is that there is some effect of neighbourhood on offending but the exact nature of this has been obscured by a lack of data and by an 'area' being able to be classified in a myriad of different ways. There is, however, a great and proud tradition of more specific neighbourhood research stretching back to the early days of the University of Chicago and predates the emergence of DLC approaches by some way. The focus of these theories was not the individual; rather they operated at a macro level, concerned more with total offending rates or locations that offences occur, and the characteristics of the area that can give rise to these events. One of the early key concepts that was used to cut across different community contexts was that of *social disorganisation*.

Between 1837 (when it was first chartered as a city) and 1910, Chicago's population grew from just over 4,000 to 2.2 million, and was for a period the fastest-growing city in the world (Encyclopaedia of Chicago, 2004). This growth, with associated social and financial opportunities, came coupled with increases at the other end of the income distribution: poverty and social problems mushroomed. This led to Chicago becoming a natural laboratory for sociological research in what came to be known as the 'Chicago School'. Based in the University of Chicago, sociologists such as Robert Park and Ernest Burgess laid the foundation for urban sociology by collecting both statistical and qualitative data and, drawing analogy with zoological and botanical observations, proposed a model of delinquency as a 'social product' of the city environment (Park and Burgess, 1920). This shifted the theoretical focus of Chicago's criminological work from an emphasis on individual pathology as the cause of crime, to a social pathology rooted in the social, cultural and structural forces accompanying the social changes taking place (Lanier and Henry, 2010).

Clifford Shaw and Henry McKay, two of the School's most important members, used Burgess's (1925) model of a concentrically-zoned city structure and, believing that "the best basis for an understanding of the development of differences among urban areas

may be gained through the study of the processes of the city growth" (Shaw and Mackay, 1942, p18), examined 56,000 court records from 1900 to 1933, creating "spot maps" of delinquent juvenile residences across Chicago. These they overlaid with community factors, such as demolished building locations, vagrancy rates, incidences of tuberculosis, infant mortality rates, levels of mental disorders and land usage. Analysing these physical, economic and population characteristics in comparison to the "more subtle differences in values, standards, attitudes, traditions and institutions" (ibid., p164), they found community problems concentrated in zone 2, the 'zone of transition', matched the higher official crime rates and that these patterns were consistent over the 33 years from which they collected data. These results were robust regardless of the inhabitants of the zone, whose composition was changing in each period through fresh waves of immigration. They argued that this second zone, just outside of the Central Business District, was dislocated from formal institutions with informal social controls eroded by the diversity and population renewal. The consistency of the results over the time periods studied implied persistence tied to the landscape which could be ultimately traced to socioeconomic forces.

These results reinforced the social disorganisation concept, first developed by Thomas and Znaniecki (1920), that the crime rate would be higher in areas in "which there is little or no community feeling, relationships are transitory, levels of community surveillance are low, institutions of informal social control are weak and social organisations are ineffective" and "delinquency patterns themselves become competing lifestyles as a means of surviving and as a way of obtaining income, intimacy and honour" (Lanier and Henry, 2010, p229). Moving beyond the basic disorganisation model, social ecology theory utilised Burgess's ecological analogy that order is found in stable zones of the eco-structure, suggesting that "crime and deviance were simply the normal responses of normal people to abnormal circumstances" (Akers, 1994, p142).

Subsequent work by Shaw and MacKay went on to find the same patterns in eighteen other cities leading them to conclude that "delinquency – particularly group delinquency… has its roots in the dynamic life in the community" (Shaw and Mackay, 1942, p435). Whilst influential, especially in terms of the effect of economic composition

of communities (Bursik, 1988), there have been many notable and strong criticisms – from Alihan's (1938) observation of the ecological fallacy inherent in the Chicago School work, to Kobrin's (1971) criticism that the counterfactual was not adequately addressed (i.e. that desirable areas were not shown to be highly organised). Bursik (1988) highlights the tautology inherent in the work – that high crime areas are a result of high delinquency traditions. Robison (1936) questioned the reliance on official records of crime rates, with Johnstone's (1978) self-report study finding different patterns not reflected in official figures. Similarly, Baldwin and Bottoms (1976) have argued that it is important to scrutinise how offending (rather than offender residence) could be plotted; the two may be correlated but need not correspond. One of the backbones of the work, and probably one of the principal findings considering the political backdrop in which the work was set, that race was independent of any neighbourhood problems was contested by contemporaneous research that found that "Oriental" residents had lower rates of delinquency (Hayner, 1933). When coupled with broader trends in criminology, after a brief period of prominence in the 1950s and 1960s, the attention paid to social disorganisation theory declined steeply and led Davidson (1981) to remark that social disorganisation "should be seen as a descriptive convenience rather than a model of criminogenic behaviour" (p89).

2.3.2 Modern Directions

Despite the criticism of social disorganisation theory and the emergence and prominence that the life course approaches took, social ecology did not disappear but developed in different but related directions. In Jane Jacobs' (1961) book, *The Death and Life of Great American Cities,* she asserts that "crime flourished when people did not know and meaningfully interact with neighbours… high levels of natural surveillance create a safe environment". This was a variant of control theory, moving social ecology theory into a holistic setting that incorporated both the physical and the social; Jacobs' theory of physical space had a profound influence. Oscar Newman (1973) saw creating safe, defensible spaces as a primary concern of the urban designer.

This social side of Jacobs' social ecology can be seen in the first of the four key components of what Bursik and Grasmick (1995) refer to as the 'systemic model of

neighbourhood crime' that started to grow through the 1980s and 1990s. Within this first component it is the networks and ties of the components of the system, the neighbourhood and the individuals within it, that provide the vehicle for information exchange which can regulate and affect the crime rate. As with Farrington's (2005a) observation on the criminal career paradigm and DLC theory, Bursik and Grasmick (1995) state that research has predominantly focused on the private level (within families or close friendship groups) as "analysts rarely have data pertaining to the precise nature of the communications that are transmitted through these [neighbourhood] networks" (p114). Despite the lack of research, they extend the scope of this feature to a looser "parochial level of control" where fellow neighbours who do not normally share the same intimacy as the private network may keep an informal eye on local children or raise the alert of threats such as the presence of 'undesirable' outsiders.

The second feature of their systemic model of social ecology is the recognition that social organisation can take many different forms which may not be immediately apparent, and will depend on the size, the heterogeneity of the area (by ethnicity, socioeconomic status, length of residence, etc.), the density of the network and the content or nature of the bonds of the network. The directness (unidirectional or reciprocal), durability and the frequency of the contacts also play a part. Again, practical data collection reasons have meant that many of these "variations in 'systemness'"(p122) have not received a full analytical study. A further feature of the systemic model is that networks change over time, dependent on the addition or loss of members, without the complete dissolution of the system. Differing significantly from the Shaw and McKay model, this explicitly acknowledges that neighbourhood relational networks are embedded in broader ecological structures of the urban landscape. The final feature is that the system is open, meaning not only that it is engaged in interchanges with the environment, but that this interchange is an essential factor underlying the system's viability, its continuity and its ability to change (Buckley, 1967, p52).

2.4 The Development of Collective Efficacy Models

The increasing analytical focus on neighbourhood effects through the 1980s, highlighted by Bursik and Grasmick (1995), had numerous triggers, from the policy relevance of the 'broken windows' theory of policing by Wilson and Kelling (1982) to Blau and Blau's (1982) work on inequality and violence, as well as a broader emergence into the mainstream of the social capital concept (Putnam, 2000; Halpern, 2005). Motivated by this upturn in papers looking at neighbourhood effects in general (not just in relation to crime) Sampson, Morenoff and Gannon-Rowley (2002) set out to review the 'neighbourhood effect literature' related to problem behaviours and health related outcomes. They identify, review and assess over 40 relevant studies, evaluating the salience of social-interaction and institutional mechanisms to account for neighbourhood level variation over a variety of phenomena such as delinquency, violence, depression and high-risk behaviour. They find the strongest evidence for neighbourhood processes when looking at crime-related outcomes, and also unpick a number of methodological issues. They consider one of the most important 'first order' findings to be that survey response items can yield reliable and valid measures of neighbourhood social ties, and call for the development of a set of ecological assessments of social environments.

Although they found very little consistency in how neighbourhood mechanisms are measured, they highlight four related but independent neighbourhood mechanisms:

- Social ties/interactions – seen as related to social capital or "resource that is realized through social relationships" (p457), which shall be discussed below.
- Norms and collective efficacy – meaning the willingness of residents to intervene dependent on conditions of mutual trust and shared expectations.
- Institutional resources – referring to the quantity, quality and diversity of institutions (especially non-profit and civic-based organisations)
- Routine activities – land use patterns and the ecological distributions of daily routine activities.

Using these findings Sampson (2006) develops a collective efficacy theory, seeing it as a "critical ingredient for understanding neighbourhood crime and more general aspects of community wellbeing" (p39). 'Collective efficacy' is fundamentally about the activation of social ties based on repeated interaction, and therefore expectations about the future which generate norms outside of the kinship/friendship realm. It has two distinct components – shared expectation of social control and social cohesion/trust (Sampson 1999, Odgers et al., 2007). Importantly, Sampson et al. (2002) find that these two dimensions can be tapped through survey response questions: for the first dimension questions such as "how likely would your neighbours be to take action if a) children were skipping school, b) a fight broke out, c) they found a lost wallet" and the second dimension, cohesion and trust, questions such as "people are willing to intervene for their neighbours" or "this is a close knit neighbourhood" can yield reliable results.

Collective efficacy should be seen as a mediating effect between structural backgrounds, the number of organisations, the routine activities of residents and spatial dynamics of land use: collective efficacy is a higher order mechanism than these other neighbourhood characteristics and as such is not confined to a predefined spatial setting (although in practice these will often be relevant). For example, neighbourhood deprivation and neighbourhood collective efficacy do not necessarily go hand in hand and research has shown that collective efficacy mediates the relationship between these two for a range of outcomes, not just in criminology (Cohen et al., 2006; Cohen et al., 2003; Sampson et al., 1997; Xue et al., 2005). The relationship between this concept and social disorganisation/social ecology is clear to see although in many ways it is the opposite of its predecessor, in that it involves the presence, rather than absence, of informal social control (Pratt and Cullen, 2005; Taylor, 1997).

Collective efficacy has shown robust associations with antisocial outcomes among adults, including violent crime (Sampson et al., 1997), partner violence (Browning, 2002), and homicide (Morenoff et al., 2001). Bradford and Myhill (2015) use the Crime

Survey for England and Wales panel experiment[4] to investigate trust in the police and the criminal justice system (CJS) more broadly and observed that perceived increases in collective efficacy over the span of the data available to them (from the first quarter of 2010 to the first quarter of 2011) was associated with an increase in confidence both in the police and the CJS. Sutherland, Brunton-Smith and Jackson (2013) found that collective efficacy was negatively related to police recorded violence but did not mediate against the structural characteristics of the neighbourhood, against their expectation based on other work. They attribute this to the specific context of London as a city, in part due to data lags and also to possible missing domains in their investigation such as the role of social and political institutions or the quality of police-public interactions. A later piece by the same trio of authors (Brunton-Smith, Jackson and Sutherland, 2014) using the same London-based METPAS data (Jackson et al., 2012) identified that collective efficacy does, however, partially mediate many of the statistical effects of structural characteristics of the neighbourhood on *beliefs* and *worries* about violent crime.

Odgers et al. (2009) found that collective efficacy was negatively associated with levels of antisocial behaviour at school entry but only in deprived neighbourhoods and this relationship held after controlling for neighbourhood problems and family-level factors. They call for the work to be repeated for adolescents, as they hypothesize that it is during this period that "children move through their primary school years and begin to have more direct exposure to members of the community and institutions, neighbourhood-level effects are likely to be transmitted both directly and indirectly, through a series of complex and age-dependent pathways" (p954). This echoed the work of Ingoldsby and Shaw (2002), which suggested that middle childhood may represent a critical developmental period during which children are at heightened risk of neighbourhood-based effects on antisocial behaviour problems.

Pratt and Cullen's (2005) meta-analysis of macro-level (community) predictors and theories of crime identified over 200 empirical studies conducted between 1960 and

[4] The CESW panel experiment was conducted in the spring/summer of 2011 and re-contacted around 1,500 respondents to the 2009/10 and 2010/11 waves of the original survey.

1999 and found social disorganisation to be amongst the theories that had received the strongest empirical support. As a headline message this is surprising, considering the criticisms this theory received during the 1980s; Pratt and Cullen explain this by noting that they allowed an area's racial heterogeneity and rates of family disruption to correspond to social disorganisation and these in part drove the finding. The addition of the few 'collective efficacy theory' studies to social disorganisation significantly (statistically) strengthened this support. When weighting for the small sample size (just 13 papers) and analysing the effect of each concept independently, the authors rank collective efficacy as the fourth most relevant predictor domain for variations in crime rates, ahead of traditional suspects such as family disruption, inequality, racial heterogeneity, age effects (which came a surprising 16th) and poverty levels (which placed eighth).

2.5 A Note on Social Capital

Putnam (2000) defines social capital as "connections among individuals – social networks and the norms of reciprocity and trustworthiness that arise from them" (p19) and primarily in terms of personal ties and group memberships. Crime related examples are littered through pretty much any discussion on social capital; as Putnam (2000) puts it, "if the crime rate in my neighbourhood is lowered by neighbours keeping an eye on one another's homes, I benefit even if I personally spend most of my time on the road and never even nod to another resident" (p20) and it's the rise in the rate of publications with 'social capital' and 'neighbourhood' that gives Sampson et al. (2002) their time frame for which to retrieve the more than 40 studies that they review.

Halpern (2005) claims all work defined under social disorganisation/social ecology/collective efficacy as a subset of social capital; "it is not my intention to encourage criminologists to give up their current terminology but it is important to recognise that the social phenomena that they explore are the same as those examined by economists, health professionals, educationalists and so on, namely social capital" (p140). This annexation of collective efficacy into social capital does criminology a disservice. Social capital as a concept is too broad and blunt to allow the nuanced understanding necessary for the study of delinquency, with Halpern's

conceptualisation often reaching as far as labelling and anomie theory. Sampson (2006) clarifies:

> Collective efficacy therefore helps to elevate the "agentic" aspect of social life over the perspective centred mainly on the accumulation of stocks of social resources as found in ties and membership (i.e., social capital). This conceptual orientation is consistent with the redefinition by Portes and Senenbrenner (1998) [sic, it should be 1993] of social capital in terms of "expectations for action within a collective". Distinguishing between the resource potential represented by personal ties, on the one hand, and shared expectations for action among neighbours represented by collective efficacy on the other, therefore helps clarify the dense networks paradox: *social networks foster the conditions under which collective efficacy may flourish, but they are not sufficient for the exercise of control* (p39, original emphasis).

That is not to say that work on social capital has nothing to offer; its high profile has certainly added to the attention paid to collective efficacy (Sampson, 2006) and there are many methodological crossovers (Halpern, 2005) which can aid and inform its criminological counterpart. One area where the lexicon of social capital can be particularly useful is where social capital 'goes bad', the formation of the (criminal) gang, but there are strong enough distinctions between the two to mean that the separation should remain.

2.6 Situational Action Theory

Wikström's Situational Action Theory (SAT) is best elucidated in his and his team's 2012 publication *Breaking Rules: the Social and Situational Dynamics of Young People Urban Crime* (Wikström et al., 2012) that claims to be "one of the most comprehensive studies of young people, their crimes, and its causes ever conducted" (from the jacket cover on the 2012 hardback) and in some respects this is a difficult claim to refute, due in no small part to the extent of the data collected. Set in the "modestly sized English city" of Peterborough, as much for its proximity to the Cambridge-based research headquarters as for its parsimonious size and diversity of physical and social characteristics, this is the first book from an expected series of publications drawing on a quite astonishing array of unique data.

Informed by a theoretical framework that attempts to integrate the "previously fragmented discipline of criminology" into an adequate theory of crime causation (more on this later) the data comprise an annual longitudinal randomly selected representative sample of around 700 children aged 11 to 12 in 2003 over 5 years to 2008, with a remarkable retention rate of 97%. The respondents were administered a detailed questionnaire asking a range of questions about their lifestyle, family, school, neighbourhood, work, personal issues such as motivations, beliefs and decision making processes (such as violent behaviour vignettes), substance use and self-reported offending across a range of different offence types. To supplement this already extensive data, a range of other sources were drawn together. Police National Computer (PNC) records were retrieved for the respondents and an innovative 'space-time budget' diary was recorded over four days (two weekdays, a Friday and a Saturday) sometime in the first quarter of each year of the study, which is assumed to be representative of the entire year (within the confines of practical achievability). This captured the traditional items of an hourly time budget diary as well as recording the location at which the events mentioned took place. Further, in 2005, the team conducted a small area community survey, differentiating at the unprecedentedly precise level of Output Area (roughly 300 residents). With over 6,615 respondents to this postal questionnaire (averaging 13 per output area) this component captured features of the immediate neighbour characteristics of the whole city such as the localities' levels of trust/cohesion, problems with youth and neighbours, issues around local services as well as a range of variables on the demographics, opinions and lifestyles of the area survey respondents, who, crucially, were independent of the 700 core longitudinal respondents. Information on the physical environment was further enriched with official data surrounding residential instability, concentrated disadvantage, ethnic diversity (or population heterogeneity), non-residential land use and the number of lone parent families. The totality of this data represents just phase 1 (adolescence) of the Peterborough Adolescent and Young Adult Development Study (PADS+), with phase 2 (not reported in the book) extending this data in young adulthood.

The goal of SAT seems to be to draw together the lessons from the tradition of neighbourhood analysis that started with the early Chicago School and progressed

through to the work of Sampson and colleagues with the developmental and life course theories that have allowed the range of risk factors that are now well known to be understood. Wikström et al. (2012) write that SAT "aims to incorporate and develop key insights from routine activity theory/crime pattern theory [Cohen and Felson, 1979 and Brantingham, 1993, respectively] with social disorganisation/collective efficacy theory within an emergence-selection framework" (p41). The observation at the base of this theory is that crime is ultimately an interaction between a person's criminogenic propensity and criminogenic exposure at a location, and given the right 'perception-choice' processes crime happens within a combination of these two elements. If a person with zero criminogenic propensity is exposed to any number of settings the outcome shall always be the same, no crime. Whereas a person more inclined to engage in criminal behaviour will do so only when the conditions for that action are perceived and then chosen.

In trying to set these basic building bricks into a system that unites the disjointed fields of criminogenic theorising and research, Wikström tries to define some base units from which to work. The first is a definition of crime, namely through morals rules: "a moral rule is a rule of conduct that states what is the right or wrong thing to do… The law is a set of moral rules of conduct. Acts of crime are acts that breach moral rules" (p12). This definition based on morals has caused some lively debate, with Felson (2014) citing attitude–behaviour research that has "repeatedly found that moral attitudes do not simply produce moral behaviour" (p254), questioning the analytical time frames and concluding that "without the morality baggage, the Peterborough descriptions fit well with prior routine activity findings" (p257). The response by Wikström et al. (2015) states that "Felson significantly misrepresents the scope and content of Breaking Rules and makes misleading claims" (p115) and fundamentally disagrees with the subsumption of SAT into Routine Activity Theory. If the moral rule definition is accepted, as Wikström claims is possible, the theory claims one need specify what moves people to engage in this reading of crime, to specific personal and environmental factors that affect crime engagement, and to look at the social condition and individual developments that brought people to this position. This latter element, the emergence of the social condition and the personal histories of the individual, are defined as

outside the scope of the 2012 work and shall be tackled in further publications (as noted on page 50 of the book). This leaves the work assuming, for the presented publication at least, that there exist different kinds of people with different criminal propensities and disparate types of areas with different criminogenic characteristics: to analysis of the "context of action" rather than the "context of development" (p35).

How these criminogenic (individual) propensities and (area) characteristics are defined clearly echo the goal of joining the two schools. A young person's individual criminogenic propensity is based on their morality and their self-control and hence follows a general theory of offending paradigm, whilst the criminogenic characteristic of the area depends on the moral context of the setting (not locations, as the same physical location can have different contexts depending, for instance, on the time of day or the presence of a responsible adult). Sampson et al.'s (1997) measure of collective efficacy from the small area community survey is used as a start point to judge the moral context of a location, with the 25% of output areas with the lowest scores defined as having poor collective efficacy. The (moral) context in which someone finds themselves in said location is added to the definition, with special consideration of the city and local centres (mainly as these are areas with large non-resident populations and are hence less likely to be cohesive and trusting), by restricting the definition to time spent in unstructured peer-oriented activities (i.e. activities with peers outside of work or school that are unsupervised by adults) in the city or local centres, or areas with poor collective efficacy.

The substantive findings of the work reflect the focus of the analysis on the 'context of action' and are mainly at the area level (such as that the effect of social disadvantage of an area on the crime count is fully mediated by collective efficacy or that social disadvantage, ethnic diversity and residential instability lead to worse collective efficacy) or the offence location (such as that young people rarely offended in the strong moral context that characterise family, school and work-related activities, or in those situations with strong enforcement - typically an adult supervisor), but there are also some that fall closer to the developmental side. For example, the work finds that no respondent living in the highest collective efficacy areas had a criminal record whereas

15% of those who lived in the poorest areas did (and the same pattern held true for self-report), that those with low crime propensities are situationally resistant to acts of crime and for those who did enter into criminal acts, provocation had a greater effect than the monitoring of the situation (at least in relationship to violent acts). The next publication in the series will "investigate the factors that influence the emergence of young people's different crime propensities, the role of changes in young people's crime propensities and the criminogenic exposure for their crime involvement and criminal careers, and the factors that influence such changes" (Wikström et al., 2012, p50).

2.7 Summary

There are two distinct seams of investigation that can be traced through the criminological literature. One is the analysis of the individual, their behaviours and their risk factors, the other the characteristics of the area and the behaviours of people within them. Through a symbiotic relationship between the empirical and the theorising under the banner of Developmental and Life Course Theory a great deal of knowledge about risk factors at the individual and family level have been uncovered, though for reason of data collection limitations and focus this field has delivered less in the way of concrete findings of the impact that neighbourhood processes play on offending. On the other hand there is a strong evidence base that deprivation, social disorganisation, disorder and/or collective efficacy have an effect on the amount of offending in a specific area. Consequently, this has led to two closely related but fundamentally distinct sets of knowledge operating at different levels; one focused on the individual, the other makes inferences and statements at a more macro-level.

The reason for this separation was recognized as far back as the 1930s (Merton, 1938) and the two streams of work have progressed to talk to different policy contexts (Messner, 1988) but the separation has been one of convenience only. The theoretical debates that have underpinned the developmental and life course criminology that led to a good deal of the risk factor type understanding have always allowed some level of environmental components to be included. Often this has been of the form best expressed by Wikström's Situational Action Theory, that there is an interaction between a 'bad person' and a 'bad place' but there is also clear reference to the characteristics of

the developmental context of the individual and how that relates to the development of any delinquent behaviours. There have been some crossovers between the two schools, notably Odgers et al. (2009), which calls explicitly for more research at a slightly older age group than in their study (aged 5 to 10), but it remains the case that the higher level risk factors, especially with concern to area level, are less well understood and cannot always be separated out from the individuals that live within the area: "it is difficult to determine to what extent the areas themselves influence antisocial behaviour and to what extent it is merely the case that antisocial people tend to live in deprived areas" (Welsh and Farrington, 2012, p60).

The data and methods from which risk-factor findings have been derived hold some culpability for the lack of clarity about the role that area character plays with regards to delinquent behaviour. The majority of the classic studies as highlighted in Table 2.1 have been constrained to either specific cohorts or small areas, or at the very best a specific city. Sampson (2012) and Wikström (2012) have gone to extraordinary lengths to be able to collect deep data on the specifics of Chicago and Peterborough respectively, and due to the wealth of data collected have been able to analyse a myriad of different research questions but have not yet connected exactly with the subject of this investigation. There is space therefore to address this gap, to try to use existing, standardised data with appropriate statistical techniques grounded in the theories that have underpinned both of these schools of thought to examine the relationship between how the area in which someone lives and its characteristics shape delinquent behavioural development.

Chapter 3

Research Specification

3.1 Overview

The previous chapter demonstrated that a great deal is known about risk factors that correlate with delinquent behaviour. On the individual level these constructs, such as low self-control, impulsivity and low empathy, have been well researched. Similarly, family level correlates have been robustly demonstrated to include large family size, criminal families, poor parental supervision, childhood physical punishment, low parental reinforcement, erratic discipline and inconsistent rewards, low parental warmth, child abuse and neglect, and parental conflict. This knowledge has typically been uncovered by small scale or geographically constrained investigations, and the data collection processes are time consuming and expensive. Less is known about 'higher' level correlates such as school, peer and neighbourhood effects. Alongside this risk factor developmental approach, a second strand of criminological work has been the long history of investigation of neighbourhood effects. For reasons of data collection limitations and policy focus the goal of many of these works has been on uncovering an understanding of processes at the neighbourhood level and less attention has been paid to how the area in which a person lives affects the individual's offending behaviour. There has been some significant attention paid to victimisation and fear of crime, but less to whether these area effects correlate with offending.

This work therefore has two main goals. The first is to investigate if these two schools of criminology can be joined together, whether light can be shed on the how the area in which a person lives acts as a risk factor for offending behaviour. The aim of this thesis is not to test one of the many theories that have been developed in the Developmental and Life Course theory of offending paradigm; rather, it is to draw on what has been common across many of these theories and to try to link it with the theorising from the neighbourhood side. There have been some tentative steps made in this direction but

more work needs to be done. This is one of the main conclusions of the Odgers et al. (2009) paper and is likely to form a significant component of the anticipated follow-up to the Wikström et al. 2012 book. The second aim is to test whether it is possible to investigate this question using existing and publically available datasets. Secondary data analysis has a number of strengths, as will be discussed in Chapter 5, and has been the subject of an Economics and Social Research Council funded initiative to boost its usage. If it is possible to extract results from an existing and pre-collected generic dataset and find substantive findings that address one of the most abiding and thorny issues in the complex theoretically rich domain of offending behaviour, then the implication could be rather wide-ranging, both within and beyond criminological research.

These two aims lead to a set of three sequential research goals that should be applicable as an analytical framework to any number of quantitative investigations and can be bluntly stated as: 1) What can be measured? 2) Can these measures be structured? 3) What are the results? To flesh these out a little and make them more specific to the current work these can be rephrased as:

1. Can criminogenic risk factors be measured from an existing dataset to include measures of both the well-established risk factors and those relating to area?
2. How should these measured factors be linked together to account both for theoretical dependencies and the limitations of the measurement?
3. What can this tell us about offending risk factors? This last question is very much the crux of the investigation as it connects the methodological questions to the substantive subject.

Before moving to express these goals in even greater granularity, so as to give a clear outline of the exact nature of the remainder of the thesis, some strict parameters need to be drawn as to what this work will and will not include. Firstly, the focus of the entire thesis will be on late childhood and early adolescence, young people aged between 10 and 16 years of age. This group are highlighted by Odgers et al. (2009) as in need of a greater understanding and it is therefore this age range that shall be concentrated on. The second sharp edge is that this work will only consider one set of higher level

factors. It has been highlighted that higher level influence on offending can include area, school and peer effects but to attempt to answer all three questions in one would be too large a task, and would cloud understanding. The final parameter to put in place is that the work will focus solely on offending behaviour rather than the more general deviant acts that can be classified as anti-social behaviours. Anti-social behaviour in a UK context was defined in the 1998 Crime and Disorder Act as acting in a "manner that caused or was likely to cause harassment, alarm or distress to one or more persons not of the same household" (section 1a) and as such is a relatively new distinction. DLC theories typically define types of delinquent behaviour in the process of defining what it is that affects their perpetration. The focus here shall be on what is classically understood to be offending, i.e., committing an illegal act, to side-step the need for such a debate.

3.2 Specific Research Questions

The following sets of questions try to unpack even further the three headline goals of this work into something like a set of answerable sub-questions. The questions are presented here separately as methodological and substantive, and to some extent the work will split between these twin goals but there will be a natural interaction between the two strands of investigation. The questions will not be answered separately but will be braided together and are presented as distinct to delineate the work more thoroughly.

3.2.1 Methodological Questions

Methods Question 1 - can a suitable range of risk factors be accurately measured from an existing data source?

This question can be quickly broken down into two subparts: are there any suitable data that are available and how can relevant components be measured from them? Before embarking upon any analysis it is vital that the data to be used are checked for suitability. This section will start with a thorough review of possible data sources before an in-depth consideration of the Home Office's Offending, Crime and Justice Survey (OCJS). This is the first and to date only national longitudinal, self-report offending

survey for England and Wales. The structure and limitations of the dataset will be investigated, with a detailed discussion of the reliability of the methodology in the questionnaire used to capture offending behaviour.

The second component necessary to answer this question is to consider the domains that have previously been seen to be relevant to offending behaviour and assess both their theoretical importance and their possible measurement. In answering this part of the question it will be necessary to measure constructs at the individual, family and 'higher' levels and for each level there will be a discussion as to what can be achieved from the existing data. Directly related to this question is the method that shall be applied to the work. Measurement is a tricky business and without proper attention to detail misleading results are highly probable (Hand, 2004) and the methodology used can play an important role in this. It seems relatively clear that standard regression techniques that are usually applied to survey data will not be adequate to capture the nuances of the measurement needed. As many of the risk factors in previous work (as will be discussed in Chapter 6) are based on a range of questions that need to be combined, a natural choice for the methodology is structural equation modelling, as it allows detailed measurement via factor analytical approaches.

Methods Question 2 - how should these factors be modelled to gain an understanding of the risk factors associated with offending?

Once data has been found to work with, and the measurement of the relevant factors arrived at, the next step is to create models that accurately reflect the structure of how these constructs interact, paying attention to both theoretical dependencies and measurement limitations. This means that the causal processes between model components will need to be examined and a structure defined that holds together from both the theoretical and analytical point of view. As with the measurement step in methods question 1, structural equation modelling will be shown to be an appropriate methodology as it allows detailed structural implications to be incorporated into the work.

Methods Question 3 - what are the limitations of both the data and the model?

The second methodological question will define a structure in its very simplest form with a focus on trying to pin down how the constructs found and defined in the first question should be linked together. Once this has been finalised it will be possible to add more parts to the model and assess how they interact with offending behaviour. This will take a number of forms dependent on what is available in the dataset but should include analysis of different sub-types of offending, adding controls to the model and trying to incorporate some element of the longitudinal aspect of the OCJS. The focus from a methodological point of view, rather than the substantive, is on how well these models function, where they start to break down and why.

3.2.2 Substantive Questions

The substantive focus of the work will match to the methodological findings an interpretation that makes sense criminologically. Without this interpretation the methodological questions are meaningless correlations of numbers in a dataset and the substantive, theoretical relevance of the findings will be constantly woven into the discussion to give meaning.

Substantive Question 1 - How do the measurable risk factors relate to previous definitions?

This is very closely associated with the first methods question, although it means interpreting the measurements made in terms of previous operationalisations that have been used in DLC work. It some respects this could be could be considered a 'pre-substantive' question as it does not generate a hypothesis. It is, however, a vital step as it will not only mean that the measurement models are representing something useful but will also go some way to separating out the nuances of the achieved measures and feed into the structural methodological work.

Substantive Question 2 - Which risk factors affect offending?

This is clearly the absolute central part of the thesis. There are some clear expectations that these effect sizes should follow. Criminogenic propensity, self-control, empathy or impulsivity (the individual's characteristics), has been the bedrock of developmental criminology and in all theories reviewed in chapter 2 have been a central concern. It

must follow, therefore, that this shall have the greatest bearing on offending behaviour. It is not clear whether family level components should have the next largest effect, as although they have been clearly found in the literature and seen to be of high importance there has been relatively little comparison to the size of area level effects, and there have often been problems of suspected selection biases clouding understanding. That said, on balance, if one set were to be chosen it has to be expected that family level size effects will be larger than area if only because these have previously been seen quite clearly, whereas area effects have remained obscured. A key concern in answering this question will be in the operationalised differences that the risk factors take, dependent on how they are measured. This will need to be worked through the narrative.

Substantive Question 3 - How do the effects of these risk factors vary by the type of offending? And how do they vary over time?

Linked to the final methodological research question, and assuming that suitable models can be formulated, this question asks for the model to extend to try to gain knowledge about wider questions. This could fall under the previous question as it will still be looking at risk factors and offending, but this is meant to capture the need for interpretation of results beyond the initial modelling phase.

3.3 Summary

The questions here are therefore along two strands of work, the methodological and the substantive, though the dichotomy is rather false. Both of these strands of work will be interwoven and findings will complement each other throughout. Indeed, many of them will only be possible to approach when considered in partnership. Creating measurement and structure from a methodological point of view will only be possible when firmly grounded in theoretical debate and related to criminological work that has gone before. Similarly, any substantive findings will be constrained and only applicable in the limited sense that the methodological limitations allow. These constraints will be discussed clearly and findings will be phrased in their terms.

Chapter 4

An Introduction to the Offending, Crime and Justice Survey

4.1 Possible Data Sources

As a piece of quantitative data analysis focused on trying to pull together previously disparate parts of criminological theory, including more modern developments around area level effects (most notably collective efficacy theory), and with a secondary goal of providing a framework to get maximum analytical value from existing data sources, the first step of an investigation is to consider from both a theoretical and practical perspective which existing datasets could be used to answer the research questions. Dale, Arber and Proctor (1988) list six questions that need to be asked of a dataset before it can be used:

1. What was the purpose of the original study, and how does it match the research questions?
2. What information was collected and is it relevant?
3. How was the sample drawn and are the subjects relevant for the research?
4. Who collected the data and were their procedures adequate?
5. Which population does it represent?
6. When the data were collected and are they still relevant?

First and foremost, in that the goal of the analysis is to investigate the effect on offending, an ideal dataset for this investigation would have been designed *specifically* to investigate criminal behaviour, rather than merely have had a range of offending questions added at some point in the study. Capturing people's offending behaviour is a difficult and sensitive area (see section 4.6) and hence a design to specially capture offending as accurately as possible is needed.

The data collected must have information incorporating both offenders and non-offenders to allow comparisons to be made, preferably with a wide range of measures

of offending behaviour to allow varied analysis (although one may expect this if indeed the survey were designed with offending research as a primary goal and followed current best practice). To be relevant to the investigation goals, key components of the research question need to be matched as closely as possible by key elements contained in the dataset: appropriate geographical/community level identifiers and survey response items relating to Sampson's (2002) collective efficacy theory must be present, along with variables about the lifestyles and positions of the respondents, such as their attitudes, socioeconomic status, details of their family background and situation, their ethnicity, gender, age, and perhaps information about their schooling/academic attainment. These are to allow controls for Developmental and Life Course Theory (DLC) risk factor-type components.

The sample would ideally be contemporary, collected at least within the last 10 years, so as to give results and policy implications relevant to the current time period and to take account of any recent survey methodology improvements that have increased the reliability of recording practices (see, for example, the self-report offending questions as discussed later). It should be representative both of possible respondents and any geographical area that it professes to cover and must have been drawn with a non-biased sampling frame to allow generalisability. To be most relevant to the research gaps identified in the background and research specification, it should ideally range in age from children/young adolescents to young adults as a minimum and at least cover the ages 10 to 16 (Odgers et al., 2009; Nagin et al., 1995).

Clearly this is a large undertaking and such a procedure could only be carried out by a professional and well-funded organisation with experience in large scale data collection and questionnaire design, with appropriate quality assurance controls, sampling frames and interviewer availability. Ideally, the data would be longitudinal, allowing within-person effects to be understood and causal natures to be more clearly unpicked (which is especially pertinent for trying to access any form of area effects, see Sampson, 2002).

The criterion that the dataset must hold reliable offending behaviour measures rules out a good number of the 'traditional UK data sources'. The 1970 British Cohort Study has limited self-report contact with the police and offending behaviour questioning in

the 1999/2000 sweep only. These consisted of the number of times the respondent reported being moved on by police, questioned, warned, taken to a police station, cautioned or found guilty by a court. The respondent was asked to cast their mind back to 1986, reporting any contact with the CJS since then; as such, it captures historical information from age 16 through to 30, which will almost certainly be subject to recall problems (again, see discussion in section 4.6). Similarly, the National Child Development Study (NCDS) does not have the correct focus, with questions on offending at the 42-year sweep, asking only of the 10 years previous to that point. Whether the respondent had been arrested between any two sweeps is available (but with incomplete coverage) but would only capture criminal behaviour from 30 years old, and hence miss the adolescent phase. The NCDS also has parental reports of trouble with the police at age 16, but with a high level of missing data (around a third, see Hobcraft, 1998) and these data are old, relating to adolescence in the mid-1960s. The Longitudinal Study of Young People (LSYPE) in England covers 'risk behaviours' but does not have the level of questioning needed for a detailed understanding of offending.

The Avon Longitudinal Study of Parents and Children (ALSPAC) similarly does not have the required level of offending questioning granularity, its goal primarily being to investigate health outcomes. The ALSPAC also suffers from being located in just one area; this would not necessarily rule it out as a data source, but the lack of offending detail does. There are two UK-based data sources that could both provide a solid basis for the investigation, although both are constrained to one city context. The Edinburgh Study of Youth Transitions and Crime (ESTYC) is a large scale, longitudinal cohort study of adolescents who started secondary school in Edinburgh in the autumn of 1998. It integrates a number of diverse data sources from young people interviews, teacher assessments, parental interviews, school records, social geography, a resident survey and police-recorded crime data. However, only the first four waves covering from age 9 to 14 (1997 to 2001) are currently available to researchers outside the core team. Similarly, the impressive array of data collected and arranged in the Peterborough Adolescent and Young Adult Development Study (PADS+) is still in development and

is only available to internal researchers at Cambridge University, with a significant publication schedule ahead of it (Wikström et al., 2012).

Consideration is therefore restricted specifically to offending related datasets that are freely available and, with a focus on analysis of risk factors at an area level, very few sources are suitable and are for the main quickly ruled out. In the Arrestee Survey, for instance, each member was the subject of arrest and therefore the counterfactual (i.e., non-offenders) are not present. The Youth Lifestyle Survey's focus was primarily to investigate offending amongst young people and how this differs by lifestyle and demographic factors, but this is cross-sectional with only two sweeps (in 1992/93 and 1998/99) and hence fails the longitudinal and the contemporaneous criteria. The Crime Survey for England and Wales (CSEW, formerly the British Crime Survey) could provide an ideal basis with its large sample size, strong methodology, its range of questions on the respondents' life circumstances and geographical quantifiers. Its primary focus, however, is on victimisation, avoiding the methodological issues of self-report offending, and an offending measure is not present.

4.2 Introduction to the Offending, Crime and Justice Survey

The Home Office recognised the significant gap in data on offending in the general population (as opposed to particular groups such as convicted offenders) and between 2003 and 2006 commissioned the Offending, Crime and Justice Survey in response. This four-year rotating panel study was the first national longitudinal, self-report offending survey for England and Wales, and was conducted by the National Centre for Social Research (NatCen) and the British Market Research Bureau Limited (BMRB)[5]. The stated aim of the study was to provide a solid base for measuring prevalence of offending and drug use in the general population of adolescents (aged 10 to 25) in England and Wales and the survey holds measures of self-reported offending, indicators of repeat offending, drug and alcohol use and questions on the links between them, a range of lifestyle, health, socioeconomic, neighbourhood and attitudinal

[5] NatCen were the primary contract holders and the primary data collection and processing agency.

questions, along with information on the nature of offences committed, such as the role of co-offenders and the relationship between perpetrators and victims. Its original purposes cover this work's research goals well, the information is relevant and contemporary, and it represents the correct population.

Although there have been relatively few publications based on the OCJS, most of which have been produced by the Home Office, it has shown that detailed statistical analysis of offending can be conducted using it. Three publications are very similar in style and report on the extent and trends of offending, anti-social behaviour, drug/alcohol use, contact with the CJS and victimisation of respondents to the survey (Budd et al., 2005; Wilson et al., 2006; Roe and Ashe, 2008). These should be considered as regular Home Office outputs containing standardised information on the datasets. Three more specific reports include a detailed longitudinal analysis (Hales et al., 2009), a paper on delinquent youth groups (Sharp et al., 2006) and a piece on minority ethnic groups and crime (Sharp and Budd, 2005). The first of these three found that measurement and identification of risk factors is important to target interventions at those most likely to go on to offend, with the analysis highlighting family, peer group and school factors as important influences on the behavioural trajectories of young people during their teenage years. The delinquent youth group paper found that membership of a 'gang' increases the risk of offending when controlling for other factors such as temperament and other lifestyle risk factors.

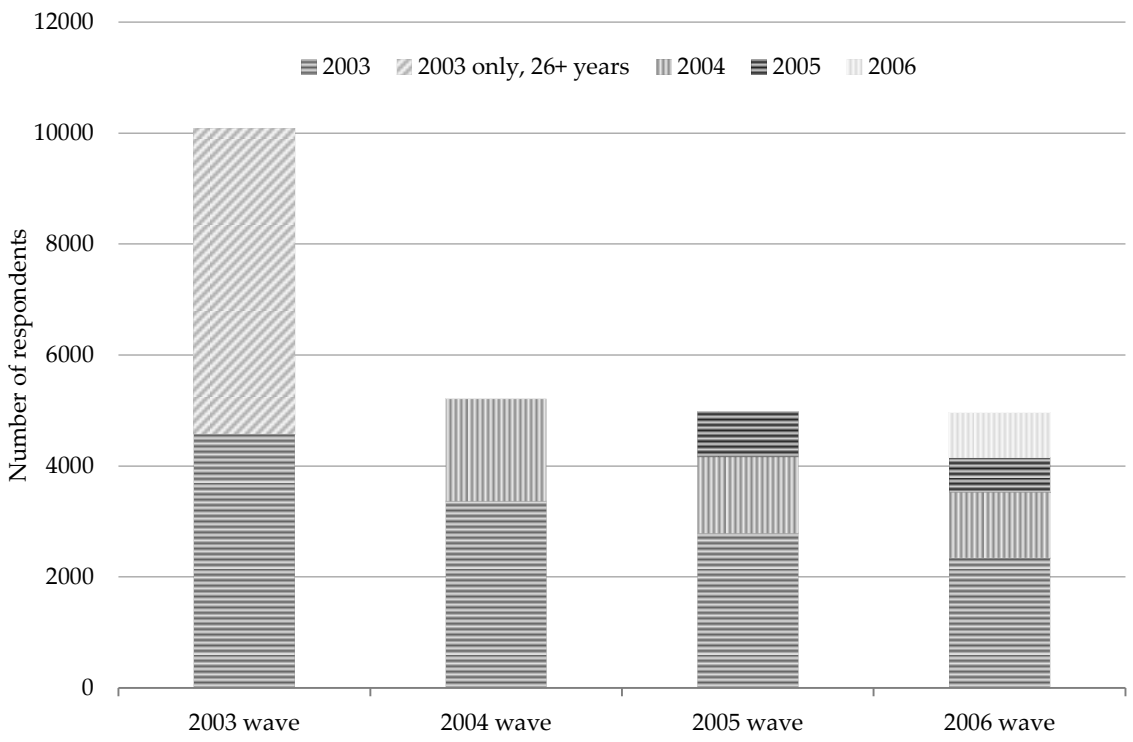
Each wave of the survey was designed to be nationally representative and used postcode address files as the sampling frame. Despite consideration of problems with the coverage of this sampling frame, sampling error and some non-response bias, it was judged to be broadly representative of the national population that it attempted to capture. The questionnaire contained the following modules (brackets show method of collection) with most questions being asked of all respondents, including children:

- Household grid (CAPI)
- Socio demographic (CAPI)
- Neighbourhood (CAPI)
- Attitudes to Criminal Justice System (CAPI)

- Contact with Criminal Justice System (part 1) (CAPI)
- Victimisation (CAPI)
- Anti-social behaviour (audio-CASI)
- White collar/hi tech crime (audio-CASI)
- Offending – count/follow-up (audio-CASI)
- Offending – nature (CASI)
- Contact with CJS (part 2) (CASI)
- Domestic violence (CASI)
- Drinking (CASI)
- Drug use (CASI)
- Health, lifestyle and risk factors (CASI)
- Reactions to the survey and re-contact (CASI)

The first wave covered around 12,000 people aged between 10 and 65 living in private houses in England and Wales and as such did not include people living in institutions such as prisons, the armed forces, or the homeless and as such it is acknowledged that it misses some of the highest rate offenders. Subsequent waves followed up with respondents aged 10 to 25 years with around 5,000 respondents per wave, with just over 2,500 in all waves. The reduction from all ages in the first wave to 10- to 25-year-olds afterwards was to focus resources and attention on those most at risk of offending (OCJS technical guide, 2004). This fits well with the aims of this research. Figure 4.1 gives a graphical representation of the structure of the achieved samples.

Figure 4.1 OCJS achieved sample, by year of entry to study



Source: Cross-sectional OCJS, waves 2003 to 2006

4.3 Nothing is Perfect: OCJS Limitations

Overall, then, the OCJS gives a good source upon which to build an investigation and although it passes most of the criteria listed by Dale et al. (1998) in relation to the goals of this study, there are some aspects that could be wished for that are not present. The most notable problem, and indeed an issue that can be levelled at the majority of nationally administered surveys, is that there is no triangulation of response. Put another way, each piece of information is collected only from the respondent; there are no measures from other sources such as teachers, other residents or the police. Hirschi and Gottfredson (1993) argue that behavioural studies of self-control, rather than attitudinal measures, are more indicative of an individual's true level and explain a greater amount of the variance in delinquency and crime. By contrast, Arneklev et al. (2006) claim quite the reverse, that "the attitudinal indicator of low self-control is a relatively stronger predictor of crime than imprudent behaviour" (p41); taking the respondents' view *only* gives the possible presence of measurement bias. Similarly,

although at first reading it seems that Sampson et al. (2002) believe that "the important point is that neighbourhood processes can and should be treated as ecological or collective phenomena rather than as individual-level perceptions or traits" (p456) a more nuanced understanding of the work surrounding their definition of 'ecometrics' is that every effort should be made to approach the neighbourhood context, size, shape and definition from as many perspectives as possible, depending on availability, what is to be researched and the focus of the analysis: a central point of Sampson's 2012 round-up of the Project of Human Development in Chicago Neighbourhoods (PHDCN) and related work, *Great American City*. This criticism of a lack of 'second opinion' in survey data is not rare and the implications/solutions will be discussed in detail later in this thesis. It is more pertinent here to focus on some more practical issues with the OCJS data.

For example, although in the 2004 OCJS User Guide the intention is made clear that "in order to ensure comparability between survey years, much of the questionnaire remained the same as in 2003" (p13), in reality there was a significant amount of variation. In many cases this took the form of additional questions being added to a pre-existing battery: extra elements around neighbourhood collective efficacy were added, self-perception grew new parts (*I like taking risks*) and questions related to family situation (such as *My parent(s) make sure I do my homework* and *My parent(s) expect me to be in at a certain time on Friday or Saturday nights*) appeared. In other cases the coding and ordering of questions was changed (for example, multiple choice lists' suffixes had been misapplied, and this was corrected into the 2004 survey). Overall, of the 1,803 questions in the 2003 questionnaire, only 1,361 remained unchanged in the 2004 sweep, with the 2004 dataset swelling to include 2,369 questions. Fortunately, the design had settled down by 2004 and from 2004 to 2006 the questionnaire design is much more stable, with around 95% of the questions that appear in 2004 also appearing, unchanged, in subsequent waves.

The sample design also changed between 2003 and later editions of the survey. Most strikingly, the target population was reduced from 10- to 65-year-olds to 10- to 25-year-

olds[6]. This is clearly shown in Figure 4.1, where an additional 5,505 people aged over 25 are represented by the top section of the 2003 wave bar. This in itself is not a problem as the key age group of interest is those for whom offending is most pertinent, i.e., the younger aged (Nagin et al., 1995) and is the focus of this current work. What is troublesome is the attrition between waves - although the retention rate between two waves is generally very good (over 70% between any two sweeps), the largest singular loss of members happened between the 2003 and 2004 sample, where over 1,200 of the core 10- to 25-year-old population (or 26% of the sample) were not re-interviewed and a large number (1,842) were added as a fresh part of the panel. The knock-on effect of this early sweep one to two attrition means that constructing a 2003 to 2006 longitudinal dataset leaves a sample of 2,343 (slightly smaller than the size of the Home Office constructed longitudinal dataset used by Hales et al. in 2009 due to the missing 25–year-olds issues noted in section 4.4.1). If, however, the process of joining the sweeps together begins with the 2004 sweep, with lower year on year attrition (less than 20%), the achieved size is 3,528; over 50% larger than the possible 2003-start dataset. Coupled with the change in questionnaire design, the construction of a longitudinal dataset capturing all four waves is problematic and, despite the discomfort caused by ignoring data, it was decided to focus on the 2004 to 2006 data and create a longitudinal dataset on this basis.

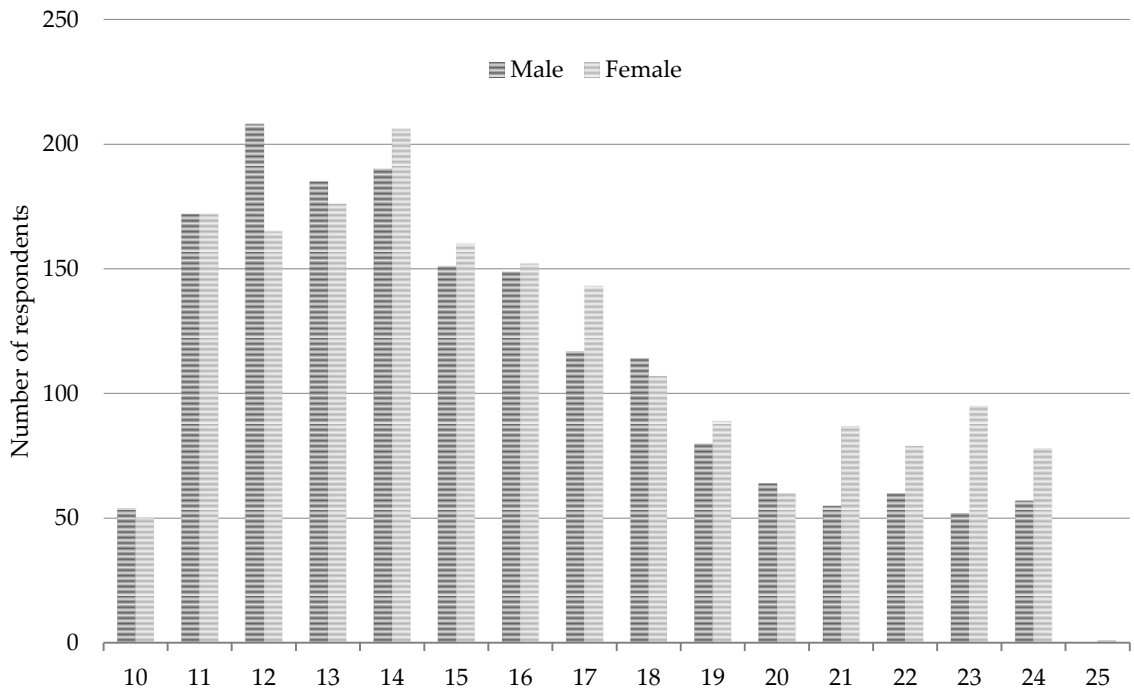
4.4 OCJS 2004 to 2006 Dataset Basics and Weights

4.4.1 Basics

Of the 3,528 respondents who were able to be matched together to form a 2004 to 2006 longitudinal dataset, 48.4% were male and 51.6% female. Figure 4.2 presents a split of the data by age and gender.

[6] By that it is meant 25-year-old respondents in 2003. Those who formed part of the longitudinal sample were followed up even if they were over 25.

Figure 4.2 Number of respondents in 2004 to 2006 longitudinal dataset, by gender and age (in 2004)



It is worth noting that:

1) There is an over-representation of the respondents aged 11 to 17. This is deliberate and due to the "young person boost sample in 2003", which was conducted to increase the number of young people due to their particular analytical importance and the fact that they are often a slightly harder to reach group (OCJS user guide, 2003).

2) There's only one 25-year-old in the final dataset! This can be nothing less than a mistake on the part of the data depositors. In the 2004 and 2006 data that was deposited in the data archive there were two versions of the dataset, the 10 to 25 version and a separate one holding the information on any members who were 25 at a previous wave and now over 25. In 2005 this dataset is not present, with just the 10 to 25 version available, and is not recoverable from the NatCen systems. This results in respondents aged 25 years old in 2004 (and hence 26 in 2005) being lost

from the dataset, with the exception of one female who was presumably interviewed at different times of the year, either side of her birthday.

Technical work on the Home Office constructed longitudinal data found that attrition was not related to reported offending, in fact the "main difference in retention rates by socio-demographic characteristics was that of age... those aged 18 and over were more likely to fall out of the sample, and this was presumably due to their leaving the parental home. On the other hand, the sample members aged 10 to 15 were the group most likely to be interviewed again, and this was linked to their family situation and compulsory attendance at secondary school" (Hales et al., 2009, p41).

4.4.2 Weights

As datasets produced primarily for Home Office usage, and to inform policy and highlight trends in self-reported offending on a cross-sectional basis, as is standard practice NatCen generated weights for each of the sweeps. A 2003 to 2006 longitudinal dataset for use with the accompanying longitudinal Home Office publication (Hales et al., 2009), holding just under 2,500 respondents, also has weights calculated and released with the data (along with a severely reduced set of response items). The Home Office states that the weights were important to correct for oversampling of those of a younger age, different ethnicities and some regional differences (ibid.).

The 2004 to 2006 dataset constructed for this analysis does not have weights for two reasons. Firstly and pragmatically, there are no entirely appropriate weights that *could* be used: the 2003 weights would not be appropriate, as they would be missing around 1,000 respondents who entered the panel in 2004, and the cross-sectional 2004 weights would also not be suitable as the longitudinal dataset misses members of the cross-sectional data that were not followed up. The longitudinal weights would miss those who were not in the original 2003 data but appear in 2004. This means weights would have to be constructed specifically for the dataset from scratch. Although appendix F of the OCJS technical report (2004) gives good methodological information on how the weights were calculated (a combination of selection, non-response and calibration weights), without knowing the exact selection and non-response fractions originally

applied to the data, nor the exact sampling procedures, creating weights for the 2004 to 2006 dataset would be at best an imprecise process and at worst wholly inaccurate.

The second reason why weights may not be wanted, regardless of their attainability, is that there is considerable debate as to whether using weights when modelling is even at all appropriate. When using surveys to analyse univariate characteristics about a certain target population, it is clear that survey weights are rather vital (see Chapter 17 of Leeux et al., 2008), but there is no clear answer as to whether it is necessary when using a survey for modelling work. In the case of ordinary least squares (OLS) regression it is established that when the weights are solely a function of the independent variables then un-weighted OLS estimates are preferred as they are unbiased, consistent and have smaller standard errors than weighted OLS (Winship and Radbill, 1994). Given that analysis of the OCJS attrition found that the drop-out was unrelated to whether the respondent claimed to have offended (Hales et al., 2009), longitudinal dataset members' non-response was unrelated to reported offending and extrapolating the same logic as cited by Winship and Radbill, un-weighted analyses should give more reliable results. Indeed, Gelman (2007) states that "survey weighting is a mess. It is not always clear how to use weights in estimating anything more complicated than a simple mean or ratios" and "even if weights are present, it may not be wise to apply them to the analysis" (p153).

From both a practical and methodological angle it is clear that weights are unneeded and analysis shall proceed on an un-weighted basis. The impact of this will mean that any descriptives about the data will not be generalisable to the population (i.e. the population of England and Wales) in the strictest sense. This would be especially the case if discussing age and ethnicity splits, but less so when considering overall offending rates, due to the oversampling of the former two characteristics. Descriptives have been presented elsewhere (Budd et al., 2005; Wilson et al., 2006; Roe and Ashe, 2008) and are not the focus of the current work and the modelling will not suffer from this limitation.

4.5 Question Structure and Format

With the exception of the offending questions (which will be discussed in section 4.6) the practicalities of the way in which questions were asked, in which order and by which method, shall be left un-discussed. NatCen and BMRB are well-established and professional organisations with excellent methodological reputations and are employed on numerous projects. One has to assume that their interview methods were adequate and one of the primary advantages of the secondary data analysis, namely high quality data, will hold true. This observation, however, does not mean that the work is free from flaws: missing values, non-response, misunderstandings, interviewer reliability and so forth are impossible to eliminate completely, but it is fair to assume that the impact of these have been minimised as far as practically possible (Bryman, 2004). It is not the intention here to criticise the data collection process. What is of relevance, though, is information and specifics about how the survey was conducted and their impact on subsequent analysis.

Firstly, presumably for reasons of time saving and respondent burden, the survey was split into two halves and a good deal of questions were divided so that half were asked of one set, the other half the other. As explained by the OCJS technical reports this was applied to respondent attitudes to the criminal justice system: "for several questions in this section the sample is split in half. Half get asked version A questions and the other half get asked version B questions. Assignment to sample A or B was random for fresh respondents… panel respondents were asked the same set of questions as previously". A range of CJS attitudinal questions[7] were asked with every respondent aged over 15 asked about street robbery, sample A (and over 15) were asked about car crime and sample B (and over 15) were asked about burglary.

---

[7] "Do you think X has gone up/down/stayed the same?", "What would stop people committing X?", "What is the main way people who commit X are caught?", "Out of how many hundred people who do X, how many are caught?", "What is the most likely thing that would happen to someone who was caught for the first (then later third) time doing X?", "How good do you think each of the following sentences would be in stopping them committing another car crime in the future? - fine of one month's salary, probation, prison less than 3 months, prison 3 years or more."

A similarly slightly complicated design was also administered in the *Health, lifestyle and risk factors* section of the questionnaire, a crucial part for the current research. The questions here were conditioned not on random assignment but on the age of the respondent. For some questions these differences are rather sensible, for example the questions about family vary slightly for those aged 10 to 16 to those aged 17 and over. "*How well do you get along with your (step) father*" is asked only of those with a (step - if appropriate) father and is asked as a constrained question "*How* did *you get along with your (step) father between the age of 10 and 16?*" of the older group (and similarly for the respondent's relationship with their (step) mother). This is clearly designed to tap the experience of the familial interaction during a certain time of the respondent's life.

For other questions the difference across age ranges is harder to explain - for instance the 2004 OCJS user guide lists two sets of questions as *Strengths and Difficulties*, one for children (i.e. those aged 10 to 16) and another set for adults (those age 17 or over). One may only hypothesise fruitlessly as to the reason for the difference, but the younger group's questions included [*How much do you agree/disagree with...*]:

- *I get upset if I see people who are sad or hurt*
- *I worry a lot*
- *I am easily bored and find it hard to concentrate*
- *I am usually helpful towards others*

Whilst the older group were asked a relatively similar but substantively different set including:

- *I get upset if I see other people suffering*
- *I am nervous in new situations. I easily lose my confidence*
- *I get bored easily*
- *I easily lose my patience with people*

There are also some areas of questioning for which there is simply no parallel between the two age groups. Often the reasoning is obvious - the assumed reduced role of the parental unit with the older respondent, for instance (*"How much would your parents*

mind if they found out that you had a) started a fight with someone, b) wrote things or spray painted a building, c) been caught smoking cannabis?" are asked only of those aged 16 or under). Other differences are a little less obvious and a response from the older group would have been an interesting piece of information. For example *"How wrong do you feel it is for someone of your age to start a fight with someone? And how wrong do you think it is for someone of your age to write things or spray paint on a building? And what about someone of your age smoking cannabis?"* are not asked of the older group, yet the reason does not appear to be apparent.

Although for the majority of questions asked differently of two different groups the split point is either the sample A/sample B random allocation or by age group 10 to 16 years and 17 years of age and over, on some occasions the question design changed based on 10- to 15-year-olds and 16 years and over. This is particularly pertinent in the *Neighbourhood* section of the questionnaire where, in a range of 13 questions asking about neighbourhood safety (crucially, an analogous group of questions used by the PHDCN to measure collective efficacy), four questions were asked only of the older group: *"I trust most people who live in this area?", "People in this area pull together to improve the area?", "People move in and out of my area a lot?"* and *"Suppose you dropped a purse or wallet in a street near where you live, with your name and address in it. How likely is it that you would get it back with nothing missing?"* The impact of these differences in questioning will be a constant theme that shall be revisited throughout the measurement section of this work (Chapter 6).

Almost all of the questions were asked with a Likert-type scale response set, the majority on a five-point *strongly agree, agree, neutral, disagree, strongly disagree* set, some questions on a four-point *very likely, fairly likely, fairly unlikely, not at all likely*, on occasion with a *yes/no* answer and even fewer with a 3 point *yes a lot, yes a little, not at all* type response pattern. All questions allowed refusal/don't know responses (and some responses are simply left blank with no data, not even missing data, added), although together these are typically only for around 30 respondents per question[8] . The

[8] The level of non-response depends slightly on the nature of the question but for most questions is well below 2% of the responses.

exceptions to the Likert-type responses were those questions that naturally lead to an interval level answer such as age or number of offence committed in the previous year.

To reiterate - the identification of the above questionnaire design intricacies is not criticism of the data collection process, to enter into which would at the very least be a hollow endeavour, as the data is collected and cannot now be altered. The debate is intended to draw the attention to some broad brush issues that shall have to be dealt with during the substantive analytical chapters. These themes and their impacts will be returned to throughout the work, both in the methods chapter and the methodological implications that the questionnaire design has enforced, and in subsequent chapters, especially the measurement sections, where the impact of these analytical restrictions shall be seen and dealt with.

4.6 A Discussion on Self-report Offending

A key part of any quantitative investigation is the concept measurement (see Chapter 6). For the majority of the concept measurements in this work, this shall be a key part of the *analytical* process, discussed in detail later in the work. One area which stands as an anomaly and deserves its own *methodological* consideration is that which will ultimately form the dependent variable: the indicator of whether respondent is an 'offender' or not.

The OCJS is a self-report study of offending - no official arrest, police, court or other types of administrative data are present. Thornberry and Krohn (2000) thoroughly review the origins of the self-reported method, charting its growth, refinement and role in criminological research at the turn of the 20th century. They identify a widely used methodology that has overcome a good deal of past limitations and claims that with some further refinements should deliver even more reliable results. Self-reported offending research began in earnest as late as the 1950s (in the US, at least), with previous criminological work predominantly based on official records such as arrest data or recorded crime. In reaction to Sellin's (1931) observation that the value of crime rates for index purposes decreases as the distance from the crime itself in terms of procedures increases, Sutherland (1949) is credited as taking the first major step toward

self-report offending research, albeit a rather "methodologically unsophisticated" (Thornberry and Krohn, 2000, p36) approach, problematic in terms of sample representivity, selection of delinquency items and failure to examine the reliability and validity of these items. Following from this, the work by Short and Nye (1957) "revolutionized ideas about the feasibility of using survey procedures with a hitherto taboo topic" (Hindelang, Hirschi, and Weis, 1981, p23) and applied a considerably more developed methodology (in terms of scale construction, reliability and validity, and sampling). This work found, for the first time, irreconcilable analytical differences in the relationship between socioeconomic status and offending. Official reports had always implied the relationship that low socioeconomic status was highly correlated with offending behaviour, but self-report questioning found the relationship to be much weaker or even non-existent. These findings not only raised questions of the independence from extra-legal influences of the juvenile justice system, but also the nature of exactly what it is that self-report is measuring.

Hindelang, Hirschi and Weis (1981) added victimisation data to the arsenal of measures and argued that self-report tapped a different domain of behaviours, missing the highest-rate offenders but capturing what Gibbons (1979) referred to as the 'dark figures of crime' i.e. those offences committed but not reported to the police. For example, violence between strangers is well reported to the authorities whilst domestic violence or violence between two people who know each other has a much lower official report rate. Adding to this, Elliott and Ageton (1980) observed that a relatively small number of young people committed a disproportionately high number of serious offences and this was being missed due to truncation of response categories. Their criticism of the prevalent design at the time was that typically this 'early' self-report method used a small (seven or nine) number of items to represent delinquency with the response categories of 'no', 'once or twice', 'several times' and 'often'; had ill-defined reference periods (such as 'since you started school'); had no measure of violent offending; and, theft was restricted to only those occasions where the item was valued over $2.

Collating their observations and refinements to date, Thornberry and Krohn (2000) go on to outline four simple procedures that must be followed to create reliable and valid self-report measures of offending:

- Self-report scales should include a wide range of delinquent acts so that the general domain of delinquency, as well as its various subdomains, is adequately represented.
- The scale should include serious as well as minor acts.
- A frequency scale should be used to record responses so that high-rate offenders can be isolated from low-rate offenders.
- Extremely trivial, non-actionable acts that are reported should be identified and eliminated from the data through the use of follow-up questions (e.g., "How much was the item worth?" or "Could you describe the fight?" and then censor the data accordingly).

If these guidelines are followed, Thornberry and Krohn (2000) conclude the same as Hindelang et al. (1981), that the "self-report method appears to behave reasonably well when judged by standard criteria available to social scientists. By these criteria, the difficulties in self-report instruments currently in use would appear to be surmountable; the method of self-reports does not appear... to be fundamentally flawed. Reliability measures are impressive (as reliable as, if not more reliable than, most social science measures) and the majority of studies produce validity coefficients in the moderate to strong range" (Hindelang et al., 1981, p114).

Finally, Thornberry and Krohn move on to the methodological techniques best suited to administration of a self-report questionnaire, noting, "One of the most promising developments in the self-report method is the advent of audio-assisted computerized interviews. This technique offers increased confidentiality to the respondent in an interview setting. Although somewhat expensive and complicated to design, the early studies indicate that it may be worth the effort" (Thornberry and Krohn, 2000, p73). Wikström (2012) contends that "self-report data (collected using rigorous and well-executed methods) provides the best and closest approximation of young people's real level of crime involvement" (p110).

As to be expected from an organisation with the reputation of NatCen, the self-report data collection process followed all of the best-practice advice that was available. The offending part of the questionnaire was conducted via audio computer-assisted self-interviewing (audio-CASI) and asked about a range of offending behaviours in a specific order[9] with the line "*Apart from anything you have already mentioned…"* reiterated before each question. Whether the interviewee had been arrested, cautioned, been to court, fined, sentenced to supervision or custody was asked of everyone, regardless of whether they had reported any offending in the previous section. Nested time frames were used, asking first if the respondent had ever committed the offence, then whether they had committed it in the last 12 months and finally if they had in the last four weeks. For the 'ever' and 12 month time frame the respondent was allowed to identify any number of offences on a frequency scale, whereas for the four-week time frame the respondent was asked simply whether they had committed the crime in that last four weeks or not (binary yes/no response). The utility of the 'ever offended' time frame is rather contested, with Junger-Tas and Marshall (1999) arguing that "the researcher has to weigh efficiency against minimizing error related to telescoping [i.e., incidents are recalled as occurring more recently than they actually did]" and "it would seem that answers concerning a short time span, such as the past six months, would be more accurate than answers concerning longer time periods" (p341). It would be sensible, therefore, to concentrate analysis on just the shorter time periods of 12 months.

Thornberry and Krohn recommend using follow-up questions to validate any initial responses to the offending question to 'clean' inaccurate responses. The OCJS followed

---

[9] Order of crimes asked about: 1. Stealing and driving away a vehicle 2. Attempting to steal and drive away a vehicle 3. Stealing parts off the outside of a vehicle 4. Stealing from inside a vehicle 5. Attempting to steal from inside or parts off the outside of a vehicle 6. Deliberate damage to a vehicle 7. Breaking in to someone's home with the intention of stealing or damaging property 8. Breaking into a non-domestic building (office, shop, school, etc.) with the intention of stealing or damaging property 9. Criminal damage/arson 10. Use of force, violence or threats to steal from a business 11. Use of force, violence or threats to steal from a person 12. Stealing from someone without use of force, violence or threats 13. Stealing from respondent's workplace 14. Stealing from respondent's school/college 15. Stealing from a shop without using force, threats or violence 16. Other thefts 17. Use of force or violence which injured someone 18. Use of force or violence which did not result in injury 19. Selling of Class A drugs 20. Selling of other (non-Class A) drugs

this advice as well as capturing a richer understanding of the specifics of any offending. Any respondent who reported offending in the previous 12 months was asked to remove the audio-CASI equipment and the interviewer, using the conventional CASI techniques, covered detailed questions on the nature of the six most serious[10] offences committed in the last 12 months and these were used to validate and clean the data, as well as providing detailed information for analytical work. Although the questions are not directly relevant to the current work and the exact questions varied according to the type of crime, in general the following information was collected in the nature of offending module and it is clear that such a set adequately covers the requirement of follow-up questions to be able to clean the initial responses:

- Whether the crime was committed in respondent's local area
- Type of building/shop (non-domestic thefts/burglaries only)
- Time of day/evening that it was committed
- Nature and value of items stolen (thefts, burglaries, robberies)
- What was done with stolen items – e.g. if sold on (thefts, burglaries, robberies)
- Age, gender, ethnicity and relationship or connection of the victim(s) to the respondent
- Type of force or violence used (where relevant)
- Whether committed the crime by themselves or with others, the number of co-offenders, details of their age, gender, ethnicity and their relationship to the respondent
- Why the crime was committed
- Whether crime was alcohol or drug-related
- Whether crime was planned or not
- Choice of target – why did they choose to burgle a particular building/steal from a particular car/person etc.
- The respondent's perceived likelihood of being caught
- Whether the police found out and the consequences of this

[10] As selected by the Home Office to match policy and perceived severity interests; the ordering can be found on page 24 of the 2004 OCJS technical report

- Whether respondent considers they will ever commit this offence again

Considering the stated aim for the OCJS was to be the first nationally representative survey of self-report offending, the methodological steps that were followed and the reputation and professionalism of both of the centres that conducted the work (NatCen and BMRB) as well as the governmental body responsible for the publication of the results (the Home Office), it is fair to assume that, despite the noted problems with reliance purely on self-report offending, this source should provide as accurate data as possible.

4.7 Self-report Offending, Basic Descriptives

Although the rates of reported offending are the focus of reports elsewhere (see Budd et al., 2005; Wilson et al., 2006; Roe and Ashe, 2008), it is worth recalculating and representing the findings based on the longitudinal 2004 to 2006 dataset to be used in this work. These measures will ultimately form the dependent for this quantitative investigation and therefore detailed knowledge of their reported rates is vital. First up, the overall number of self-reported offences in the previous 12 months is given in Table 4.1.

Table 4.1 Self-report offending, 2004 to 2006, longitudinal dataset (all ages)

	2004	**% of valid**	**2005**	**% of valid**	**2006**	**% of valid**
None	2,399	75.7%	2,535	74.5%	2,700	78.9%
One	232	7.3%	245	7.2%	231	6.8%
Two	133	4.2%	172	5.1%	121	3.5%
3 to 5	185	5.8%	198	5.8%	153	4.5%
6 to 10	87	2.7%	114	3.4%	92	2.7%
More than 10	134	4.2%	137	4.0%	124	3.6%
		% of all		**% of all**		**% of all**
Missing	358	10.1%	127	3.6%	107	3.0%

As can be seen, the majority of respondents in each wave do not report offending. Of the remaining (approximate) quarter the most common reported level is just one offence, yet there is a non-trivial number who report rather frequent offending, with over 3.5% of the sample in any wave reporting committing more than 10 offences.

Between the 2006 sweep and both of the earlier time points there is statistically significant difference in the proportion who reported that they had not offended in the preceding 12 months. The rate is larger in the 2006 sweep (significant at the 1% level), though whether this is due to a true fall (perhaps attributable to an aging cohort) a type one error or due to respondent fatigue and being asked the same question over the year is unknowable. Despite the statistical difference the same substantive pattern can be seen in each sweep - between 75% and 80% of respondents report no offending in the preceding 12 months and there are no significant differences between the numbers reporting one, two or more offences. The only other significant difference is in the unusually high number of missing values in the 2004 sweep, the reason for which remains a mystery.

Table 4.2 shows that, despite the identification of around 75% of the sample not reporting offending in any one year, over the period 2004 to 2006, less than 60% consistently didn't report offending. 22% reported offending in just one year, 10% in adjacent years, with 2.4% reporting offending in 2004 and 2006, but not in 2005. 7% of the sample reported offending in all years and, interestingly, only six of this group of 217 said they had committed just one offence in each year. This latter group would fit Elliott and Ageton's (1980) relatively small number of youths who commit a disproportionately high number of crimes.

Table 4.2 Self-report repeat offending, longitudinal dataset

	Freq.	**% of valid**
Never	1,768	58.47
04 only	259	8.56
05 only	216	7.14
06 only	187	6.18
04 and 05	180	5.95
05 and 06	125	4.13
04 and 06	72	2.38
All years	217	7.18

Looking at the correlation[11] between years (Table 4.3) a moderately strong positive correlation is seen between years, with reports of offending from adjacent years correlating most strongly. This shows there is a good level of variation in the reported levels of offending, with the most stability exhibited between adjacent time periods.

Table 4.3 Pair-wise Spearman's correlation between number of offences per year

	Number of offences 04	Number of offences 05	Number of offences 06
Number of offences 04	1		
Number of offences 05	0.439	1	
Number of offences 06	0.304	0.441	1

The individual offences types as listed in the survey questionnaire provide a fine-grained view of the reported offending of the respondent (as should be the case from a methodological data collection point of view). For many types within the survey, however, there were very few respondents reporting having committed the offence. To aid with analysis, 10 summary offences types were created by the data creators and supplied as derived variables with the dataset. These capture and summarise the 20 raw offences. Details of the groupings are given in Table 4.4, along with the number of people reporting having committed the lowest level offence grouping (i.e. the original question) once or more in the 2004 sweep. The categories *violent, property* and *serious* offences constitute 'higher level' or second order aggregate groupings, and capture a combination of offences that are themselves first defined into the seven lower groupings.

Table 4.4 Offence categories grouping and original question reporting rate for 2004

	1+ in 2004	**% of valid**
Vehicle theft offences - including:		
• stole a vehicle	16	0.5%
• tried to steal a vehicle	10	0.3%
• stole from outside of vehicle	37	1.1%

[11] Spearman's rank correlation coefficient is presented to account for extreme responses. A very similar pattern is seen when using the raw number of reported offences and top coding all reported frequencies over 10.

	1+ in 2004	% of valid
• stole from inside vehicle	10	0.3%
• tried to steal from outside or inside a vehicle	11	0.3%
Criminal damage offences - including:		
• vehicle damage	71	2.0%
• damaged property	135	3.9%
Burglary offences - including:		
• domestic burglary	5	0.1%
• commercial burglary	31	0.9%
Robbery offences - including:		
• commercial robbery	6	0.2%
• personal robbery	3	0.1%
Other theft offences - including:		
• theft from person	31	0.9%
• theft from work	121	3.4%
• theft from school	252	7.5%
• shoplifted	88	2.5%
• other theft	83	2.4%
Assaults - including:		
• violence with injury	391	11.3%
• assault no injury	415	12.0%
Drug offences - including:		
• dealing class A drugs	19	0.5%
• dealing other drugs	102	2.9%
Violent offences - including:		
• all robbery, all assaults	585	17.2%
Property offences - including:		
• all vehicle theft, all burglary, all other theft	447	13.7%
Serious offences - including:		
• stole a vehicle, domestic burglary in last year, commercial burglary, commercial robbery, personal robbery, theft from person, violence with injury, dealing class A drugs	423	12.3%

It can be seen that the reported levels vary significantly between offence types (displayed in Table 4.5, along with the Spearman's correlation between years), from just 0.3% (9 people) reporting having committed a robbery offence to 17.2% (585 people)

reporting an assault offence in 2004. Of the higher level offence categories (violent, property and serious offences) there is a higher apparent rate of offending, but in reality these are driven by a smaller subset of the original offences types. Serious offending is predominantly (over 90% in each year) violence with injury. Almost all violent offending is violence with injury or assault with no injury (as reported robbery was such a rare event). Property offences are almost universally 'other theft offences' which are themselves predominantly either theft from school or work. Yearly correlations are typically of the range seen with all offences (Table 4.3), once the large volatility of offences with low report rates such as burglary and robbery are discounted.

The impact of these observations is important from a substantive point of view in that it demonstrates that there is considerable variation over the offence types committed by the cohort, but the majority of offences committed are of a relatively low 'seriousness' (despite the analytical difficulties in measuring such a concept - see, for example, Francis, Soothill and Dittrich, 2001). This will be analytically important in later modelling work. Prediction of rare events leads to large standard errors and imprecise estimates, rendering models unreliable, and this preliminary analysis will inform the possible future dependents. Judging from Table 4.5 it is likely that only the all offending, violent (basically assault with or without injury), property (low level theft) and serious (violence with injury) offending will hold enough variation to be able to be modelled. This will be returned to in Chapter 8.

Table 4.5 Percentage reporting one or more offence, by offence group and year, with Spearman's correlation between years

	2004	**2005**	**2006**	**Correlations**		
Number of times committed...	**1+**	**1+**	**1+**	**2004 - 2005**	**2005 - 2006**	**2004 - 2006**
... vehicle theft offence in last year	1.7%	1.8%	1.8%	0.22	0.29	0.12
... criminal damage offence in last year	5.1%	5.0%	4.0%	0.25	0.35	0.14
... burglary offence in last year	0.9%	1.1%	1.2%	0.13	0.18	0.07
... robbery offence in last year	0.3%	0.2%	0.1%	0.00	0.16	0.00
... other theft offence in last year	11.7%	12.0%	10.3%	0.33	0.33	0.20

	2004	2005	2006	Correlations		
Number of times committed...	**1+**	**1+**	**1+**	**2004 - 2005**	**2005 - 2006**	**2004 - 2006**
... an assault in last year	17.2%	17.2%	12.3%	0.36	0.37	0.24
... a drug offence in last year	3.0%	3.3%	3.3%	0.50	0.42	0.32
... a violent offence in last year	17.1%	17.3%	12.3%	0.36	0.37	0.24
... a property offence in last year	13.7%	14.5%	12.9%	0.37	0.38	0.25
... a serious offence in the last year	12.3%	12.9%	10.2%	0.33	0.35	0.23

4.8 Data round-up

A range of possible secondary data sources were considered as a basis for this investigation and judged against criteria outlined by Dale, Arber and Proctor (1988). The Offending and Crime Justice Survey (OCJS) was deemed an appropriate data source in that the original study aims match the research question adequately and with a representative sample collected by a reputable and reliable research organisation in a contemporary time period (2003 to 2006). Other data sources that were considered were the 'traditional' UK social science dataset such as the 1970 British Cohort Study, the National Child Development Study, the Longitudinal Study of Young People (LSYPE) and the Avon Longitudinal Study of Parents and Children (ALSPAC), which were ruled out, predominantly due to a lack of detailed offending questions. The Edinburgh Study of Youth Transitions and Crime (ESYTC) and the Peterborough Adolescent and Young Adult Development Study (PADS+) were rejected as they are not publically available (yet) and, although an analysis of a restricted geographical area could still give important results, were not nationally representative and as such do not fit exactly against the research question. Other data sources, such as the Arrestee Survey and the Crime Survey for England and Wales, failed either because of a lack of detailed offending variables or, in the case of the Youth Lifestyle Survey, due to the age of the data.

The OCJS was conducted by the National Centre for Social Research (NatCen) and British Market Research Bureau (BMRB) on behalf of the Home Office between 2003 and 2006. Due to its sampling frame and data collection procedures, it is representative of adolescents and young adults up to 25 years of age, and hence covers the key age group of 10- to 16-year-olds that is of interest to this work. It has a range of variables about socio-demographics, individual attitudes, lifestyle, family circumstances, personal relationships, attitudes to and experiences of the criminal justice system and, crucially, self-report offending indicators. The self-report offending questions followed best available practice as reviewed by Thornberry and Krohn (2000) with a methodologically sophisticated approach using audio-assisted computerized interviews, asking about a range of behaviours with detailed follow-up questions used to validate the information collected.

The dataset appears to have been underutilised but those reports that have used the OCJS have demonstrated that it is of a sufficient standard to be used to control for known offending risk factors and in longitudinal analysis. However, due to some design limitation, there are some (surmountable) issues: the 2003 wave was found to be different enough in terms of some key variables and its attrition rate to the latter waves as to need to be ignored, there are some differences in the question administration based on either a random allocation or age banding that need to be accounted for in the analytical approach; and finally, although the offending questions are of a good standard, only four categorisations are likely to be useful in regression-type analysis. These data will form the basis for this investigation.

Chapter 5

A Short Overview of Structural Equation Modelling

5.1 Advantages and Limits of Secondary Data Analysis

The advantages of secondary data analysis are well known and well reported, and form part of any social research methods degree programme. Amongst the most prominent positives are the costs and time saved to the researcher, who does not need to contend with often laborious and complicated data collection processes, with the resultant data of high quality and often of sufficiently large sample size so that sub-group or cross-cultural analysis may be possible. At a more immediate level the researcher should have more time to engage with the data, understanding its nuances and the act of reanalysis of the data may lead to fresh insights and new interpretations (Bryman, 2004). Bryman also argues that it is an obligation to the participants within the survey that the research community should try to extract maximum value from any collected data — a relatively minor point, but nonetheless pertinent with the Offending, Crime and Justice Survey (OCJS) due to its relatively low usage rate.

A further advantage of secondary data that is often overlooked is touched on by Kiecolt and Nathan (1985) is that widespread familiarity with data not only allows the researcher/research team conducting the analysis to work to a high standard in less time, but also allows the wider community to understand the findings and limitations of the work with less effort. For instance, the American General Social Survey (GSS) is widely used, and work based on it can be reported and understood more easily than new data sources where a good deal of any publication may need to include explanation of exactly what has been captured, the limitations of the data and who exactly the population represents. Applying this logic to the OCJS may seem slightly erroneous given its aforementioned low usage, however the overall methodology and many of the questions that were used to form the survey were in common usage; the OCJS technical report (2003) states that the survey was conceived as the third part of the

triumvirate including the established Crime Survey for England and Wales and the now defunct Citizenship Survey, sharing many of the same design features and with analogous question modules, target population and sampling design.

Although the advantages of secondary data work clearly outweigh the negatives, there are some clear issues that need to be kept in mind. Again Bryman (2004) furnishes us with a list including the fact that the researcher's lack of familiarity with the data coupled with its possible complexity can lead to errors of understanding. There is also no control over the data quality, which may sound paradoxical given that high quality is one of the main attractions to secondary analysis, but what is meant here is that it is the quality across certain variables that may be problematic (for instance, the Labour Force Survey in its basic form is not suitable for sub-regional analysis and certain adjustments need to be made before progressing). Kiecolt and Nathan (1985) expand this point, noting that sometimes only aggregate level data may be present and that any errors in the original work are forgotten and/or unknown, and analytical strategies to deal with these that the original researchers followed may be missed. As noted in section 3.3 the OCJS has issues with missing data, with question differences across ages or through random allocation: to handle these issues detailed attention to the technical reports is needed, and to err on the side of caution when problems arise.

Kiecolt and Nathan (1985) finish their piece with two further, linked warnings about the possible consequences of over-reliance on secondary data research. The first is that without new items and sources being added to the researchers' armoury then there is a risk that sources will be over-analysed, inhibiting creativity. Whilst this warning is no doubt true, it seems clear that developments since their work was penned to the current day show that their call for new data sources has been heeded and social science has not stagnated. The second of the two warnings, however, is certainly still pertinent: "Some researchers obtain a dataset, apply a currently popular statistical technique, and then look for a problem to investigate. Without theory, however, the utility of social research is called into question.... the 'data set in search of analysis' approach yields only trivial findings" (p135).

This warning against atheoretical data-led analysis is important, and can be found to be echoed across a number of fields. A particular fine and relevant restatement comes from the realm of cultural criminology and Hayward and Young, who state that "data that are in fact technically weak... and, by their very nature, contested, blurred, ambiguous, and unsuited for quantification, are mindlessly churned through personal computers", they lambast the trend towards "ill-developed theory, regression analysis... followed by inconclusive results" (Hayward and Young, 2007, p114). For the current work this warning must be constantly revisited. This is secondary data analysis using a modern statistical technique and is therefore prone to these criticisms at the most fundamental level. The weakness of the data, the steps taken to deal with these and the theoretical grounds on which these steps were made, and the exact nature of the constructs that are being created need to be acknowledged, repeated and adhered to at each stage, with any resultant limitations openly confronted.

5.2 Introduction to Structural Equation Modelling

The majority of statistical methods applied in the social sciences involve exploratory statistics, calculating a mean or variance, checking the equivalence of two proportions, testing whether a correlation has a statistically significant difference from zero, or whether a certain regression model and its coefficients have explanatory power over a certain dependent. It is common practice to assert a hypothesis that there will be a non-zero correlation and then test to see whether this holds true. This by itself holds limited intrinsic value, knowledge that a correlation between A and B is non-zero does not lend much support to a theory about the nature of the relationship between A and B, as any number of substantive theories may give this result. Structural Equation Modelling (SEM) is a different way of approaching these problems, with an emphasis on forming theoretically based models and then testing their validity against observed data.

Historically, the framework for SEM began with the factor analytical models of Charles Spearman (1904) combined with path analysis techniques, or simultaneous equation modelling, of the geneticist Sewall Wright (1918 and 1921). Spearman's initial factor analysis was concerned with the observation that children who performed well on one test of mental ability tended to do well on alternative tests. He hypothesised that

individual scores were manifestation of some underlying general ability, and although scores would not correlate perfectly the similarities across different instances of similar experiments pointed to this general ability level. Path analysis, on the other hand, sought to infer how the patterns of correlations should be among a set of variables if they had a certain specified causal relationships between them. It was at first illustrated by reference to determining the relative importance of interrelated factors affecting guinea pig birth weight, from gestation periods, rates of growth and litter size (Wright, 1921). Although many analysts worked in related realms, with analogous techniques (such as Keynes, 1936 and Bock and Bargmann, 1966) the analytical combination of the two schools originated with Jöreskog (1969), who led the development of simultaneous linear equations *with* latent variables. Crucially for implementation, Jöreskog developed and coded an algorithm to estimate the parameters and chi-squared goodness of fit tests. The program, known as LISREL for "linear structural relations", united the latent variables of factor analysis with the path analytical techniques and initiated a methodological revolution (Mulaik, 2009).

SEM is thus a collection of statistical techniques that allow a set of relationships between one or more independent variables, either continuous or discrete, and one or more dependent variables (again either continuous or discrete) to be examined. It is a general model of many commonly employed statistical techniques including analysis of variance or covariance, multiple regression, factor analysis, path analysis, item response theory, econometric models of simultaneous equation and non-recursive modelling, multilevel modelling, and latent growth curve modelling. Confirmatory factor analysis simultaneously combined with a regression is often given as an intuitive introductory example of the method but much more intricate systems are possible. At the simplest level this could be a relationship between an outcome variable and a set of measured variables (simple linear regression) but more complicated models allow for the presence of latent variables as either the outcome, independents or intermediaries. The relationships between the variables in SEM may influence one another reciprocally, directly or through other variables, which can be manifest or themselves latent. The presence of the latent variables allows for a reduction in dimensionality to aid interpretation of the structure of the data and, with the appropriately acknowledged

ontology (to be discussed in section 5.4), increased reliability of the measure of underlying concepts or causes. Appropriately constructed SEM can also account for the reliability of indicators, group differences, multilevel modelling and, although it is primarily a cross-sectional technique, the relatively new approach known as latent growth curve modelling can allow longitudinal data analysis to be conducted (Wang and Wang, 2012; Byrne, 2012; Bartholomew, Knott and Moustaki, 2011; Mulaik, 2009).

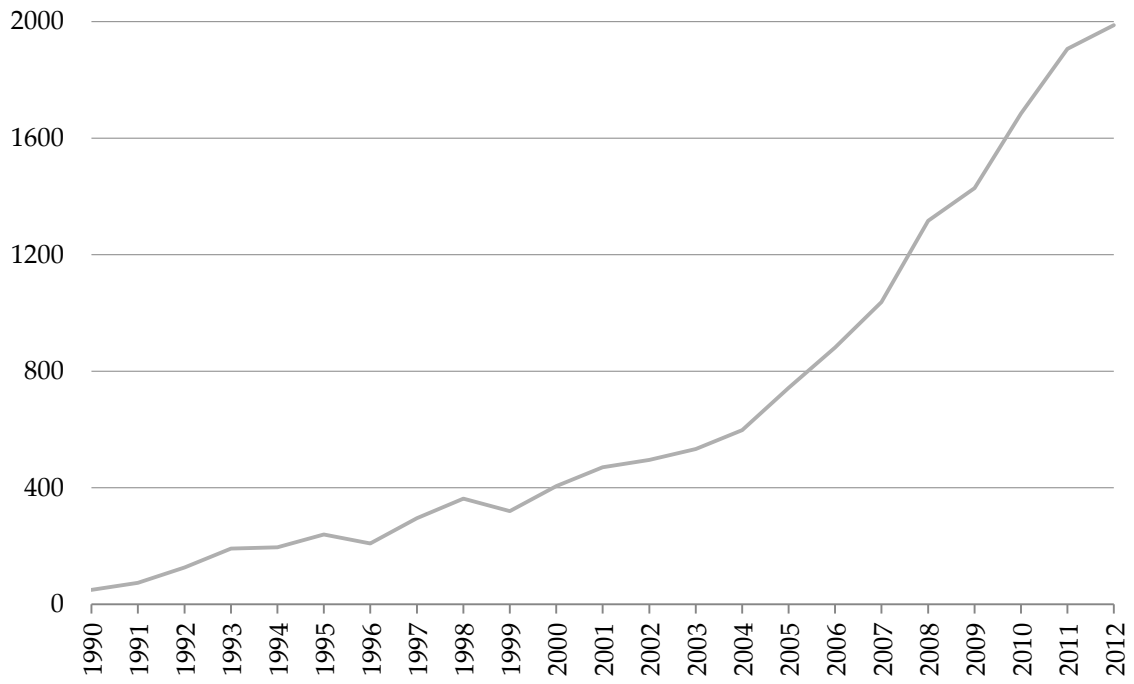
A cynical view of SEMs is that their popularity in the social sciences reflects the legitimacy that the models appear to lend to causal interpretation of observational data (though in reality such an interpretation is equally problematic for other kinds of regression models). A more charitable interpretation is that SEMs are close to the kind of informal thinking about causal relationships that is common in (criminological) theorising, and therefore, these models facilitate translating such theories into data analysis. In keeping with the warnings of atheoretical secondary data analysis in the previous section, it is of high importance that SEM be used primarily as a confirmatory technique rather than an exploratory method; "one cannot do SEM without prior knowledge of, or hypothesis about, potential relationships among variables... Planning, driven by theory, is essential to any SEM analysis" (Tabachnick and Fidell, 2007, p682). Offending behaviour is an ideal area for the application of SEM as there is a long tradition of empirically tested theoretical arguments. Many of the questions that were used to test these arguments have become commonplace and reduced versions or slight variants are included in OCJS.

SEM is an increasingly popular technique in the social sciences, with a steadily increasing number of journal articles found on Web of Science citing it as the methodology of use in the abstract since 1990 (Figure 5.1[12]). It has been applied to an extremely wide variety of research questions, including (to name a few) investigations into attractiveness (Riggio et al., 1990), memory capacity (Conway et al., 2002), the enticement of online consumers (Chen et al., 2002), the psychological bases of ideology

[12] A search for the terms "structural equation modelling" or "structural equation modeling" (to account for the American spelling) in the article abstract within the social science domain in October 2014

and prejudice (Duckitt, 2002), autism heritability (Lichenstein, 2010), ethical workplace decisions (Kish-Gephart, 2010), the big five personality traits (Marsh et al., 2010), immigrant youth adaptation (Berry, 2006) and teacher burnout (Hakanen, 2006).

Figure 5.1 Prevalence of structural equation modelling cited as methodology in the social sciences on Web of Science, 1990 to 2012



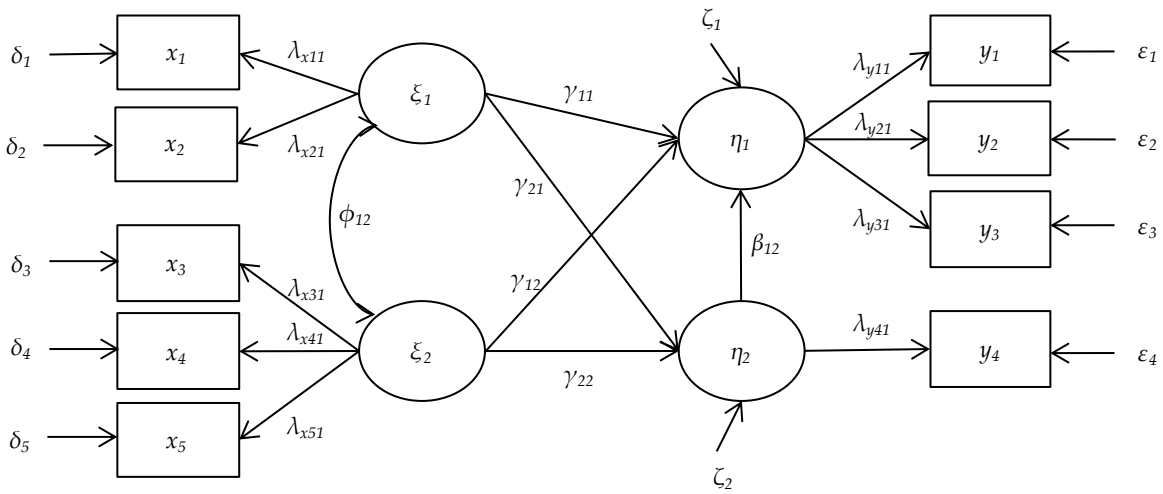
A wide variety of computer programs are available for SEM analysis. The original, LISREL, developed by Jöreskog in the 1970s has been added to by EQS (Bentler, 1995), AMOS (Arbuckle, 2006), various SAS packages, some open source R scripts (Fox, 2006), SPSS Amos and Mplus (Muthén and Muthén, 1998-2012). STATA12 had limited linear SEM capabilities (only continuous manifest variables) with a fine graphical user interface, and this was further developed in STATA13 which added generalised linear models and multi-level capabilities. Despite these developments, this work will use Mplus for all SEM modelling. SEM works by estimating variance/covariance matrices based on a set theoretical shape and comparing these to observed data, as such it relies on robust estimation and fitting algorithms, and Mplus has by far the widest and most versatile set of estimators for categorical data, which will be shown to be a key requirement. Brown (2006) contends that the WLSMV (weighted least squares with missing values) estimator is the best performing method when modelling categorical

data and this is only available in Mplus. The WLSMV estimator has been shown to give accurate test statistics, parameter estimates and standard errors under normal and non-normal latent variable distributions[13], the latter being extremely importantly for this work. It is also capable of dealing adequately with missing values, either missing at random or missing due to questionnaire restrictions (Asparouhov and Muthén, 2010).

5.3 Graphical Displays of SEM

Although all SEM models can be expressed in terms of matrix algebra they are more easily understood when expressed as a path diagram, and this representation is fundamental to the SEM researcher. Figure 5.2 is a taken from Wang and Wang (2012) and is an example of a hypothesised general structure for a structural equation model given in their introductory section, and this notation (or a reduced form thereof) is the SEM standard.

Figure 5.2 A fictional general structural equation model



Each different part of the model (shown by different shapes, arrows or Greek numerals) represents different empirical components:

- The ovals represent the latent variables;

[13] The term "latent variable distribution" refers to the observed ordinal distribution as generated from the unobserved continuous distribution assumed to underlie the observed categorical variable. This shall be discussed in section 5.5.

- Endogenous latent variables, i.e. those that are determined within the model (have at least one single-headed arrow pointing to them), are represented by a η (with subscripts). (Note that the arrow represented with a ϕ is *doubled-headed*);
- Exogenous latent variables ('lying outside the model' or without a single-headed arrow pointing into them) are denoted by ξ;

- The rectangular boxes display observed, or manifest, variables;
- The x and y observed variables are given different symbols dependent on whether they associate with exogenous or endogenous variables;
- The λ coefficients are the factor loadings linking the observed variables to the latent variables;
- Each of the observed variables has an associated measurement error term, denoted δ or ε depending on whether they are associated with exogenous or endogenous variables respectively. These may also be allowed to correlate (if this were the case, a two-headed arrow between the two relevant error terms would be added).

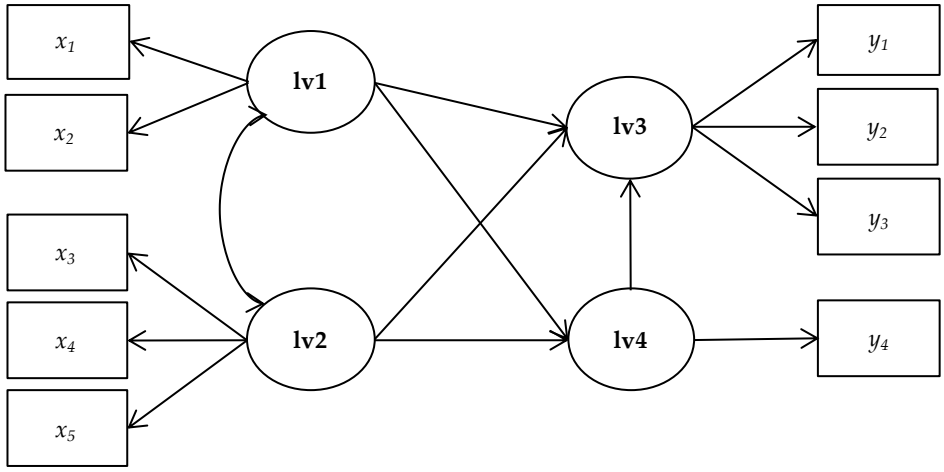
- The β and γ coefficients on the arrows between latent variables are path coefficients, altering between β and γ dependent on whether the 'causal' latent variable (i.e., the latent variable from which the arrow originates) is exogenous or endogenous;
- The ζ are structural residuals (as with regressions, nothing is predicted perfectly);
- The ϕ is a covariance between the two exogenous latent variables (there is a covariance between any pair of exogenous variables in the model).

One-headed arrows are predictive causal relationships whilst two-headed arrows are covariances or correlations. This is an important distinction, as the single-headed arrow implies that the point of origin of the arrow has some form of predictive relationship or causal effect on the arrow's resting place. In terms of Figure 5.2 therefore, the single arrows from the latent variable ξ_1 to x_1 and x_2 imply that the answers to these two observed variables are given by the respondent's underlying position on the latent variable (as discussed in section 5.4). In regard to the single-headed arrow between ξ_1 and η_1, this means that the respondent's position on ξ_1 (the exogenous latent variable) affects the respondent's position on η_1 and the position on this latent variable in turn is responsible for the answers y_1 to y_3. Two-headed arrows, on the other hand, represent correlations, or interdependencies, between components of the model. These can include correlations between the errors in the original questions (most often interpreted

as question bias correlations) but are most often of analytical importance when between two latent variables. For example, the value of ϕ_{12} between ξ_1 and ξ_2 indicates the correlation between those two latent variables, with no causal or predictive path implied or otherwise.

Figure 5.1 represents an extremely thorough and comprehensive presentation of a SEM model. In most applications, and in the current work, certain components can be left out of diagrams and notation rules can be relaxed. Most notably the error and residual terms can be omitted, leaving only an unmarked arrow or even completely deleted. Unless otherwise constrained these parts are always present, do not add to the model interpretation and have little presentational interest. The ξ and η (i.e. the latent variable labels) and the γ and β (the path labels) do not need to be identified in the notation differently. Whether the latent variable is either exogenous or endogenous means little, they serve broadly the same function and it is the interpretation of endogeneity versus exogeneity and their interrelations that is of importance. Attaching meaningful abbreviations is most useful and the distinction between endogeneity and exogeneity can easily be achieved via prose. Furthermore, in most instances when the model is set up there is no need to label the path coefficients, as they take a discrete value once the model is evaluated. They can easily be referred to as *"the path between {variable} and {latent variable}"* or even as "*{latent variable} with {variable}"*. In a model pre-estimation displaying the hypothesised structure, the path labels may therefore be left blank, and once the model has been estimated, the paths can be identified with their estimated value. A reduced presentation of Figure 5.2 can be seen in Figure 5.3 where the latent variables are named lv1 through to lv4 in lieu of substantive abbreviations (which would be given based on their approximate interpretation).

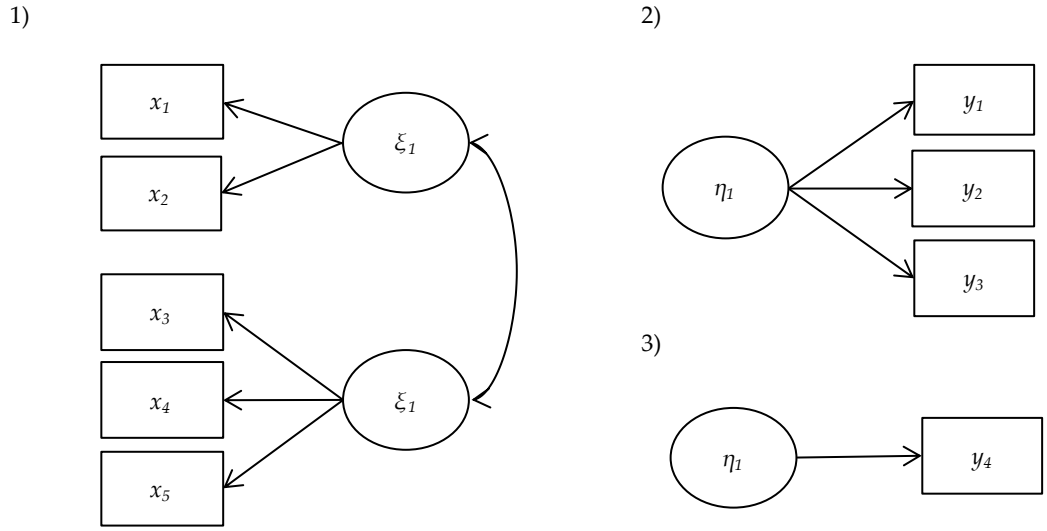
Figure 5.3 A simplified presentation of a full structural model



As the intellectual traditions from which SEM was born came from a combination of factor analysis and path analysis, so too does the modern application of SEM. The two components are the measurement model and the structural model. The first part, the measurement model, defines the relationship between observed variables and the latent components and gives a description of how well the observed data serves as a measurement instrument for the underlying constructs. These are ideally defined by confirmatory factor analysis, however, in reality there is often an element of exploratory investigation to be conducted to correct misspecifications (Asparouhov and Muthén, 2009), as will be discussed later. The SEM model in Figure 5.2 has three measurement models as outlined below and shown in reduced form in Figure 5.4:

(1) capturing the variables x_1 to x_5 and the ξ_1 and ξ_2, the associated λ values and the covariance (ϕ) between the two latent variables,
(2) is formed of y_1 to y_3 and η_1 and
(3) a single factor η_2 with y_4 (although this is not identified when estimated outside of the full structural model, also to be discussed later).

Figure 5.4 Measurement models present in Figure 5.1



The second step of the model is the structural part; once the latent variables have been constructed via the measurement model the hypothesised relationships amongst them is assessed, i.e. the γ and β values from Figure 5.2 are estimated. This procedural methodology shall be applied in this work, the measurement models defined and validated (Chapter 6), then the structural parts shall be put together (Chapter 7) before any further analysis is undertaken.

5.4 Ontology and Causality within SEM

5.4.1 Ontology

The ontological positions that SEM takes are rarely discussed within applications of the method. It is left as an implicit assumption of the technique. This doesn't need to be the case and a proper understanding and acceptance of the assumptions for any one particular ontological position can add significant strength to the discussion of the implications of any model results. This is particularly pertinent for psychological measurement where the empirical system is not clearly defined and there is considerable scope to define the shape of the investigation amongst numerous possible constructs.

Borsboom (2005) attempts to answer the question of whether psychological measurements "really measure something and, if so, what?" claiming that "after a

century of theory and research on psychological test scores, for most... we still have no idea whether they really measure something or are no more than relatively arbitrary summations of item responses" (p2). Three different types of measurement are considered: true scores, scales and latent variable systems. The conclusions reached are scathing in terms of true scores, such as IQ tests, where the measured value is taken to be a true measure of the psychological concept, in the same way as metres are taken to represent height. This most widely used of psychological measurements is found to rest on a defective philosophical foundation allowing no true possibility for error. The analysis implies that representational scales, i.e. a process of assigning numbers in such a manner as preserve basic qualitative relations in the observed world, implements a constructivist philosophy[14] whilst latent variable theory rests on a realist ontology.

The approach to the latent variables debate is based on the point made in the introduction to SEM in the previous section, that in SEM the methodology works by specifying the formal (or abstract) structure that the system is assumed to exhibit and then this is separately tested against the empirical structure in the observed information. The emphasis is that there needs to be an inferential step connecting the theoretical construct to that which is seen in the observed data: theory needs to be connected to observation, and it is here that an ontological position is needed. The operational latent variable, that which is computed within the model, is merely a symptom of the observed variables (usually a weighted sum-score, with the weights determined via the model in question). Even if this operational definition of the latent variable is seen to fit the empirical data, this fit may equally apply to the output of a data-generating mechanism. Wood (1978) gives an example of a system of data generated randomly by the flipping of a coin that "fits more than adequately" (p27) with a Rasch model usually designed to test reading comprehension via frequency of mistakes (a special case of Item Response Theory, itself a subdivision of SEM).

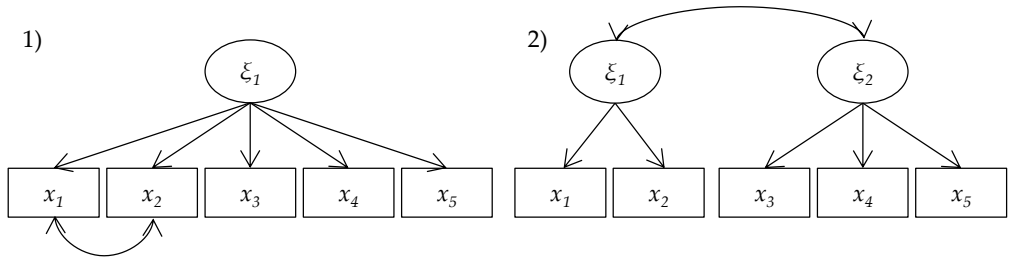
---

[14] A discussion of the arguments for true scores and scales is beyond the scope of this brief overview of the main arguments in relation to latent variables only. Although there are considerable interrelations between the argument put forward in each section, any reproduction would be just that and the interested reader is referred directly to Borsboom's highly readable, if philosophically dense, work.

The ontological position to be taken can either be realist or constructivist. That is, from a realist position, the latent psychological construct exists as an independent concept within the subject under examination and different item responses are manifestations of that one concept (with an additional aspect of error) or, as constructivists would have it, the latent measurement is 'imagined' and used merely to summarise the responses to a given set of items and aid interpretation[15]. The latter, constructivist view therefore means that the item responses define the latent variable structure and it is the fit of the formal system to the observed data that is of primary importance. Hershberger (1994 in Borsboom, 2005, p66) gives an example of two statistically equivalent models shown in Figure 5.5, the first with one latent variable and a covariance between the two items, the second with two latent variables. With a constructivist view, the fit of the model takes primary place in assessing the truth of the model, and as the two models are statistically equivalent, both of Hershberger's models would be suitable. Therefore, the correlation between the two latent variables in the second system could be estimated, but as there is nothing in the constructivist view to separate the choice of models except the fit, that there is only one latent variable is still a valid position. So a covariance between two latent variables is estimated whilst it is simultaneously held that there is only one latent variable! This lack of ability to decide which of the two constructs is 'true' cannot be determined within a constructivist framework. A realist ontology, however, could choose which fits based on the understanding that one or the other *may* be true, that "the simple reason that the formal theory implies that one could be wrong about the position of a given subject on the latent variable is *only possible on the assumption that there is a true position*" (Borsboom, 2005, p64, emphasis added) or, to quote Messick (1988), "one must be an ontological realist to be an epistemological fallibilist" (p26).

[15] Borsboom actually splits the latter into the more granulated constructivist and operationalist positions, but admits that the operationalist position is a special case of the constructivist and for this review, discussing just the broader of the two is not problematic.

Figure 5.5 Two empirically equivalent measurement models



This is not intended to imply that this is the only possible ontological position, merely this is that which is seen most commonly within structural equation modelling and that which is most logically sound. Bartholomew, Knott and Moustaki (2011) state that "the usefulness and validity of the methods… do not depend primarily on whether one adopts a realist or an instrumentalist view of latent variables. Whether one regards the latent variables as existing in some real world or merely as a means of thinking economically about complex relationships it is possible to use the method… *as if* the theory were dealing with real entities" (p3, original emphasis) but their argument implies an epistemological difference rather than one of ontology. The whole language of SEM is based around reflexive constructs and that the pattern of covariance between indicators can be explained by their regression onto an underlying construct, rather than the other way round. It is rare in the extreme to see a dialogue where the language implies that measurement of a concept, say shyness, where the items presented are thought of as a comprehensive set that define the concept. Rather it is the other way round, the items' responses are thought of as manifestation of that shyness. Of course there are other systems, especially in economics, that are formative concepts (such as socioeconomic status or area deprivation level), but for the individually perceived concepts to be defined in this work, these arguments imply that this realist ontology is that which is to be adopted explicitly rather than left to pass into the work on an implicit, hidden level.

5.4.2 Causality

An important corollary of this ontological position is that it helps constrain and clarify what constitutes a casual relation between latent variables and indicators in a standard measurement model, especially when considered as between-subject or within-subject.

A constructivist need not be concerned with this question, as to them the model is a trick for understanding theories, rather than a real relationship and thus if the model fits, the causation is immaterial. In a realist perception causality needs to be explored as the relationship between the observed and the latent, and that the latent variable is the cause for the responses follows directly from the realist position. The subject has a position on the latent variable, and this position precedes the response (from the realist assumption) and if two subjects occupy the same place on the latent variable then the response to an item should be the same (given measurement error).

The form that this causation takes, however, can be separated into between- and within-subject accounts. The former, a between-subject account, makes statements about population level differences based on between-individual variation of the position on the latent variable. This difference in position produces the variations between individuals on the item responses (plus any measurement error) hence the required variation for statistical modelling of the system. On the other hand, within-individual causation makes a fundamental misinterpretation of a measurement model as a process model, that a person must simultaneously hold a certain position on the latent variable (by the assumption of its existence) but also that they must have variation on that assumed position for the within-individual model to be constructible. The measurement model is not a mechanism that operates at the level of the individual, as the latent variable is an attribute that in most psychological situations should not change over a sufficiently short time frame, and therefore cannot also be the cause of the item responses. The answers are manifestations of the attribute. A latent model, therefore, should only be considered to apply between subjects - in keeping with the statistical maxim that between-subject conclusions should not be interpreted in a within-subject sense.

On a structural level, defining the structure of the relationship between these latent variables, the realist ontological framework can again be applied to help understand where assumption is made and how the latent constructs items relate to each other. The fundamental assumption for Mulaik (2006) and his discussion on causality in SEM is that "science is the knowledge of objects" (p106). From this realist ontological position,

enforced by the necessities in Borsboom's arguments, he outlines that a causal variable in terms of structural equation modelling is one which affects the probability distribution with which an effect variable occurs. This causal variable may determine some or all of the parameters of the distribution, so that these parameters themselves are functions of the causal variable. Mulaik (1986) clearly identifies that "variables in nature can be connected together in complex networks in which variables can be the combined effect of numerous independently acting causes and in turn the cause of numerous other variables... the simple model… which involves a simple sequence or chain of 'single connected' variables… is an inadequate representation of the complexities involved" (p91). Using the nomenclature of Pearl's (2000, p91) *Markovian parents*, Mulaik goes on to state that in the case that a single variable is the effect of a general number of causal variables, we may *"treat the causal variables jointly as a single variable"* (Mulaik, 2006, p91, emphasis added). Put another way, "structural equation modelling considers systems of variables wherein it is possible for more than one variable to be a cause of a given variable" (ibid., p87) and this set of variables can be considered jointly as one (causal) variable. An important pre-requisite is that the "assumption, often implicitly made, but should be explicitly evaluated for empirical validity when feasible, is that our models involve closed, self-contained systems of variables" (p87): closed in the sense that there are no other possible components (outside of the model) that may affect the position of the dependent. This empirical test is most often attempted by closing off extraneous causal variables, by shielding, holding constant, randomising or by isolating systems.

Within the social sciences it is particularly rare to have such closed systems and the empirical tests, controls or randomizations (see for instance Sampson's, 2012, discussion of causality in the Moving to Opportunity program) are at best an approximation. This is particularly the case when operating within the confines of survey responses, where these approximations of closure can be little more than assumption resting on the momentary truthfulness and independence of the respondent (perhaps a fair assumption given accurate survey design?). Whether a given response on a certain item is the result of the respondent's position on a certain (ontologically assumed to exist) latent variable hence falls to assumption and theoretical argument. Once the latent

variables have been defined, links between them to assess the relational structures posited by theory will need a theoretical argument made that asserts that the causal relation is of a certain direction, say from A to B, rather than a covariance between A and B. This theoretically based assumption shall be a common theme of the work and will be revisited regularly.

5.5 Fitting a SEM Measurement Model

The first step to SEM is to fit the measurement model. This takes the manifest variables and produces a model of the underlying latent variable(s). The first consideration of this step is whether a solution is identifiable, i.e. is it even mathematically possible to find a solution to the model that is theorised. The mathematical side of model identification is a complex topic and a full treatment requires technical and needless detail (more detail can be found in, amongst others, Kline, 2010). Broadly speaking, a model is identified if a unique solution can be found. If it is possible to find a unique solution it can therefore be estimated and, under certain conditions, this estimate can be tested for adequacy against observed data. A model can be under-, just- and over-identified. Just-identified models can be estimated but they have zero degrees of freedom and therefore cannot be statistically tested for fit, whilst under-identified models cannot be computed. The goal is therefore to work with over-identified models. SEM works with variance and covariance matrices, and the total number of variance/covariance elements between the manifest variables need to be greater than the number of free parameters to be computed. The number of variance/covariance data points in the matrix is $p(p+1)/2$, where p is the number of variables. When a model is analysed, the free parameters are the factor loadings, factor variance/covariances, path coefficients, residual variances/covariances and error variances. There need to be less of these in the model than there are in the original data (i.e. less than $p(p+1)/2$). In reality, this process is not as complicated as it sounds! SEM computer packages will only compute a result for identifiable models. Although a basic understanding is needed to notice that, say, a measurement model with just two manifest variables would be unidentified in general, the rest can be safely left to the black-box of the software.

Closely related, easier to understand and more important from an interpretational point of view than the technical complexity of identification is the concept of latent variable scaling, i.e. how the numbers relate to one another and what the latent variable is measured in terms of. Because latent variables are unobserved and therefore have no inherent scale, this needs to be fixed. There are two main ways of doing this. The first and most commonly applied (especially whilst modelling building) is to fix the loading (typically to 1) of one manifest variable for each of the latent variables. This in essence creates a reference variable and all other loadings onto that latent variable will thus be in terms of the scale of that reference variable. The second technique is to fix the variance of the latent variable to 1 and then freely estimate all the item loadings[16] giving a standardised result. This second methodology is enforced on this work as the original variables in the OCJS are attitudinal and categorical and hence have no scale. As latent variables are normally distributed, to discuss the latent variable in terms of standard deviation difference therefore is the only sensible way to proceed.

The categorical OCJS variables have implications not only for the latent variable scale, but also for model fit measures, the estimator that is used to calculate the parameters and subsequent interpretation. The theoretical assumption for categorical variable modelling is that underlying the given answers there is a continuous, normally distributed variable of which the categories represent only a crude measurement. In reality, the normality assumption of the unobserved variable is deeply restrictive (for example, the response pattern to the question "*How wrong is it for your age to write or spray paint on buildings?"* is 78% very wrong, 20% a little wrong and just 1.5% not wrong). This limitation is common and has led to the development of specific estimators that allow relaxation of the normality assumption, the foremost of these is the aforementioned (section 5.2) WLSMV estimator, only available in Mplus[17] (Byrne,

---

[16] A third technique proposed by Little et al. (2006), which they term "effects coding", allows one to estimate the latent parameters in a non-arbitrary metric that reflects the metric of the measured indicators. This nascent technique, however, has not been implemented in Mplus.

[17] Although STATA13 has increased functionality in terms of generalised SEM, only the asymptotic distribution free (also known as weighted least squares) is available to handle categorical answers (STATA13 SEM manual), which requires extremely large sample sizes and normally distributed underlying variables, and is comparatively inefficient.

2012). The WLSMV estimator has been shown to function with a wide range of missing data variants and is considered robust to missing data (Asparouhov and Muthén, 2010). This is the first impact of categorical data.

The second is in terms of the latent variable scale; the reference variable approach would fix the variance of the latent variable in terms of this underlying continuous variable, which is computed to have mean 0 and variance of 1. Hence the latent variable would be in the scale of something which itself is deduced within the modelling framework, i.e. the scale would be arbitrary (even meaningless). The focus should therefore be on the second method of scaling and standardised outputs, where the latent variable variance is fixed to 1. A corollary is therefore that interpretation of the factor-loading estimates is based on the squared standardised factor loadings. For example, a loading of 0.7 would square to 0.49, meaning that 49% of the variance in the (underlying continuous version of the categorical) original variable is explained by the latent variable. This, quite clearly, is analogous to the interpretation of the regression coefficient, r, and the r-squared in ordinary least squares regression.

Once an estimator has been chosen, the iteration procedure completed and a final model estimated it needs to be determined whether it suitably fits the data. This is achieved by way of fit indices that judge whether the theoretical structure that has been implied is sufficiently close to that observed in the data, and here lies the final implication of the categorical data. It has been shown that relying solely on one fit measure is inappropriate (Hu and Bentler, 1999; Marsh, Hau and Wen, 2004; Sivo et al., 2006; Barrett, 2007), and consequently all SEM packages incorporate a range of measures. The WLSMV estimator limits somewhat the choice of measures to be used and Mplus gives four different metrics by which to judge the model fit: the CFI (Bentler, 1990); the TLI (Tucker and Lewis, 1973); the RMSEA, which since the release of Mplus version 7 is accompanied by a confidence interval (Steiger and Lind, 1980); and, the WRMR (Yu, 2002). The CFI, TLI and the RMSEA are recommended (Byrne, 2012) as the core set of measures for all estimators, and with the WLSMV estimator the more usual supplementary AIC, BIC and SRMR is replace with the experimental WRMR.

The CFI is normalised (i.e. takes values from 0 to 1) with a value over 0.95 being indicative of good fit (although down to 0.9 also gives evidence of decent fit). The TLI, although not normalised (i.e. can take values over 1) is customary to interpret it in the same way as the CFI. A secondary property of this measure is that it penalises overly complex models, i.e. parameters that add little to the overall fit of the model would be detrimental to the value attained by this measure. Both the TLI and the CFI work by comparing the model to a baseline null model and assessing whether the implied model fits the data more suitably, and are known collectively as incremental fit indices. Conversely, the RMSEA (Root Mean Square Error of Approximation) and the related WRMR (Weighted Root Mean Square Residual) are absolute fit indices and determine how well the model matches the data's variance/covariance structure. Lower values indicate better fit. A value of under 0.05 for the RMSEA indicates good fit, but values up to 0.08 can be considered suitable, with over 0.1 showing poor fit. Mplus7 (released in September 2012) also gives a 90% confidence interval for the RMSEA and a probability that the value is less than 0.05. The WRMR is a variation of the continuous analogy SRMR (standardised root mean square residual) for categorical variables, is relatively experimental (Byrne, 2012) and should be treated with caution, but coupled with other measures showing good fit, a value of under 1 should be considered to be evidence of an extremely well-fitting structure. Table 5.1 summarises these fit index levels. It should be noted that each of these levels is an approximation, and a substantively meaningful model that passes two criteria whilst almost meeting the third is not a showstopper. It merely means the fit is adequate and further work may solve these issues. As with all social science there is a degree of 'close enough', especially when backed by suitably strong expectations.

Table 5.1 Summary table of fit criteria under WLSMV estimator

Index	Fit criteria
CFI	Above 0.95
TLI	Above 0.95
RMSEA	Less than 0.05 ideally but up to 0.08 can be acceptable
RMSEA confidence interval	a probability over 5% that it contains 0.05 is good, but the confidence interval should at least be close to having 0.05 as a lower bound
WRMR	Ideally under the value of 1, but this is experimental and to be used as a supplemental measure

Statistically speaking, the standardised loadings between manifest variables and latent components of a well-fitting model are rarely non-significant. SEM measurement produces a variance/covariance matrix independently of the data, and this is then compared to the observed data. As such, if a path coefficient (a loading) is statistically insignificant this represents an item uncorrelated with the supposed structure and hence would have been picked up in the fitting step. Substantively significant loadings, however, are a different story. As all loadings for debate in this work will be standardised, they range from -1 to 1 with values closer to either of these extremes representing the strongest association with the latent variable. What is considered a substantively significant is another area where considered judgement needs to be applied. With the interpretationally similar principal component analysis, Tabachnick and Fidell (2007) recommend suppressing values lower than 0.3 from a rotated solution, and this cut-off seems a relatively appropriate approximate rule in the SEM measurement context; however, all loadings need to be considered in context with each other and, again, theoretical expectations. If four out of five items have a loading of around 0.9, with the fifth at somewhere around 0.3 or 0.4, then this lower value perhaps deserves attention. Similarly – if all items are around 0.4 with one at a value approaching 1 (or indeed -1) the latent variable is probably being overly dominated by this latter component and again caution needs to be taken.

The final procedure to be discussed relates to model specification, misspecification and reformulation. This is the element where SEM can deviate from its ideal of being a completely confirmatory procedure, and can be carried out at two time points, either before the initial model or to adjust inadequate model fit, and is predominantly applied to the measurement models. For the first case, exploratory work before the model

formulation, this can be carried out via exploratory factor analysis which can give an approximate feel for how the variables group together, particularly useful when trying to replicate previously used measurement systems with slightly different, analogous variables.

At the second time point, to correct inadequate model fit, the model maybe re-specified slightly to account for substantively sensible possible reformulations. This can take the form of loading variables that had previously been placed onto one latent variable onto a suitable alternative. Similarly, reflecting on a slight misfit and a low factor loading, an observed variable may not match the latent variable as initially thought and may need to be deleted. Although correlation of error terms is generally frowned upon as it often means indicators are measuring something else, or something in addition, to the construct they are supposed to measure (Jöreskog and Sörbom, 1993, p113), if there is a likely response bias amongst a set of questions the errors within that set can be allowed to correlate. Mplus provides, upon request, modification indices (MI) which comprehensively list possible improvements to model fit that can be gained by allowing any item (manifest or latent) to load onto another/have correlated errors. Although in some instances these may help identify poorly fitting variables, or interrelations that would otherwise have gone unnoticed, these MI statistics must be treated with caution, as an increase in fit without reference to theory may be nonsensical. These steps must have a theoretical underpinning and altering a structure to gain adequate fit measures is inappropriate (Byrne, 2012). Any modifications made to an original model must be substantively meaningful and justifiable, it must be "nine-tenths theory driven and only one-tenth data driven" (Hayduk, 1987, p177) or as Little (2013) rather poetically puts it, "Much like the song of the Sirens, the allure of modification indices should be resisted. Many times, the suggested modifications are due to sampling variability and not to a 'real' process" (p156).

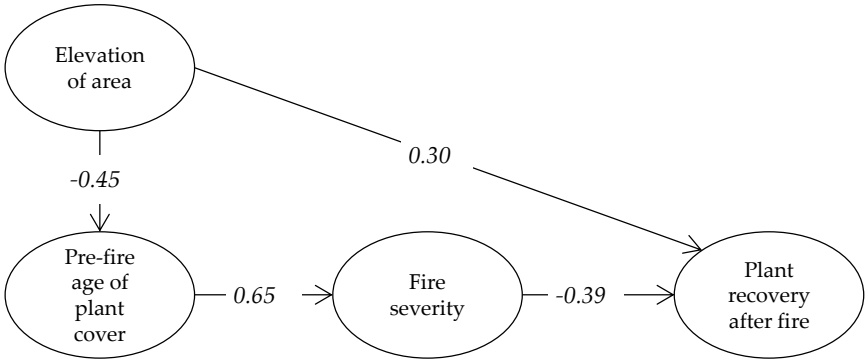
5.6 Moving to a Structural Model

Whereas Spearman pioneered the factor analytical techniques for the measurement part, Wright spearheaded the path analysis of the structural step and it is here that some of the real advantages of SEM over techniques such as multiple regression, even

when combined with prior factor analysis, become apparent. The structural part of the modelling allows item dependencies to be incorporated into the results of the analysis and this is where any theoretical perspectives on the structures and relationships in the system to be modelled are made analytically explicit and assessed in a rigorous way.

A simple way to look at the procedure is to say that it involves linking together latent (and if needed manifest) variables into a structural system via regression (single-headed arrows) or through covariance (double-headed arrows). Grace and Bollen (2005) outline an intuitive and simple real life system of modelling that details the advantages of a path analysis. The example is based on plant recovery speeds (as the dependent variable) after shrub land fires in Southern California, and how these are affected by elevation, fire severity and the age of the pre-fire growth. In a standard regression, linear or otherwise, each of the independents is assumed to be 'at the same level', i.e. they do not cause one another and their relationships are expressed solely through covariances. In a structural model, knowledge about interrelations can be incorporated. For example, in the original work upon which the Grace and Bollen paper is based, Keeley (1999) stated that they knew that the average age of the pre-fire plants was lower by 2.2 years for every 100m of altitude gained, and that the pre-fire age of plant cover (the density of the organic material) had a strong impact on the severity of the fire. These relationships form a causal chain, elevation causes younger plants to be present, and younger less dense plants in turn caused lower intensity fires, and this lower intensity leads to increased recovery speed. The severity of the fire could not possible affect the pre-fire plant cover and an area's pre-fire plant cover would not affect the elevation (unless it formed a particular high and unlikely mound of organic matter). This knowledge, however, was not used beyond descriptive analysis by Keeley. Grace and Bollen reformulated the modelling strategy and re-analysed the information using a SEM framework, with the results capturing the causal chain of reasoning (presented as a path diagram in Figure 5.6).

Figure 5.6 Standardised structural path model for plant recovery after fire, adapted from Grace and Bollen (2005)[18]



As with the original regression analysis plant recovery is regressed directly onto fire severity and elevation (the single-headed arrows pointing into the plant recovery latent variable on the right hand side of the diagram), but as pre-fire age was seen to only impact fire severity it was not directly linked to plant recovery, rather its effect was through fire severity. There are two paths from elevation to the dependent plant recovery, the direct and that via the causal chain as outlined in the previous paragraph. To fully understand how the role elevation of these components interacts with the plant recovery both the direct (elevation to recovery) and the indirect (via pre-fire plant age and fire severity) paths need to be considered. There is a direct effect and a *mediation* effect via the indirect chain; the sum is the total effect of elevation on plant recovery, controlling for other items in the model *and* their relationships. This methodology therefore elicits a clearer analytical understanding of the system than a regression approach.

In the above example the causal chains are rather clear, in a good deal of other SEM pieces the chains are much less easily defined. To quote Grace and Bollen: "While SEM permits the implications of a causally structured theory to be expressed, the analysis itself does not contribute to the establishment of causality. This must come from other information" (p288). This other information being theoretical consideration, based on

[18] Although in the original Grace and Bollen piece the model is discussed in terms of path diagrams, i.e., with manifest variables, the commentary makes it clear that the same considerations are needed when dealing with latent variable structuring. The model presented is adapted to consider that exact situation.

previous work and logical implication. A pertinent point here is the separation of endogenous and exogenous variables. As referred to in section 5.3, although it is not necessary to *label* these two types of model components differently, it is important to *consider* them as separate. An item at 'the beginning' of a causal chain is not influenced by any components within a model and is known as an exogenous variable (no single arrows pointing in or, equivalently, not regressed onto anything). As such, they represent the beginning block of the model and have direct effects on the ultimate dependent domain, and can be the start point of an indirect effect, via a (number of) mediating variables. The endogenous variables, on the other hand, are influenced by components within the model through a regression (have single-headed arrows pointing into them) and therefore, if not the ultimate dependent variable, take the form of mediating components (a full discussion of this shall be returned to in Chapter 7).

A further component of a structural model is covariance, which are represented by double headed arrows in Figure 5.2. These are typically present between any two exogenous or endogenous variables unless there is a reason to constrain them to zero. They represent simply the degree of relationship between the two variables (latent or otherwise), and their interpretation is clear. The main caution with regard to these model elements is one of the program defaults. The Mplus default is to set to zero the covariance between any two endogenous variables (i.e. those with a single-headed arrow pointing into them). Blind acceptance of this scenario is not valid and the constraint of each covariance must be considered.

Although a set of fit index levels to be attained to indicate adequate fit were outlined in the previous measurement section, when it comes to structural models these rules become looser and less concrete. The papers from which the measurement model rules were by and large derived were based on Monte Carlo simulations of misspecified measurement models, with certain loadings or covariances constrained to zero. The majority of the work follows the influential 15 indicators, three latent variables, one "true" and two under-parameterised misspecified models as found in the influential Hu and Bentler (1999) paper. In many cases the adherence to the strict rules in this work have been overstated. Before the 1999 work required levels for adequate fit were

more relaxed, and indeed this point is made in the Hu and Bentler piece. Marsh, Hau and Wen (2004) reiterate and restate this point clearly:

> "Quantitative social scientists, particularly psychologists, have long been engaged in the elusive search for universal 'golden rules'—guidelines that allow applied researchers to make objective interpretations of their data rather than being forced to defend subjective interpretations on the basis of substantive and methodological issues. Their appeal—like the mythical Golden Fleece, the search for the fountain of youth, and the quest for absolute truth and beauty—is seductive, but unlikely to be realized…. data interpretations and their defence is a subjective undertaking that requires researchers to immerse themselves in their data" (p321).

This is none more true when building structural models. Their increasing complexity means that the values attained by the fit measures with more than five factors or with 50 or so items rarely achieve what is considered 'conventional' fit (Marsh, Hau and Grayson, 2005). Overall, then, the fit indices, although still playing a role in the assessment of model suitability, should be paid less attention in the structural section than in the measurement set-up and a CFI/TLI above 0.9, and a RMSEA up to 0.08 means that work can still proceed with accuracy. Loadings of items (which should vary as little as possible from measurement models), paths coefficients and substantive meaning, however, remain vital.

5.7 Calculation of Total and Indirect Effects

Due to the categorical data present in the current study, the discussion in the measurement methodological section has already shown that these models are best defined in terms of standardised coefficients. Grace and Bollen warn against this as standardisation means that interpretations of path coefficients are then in terms of standard deviation changes. When the variable scales are interpretable this standardisation leads to more abstract measurements, although in the current work the lack of interpretability of the latent variables is a 'bridge already crossed'. Therefore, for the majority of the analysis in this work a path represented by a single-headed arrow (i.e. a regression coefficient) will be considered in terms of standard deviation changes of the arrow's origin. Returning to Figure 5.6, the coefficient of -0.39 between fire

severity and plant recovery would mean that for every standard deviation increase in fire severity the plant recovery would decrease by 0.39 of its standard deviation, holding pre-fire plant age and elevation constant.

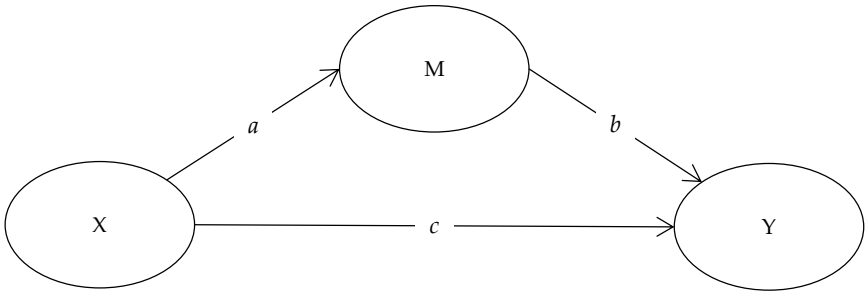
A similar interpretation would be seen for the effect of elevation; holding all else constant for a one standard deviation increase in elevation the plant recovery would increase by 0.30 of a standard deviation. But, of course, it has been seen that pre-fire plant age and fire severity are *not* constant as elevation changes, there is also an indirect mediating effect through pre-fire age and severity. The direct effect is still 0.30 but in this situation there also exists an indirect effect, calculated by taking the product of the secondary path, so (-0.45 * 0.65 *-0.39) = 0.11. The total effect is calculated as the direct effect plus the indirect effect giving 0.30 + (-0.45 * 0.65 *-0.39) = 0.41. The total effect, therefore, of a one standard deviation increase in elevation would give an increase in plant recovery of 0.41 standard deviations.

This method of calculation of indirect effects was traditionally approached by the Baron and Kenny (1986) method *causal steps strategy* and was formulated when there is just one mediating pathway, as represented by Figure 5.7. Here variable M is a mediator if X significantly accounts for variability in M, X significantly accounts for variability in Y and M significantly accounts for variability in Y when controlling for X. More recent approaches use the significance of the indirect effect, *ab*, as typically calculated via the Sobel test (Sobel, 1986) and compares the ratio of this value to its estimated standard error, computed via delta parameterisation. Mplus provides a simple method to calculate indirect effects with significance levels via a specific sub-command and although calculation of standard errors is a non-trivial matter, especially with non-normally distributed data (as in this case), it has been shown that these are accurate given sample size larger than 200 (MacKinnon et al., 2002). Despite this, Preacher and Hayes (2008) recommend using bootstrapping methods to estimate significance in the case of multiple mediation effects. However, Mackinnon stated[19] in the Mplus discussion forums that this methodology is not necessary with the WLSMV estimator

[19] See http://www.statmodel2.com/discussion/messages/11/266.html?1380580099 [accessed January 2015] for the complete discussion around the estimation of indirect effects in Mplus.

(as used here). Under these circumstances, with accurately computed significance levels, Preacher and Hayes (2008) see the introduction of multiple mediators as good practice, stating that "including several mediators in the same model is one way to pit competing theories against one another within a single model. Theory comparison is good scientific practice" (p881).

Figure 5.7 SEM with a single mediation effect



5.8 Interpretation of Model Results with a Dichotomous (or Categorical) Dependent

In this modelling work, the dependent to be used is self-report offending. This is usually a dichotomous yes/no in the previous 12 months or a categorical count variable (never, once, two or three, four or five and six or more). As such the regression coefficients to the dependent are not linear regressions and, due to the implementation of Mplus with the WLSMV estimator, is constrained to use a probit link function. Interpretations of coefficients are hence non-linear and are not interpretable in the same way as a linear regression ('a one unit change in X causes a change of β in Y') nor as a simple log odd ratio ('a one unit change in X causes a change in the log odds ratio in Y by β') as in the case of a logit link function. Rather, the interpretation is dependent not just on the change in whichever variable is allowed to alter but also on the values of the other variables in the equation, due to the nature of the cumulative distribution function (CDF) of the standard normal distribution. To be explicit and to use mathematical notation for one point in this thesis, in a latent variable modelling context for a dichotomous dependent Y, there is a continuous variable Y^* that determines the value of Y. It is assumed that:

$$Y_i^* = (\sum_{i=0}^{n} \beta_i X_i) + u_i$$

and, with u_i being a random disturbance term, that:

$$Y_i = \begin{cases} 1 \; if \; Y_i^* > 0 \\ 0 \; otherwise \end{cases}$$

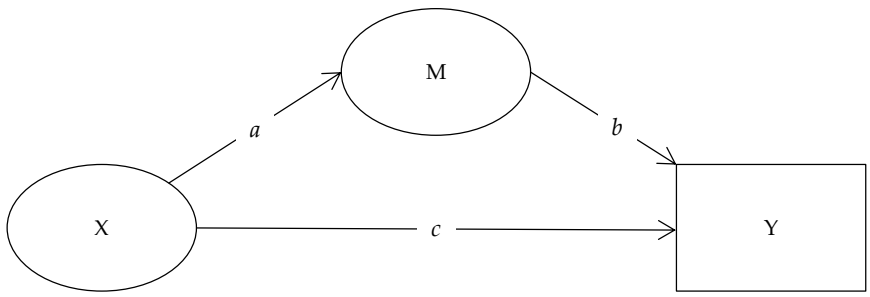
and therefore that it can be shown (Nagler, 1994) that:

$$Prob(Y = 1) = \Phi(\sum_{i=0}^{n} \beta_i X_i) \tag{1}$$

where Y is the dependent, X_i the independents, β_i the estimated coefficients, n the number of independents and Φ is the CDF of the normal distribution.

Page 492 of the Mplus version 7 manual (Muthén & Muthén, 1998-2012) gives a simple methodology for calculating predicted probabilities for either dichotomous or categorical variables involving the threshold value(s), the coefficients and the values on each of the dependents. When the latent variables involved in the calculation are standardised then the process becomes easier still and the discussion can be carried forward in terms of respondents at various different standard deviations from a score of zero (i.e. the mean value) on the exogenous variable(s). When continuous mediating variables are present, it is simply the total effect that is used as the coefficient in the calculation, as can be seen by considering Figure 5.8.

Figure 5.8 SEM with a single mediation effect and categorical dependent



Where Y is dichotomous (the similarity with Figure 5.7 is clear to see). As X and M are continuous and standardised, using (1) the equations for this system can be written:

$$M = aX + \varepsilon \tag{2}$$

$$Prob(Y = 1) = \Phi(-t + cX + bM) \tag{3}$$

Where t in equation (3) is the threshold value as in the methodology as outlined in the Mplus manual and $\varepsilon \backsim N(0, a)$ and there is no constant term in the equation (2) due to standardisation. Substituting (2) into (3) gives:

$$Prob(Y = 1) = \Phi(-t + (c + ab)X) \tag{4}$$

and therefore $c + ab$, the total effect, is the coefficient of interest by analogy to equation 1. The extension to this for a categorical dependent is similar and can be found in the Mplus user manual.

Interpreting the impact of mediating variables, via the indirect effects, is trickier and is rather underdeveloped in the SEM literature, as clearly stated by Muthén and Asparouhov (2014). There has however been more attention paid to this type of mediation in a clinical trials setting and is discussed in terms of counterfactual and potential outcomes based on randomised control trail settings (Robins and Greenland, 1992; MacKinnon and Dwyer, 1993; Pearl, 2001; Valeri and Van de Weele, 2013). Based on derivations of what are referred to as natural direct and indirect effects where the mediator, M, is continuous and with a binary X and Y, these papers show that the total effect can be considered as the sum of the direct and indirect effects and that the proportion of mediation can be calculated as the indirect effect divided by the total causal effect, despite the non-linear nature of the probit function. Muthén and Asparouhov (2014) run a Monte Carlo simulation to show that when the mediator variable is defined as a latent variable with 3 or 6 indicators, that this approximation holds true for sample sizes of 200. Therefore, despite the fact that the value that the *c+ab* term plays in the predicted probabilities is dependent on other values (and the threshold value), an approximation to compare the size of the effect for the direct component from an exogenous variable to the effect through the mediating variable, it

seems fair given the relatively large sample present in the current work to calculate this as the intuitively obvious formula:

$$\text{Proportion of mediation} = \frac{\text{Indirect effect}}{\text{Total effect}} \tag{5}$$

This will be used extensively in the current work and the explanation of its meaning will be returned to and explained with examples from the data. This, it is hoped, will clarify the meaning and make understanding easier rather than attempt to battle on in the abstract.

5.9 Concluding Remarks

This chapter is not meant as a replacement for a detailed methodological textbook or the like, but more as guidance of the techniques that shall be applied. In relation to SEM the most pertinent point is that the work can adequately and succinctly be expressed through the usage of path diagrams rather than the more complicated matrix algebra. It is this non-mathematical presentation that will be used throughout the thesis and with this style the work should be accessible to those without any formal mathematical training. Also of note are: the fit indices that are to be used for assessing measurement models; that these fit indices aren't as important once work has progressed to the structural modelling; some of the constraints on the modelling procedures in the specific case of the data that will be used in this thesis, namely standardised results and the usage of the WLSMV estimator (and hence Mplus) with both of these being enforced through the categorical nature of the manifest variables; and an outline of what is meant by an indirect pathway and how these should be interpreted using the 'proportion of mediation'.

The discussion on ontology and causation is important as it clarifies what is inherently being built into a model approached through SEM. The fact that it enforces a realist ontology does not mean that the work needs to be conducted as a strict positivist but that the language that is needed is of a particular type. The items that are used to measure a latent variable need to be described as such, not referred to the other way round as 'making the latent variable'. This will be shown to have impacts of the

definition of how the latent variables interact in Chapter 7. The causality discussion is intended to stress that SEM cannot answer questions which are fundamentally conceptual; to paraphrase Jackson and Kuha (forthcoming, 2015) "Can latent variable modelling really be used to test whether one has, in the first place, measured the 'correct' constituent parts…?... The answer to this question is clearly no. This remains a conceptual claim."

In conducting this work there will be occasions where more methodological discussion will be needed, particularly when dealing with specific effects that become apparent in the structural work (such as suppression effects) and when turning to the longitudinal modelling. This will be worked into the narrative when appropriate. Also, as work progresses through these sections, the explanation of the specifics of the modelling will gradually be reduced. So in Chapter 7, for instance, where predicted probabilities are first mentioned there will be an explanation of the correspondence of the probit result with the results of a logit model. The next time this is mentioned the explanation is reduced, until eventually it is taken as read that the nuances of the equivalence are understood. The grand hope being that the work is in some way didactic.

Chapter 6

Measurement of Self-report Risk Factors

6.1 Process and Aims

Statistics is usually concerned with the connection between inference and data. Measurement, on the other hand, is concerned with the connection between data and reality, and is the activity that produces the raw materials which statistical methods analyse (Hand, 2004). It occurs during the process of assigning numbers or symbols to represent an attribute (or attributes) of a subject, such that these numbers or symbols reflect the relationships of these attributes; a homomorphic movement from an empirical relational system to a numerical relational system. Measurement theory, as a branch of applied mathematics, is mostly concerned with how the empirical relationships between objects can be preserved under various mathematical transformations (for a mathematical approach to measurement theory see the *Foundations of Measurement* series; Krantz et al., 1971; Suppes et al., 1989; Luce et al., 1990). The field found prominence in psychology with the work of Stevens (1946), who cites a seven-year committee of the British Association for the Advancement of Science that debated the problem of the measurement of human sensation. After appointment in 1932, by 1938 they still had not reached agreement and their interim conclusion was general disagreement and a call for more time to consider the question! Stevens attempted to iron out the confusion by stating that measurement is fundamentally rooted in meaning and is therefore a semantic issue, but one that is susceptible to orderly discussion. He (Stevens, 1946, 1951) goes on to define the now well-known nominal/ordinal/interval/ratio levels of measurement, which went some way to disentangle (some of) the committee's semantic problems by separating the assignment of numerals into different realms. These differences occur dependent on the specific contexts which one is trying to measure, where the empirical relationship must be reflected in the numerical equivalent (for example, loudness, or as the committee

termed it "the subjective magnitude of an auditory sensation", could be described as "quiet" vs. "loud" though in a different context could be scaled along a decibel counter and referred to as volume).

Once the data collection process has been conducted via survey items or observational work (or the like) the discussion regarding the level of measurement and the connection between the empirical relational system and the numerical partner becomes somewhat moot. The data have been collected, the levels of measurement for each variable defined within the data collection design and the numerical codes that represent the opinions of the respondent have been collected and are unable to be changed (or least cannot be expanded); they have to be accepted. However, to which 'concepts' these questions relate is often taken with little consideration and it is not uncommon to see whole theses of work being constructed on the basis of one single question response, asked in amongst many thousands during the administration of a questionnaire; “a cursory inspection of the literature of social research will show that much of it centres on entities which are handled as if they were measureable quantities but for which no measurement instrument exists” (Bartholomew, Knott and Moustaki, 2011, p2). In many psychological situations direct measurement of an attribute in question is not possible, and one has to accept that a representation of the multi-dimensional data distribution represented by the concept can only be operationalised as a unidimensional item. In other words it is necessary "to acknowledge that the definition of the variable being measured must include some pragmatic component" (Hand, 2004, p89). This is not to mean that one should quickly accept the easiest measurement, disregarding complex constructs and taking the validity of a scale merely for convenience; a pragmatic approach does not mean a simplistic one.

Hand (2004), in his comprehensive attempt to join together the measurement literature (including contributions from mathematicians, philosophers, and the physical, behavioural, social and life sciences), when turning his attention to measurement in economics and the social sciences illustrates “the heavy, even dominant, role that pragmatic aspects take in much of the social and economic measurements. The objects and systems being described in such contexts are so complex that pure representational

measurements are rarely possible. Definition of the attribute being measured necessarily occurs at the same time as the measurement instrument is constructed" (p271). Although in a general sense this is true of all measurement in the social sciences, from league tables to retail price indices, crime rates and indices of multiple deprivation, it is also true where the social science object being measured is closer to the psychological world, where "the complex nature of personality means it cannot be measured without first building a model of it, and describing its various attributes" (p163).

Structural Equation Modelling (SEM) provides an excellent tool for exactly this type of psychological measurement procedure. As identified as far back as Bentler (1980), being able to measure a concept with multiple indicators means being able to triangulate and identify concepts with more reliability. The SEM methodological framework then incorporates the possibility of further regression-type analysis, linking concepts and checking interrelations, a strong platform for this work. As explained in Chapter 5, the first step is to create valid measurement models and this is the concern of the present chapter. The discussion in section 5.4.1 means that although the pragmatic considerations need to be considered as part of the measurement process, it does not mean that the resultant latent variable is merely a pragmatic summarisation of those manifest variables. It has to be accepted that the latent variable constructed is a true trait, that the respondent has the characteristics defined as a psychological construct. The realist ontology implies that the respondent has an underlying level of, say, impulsivity or self-control and this underlying position manifests itself as the question responses. This acceptance of the trait allows the interrelations to be more clearly identified and understood.

The data from the OCJS were not collected specifically with this research in mind; rather they were collected as a general attitudinal and self-report offending survey. The survey was designed in partnership with a panel of academics whose research interests span across the criminological sphere (OCJS technical report, 2003) and their input into the questionnaire design will have added questions that refer to their own extensive works. The restrictions placed upon the survey by cost and respondent burden

considerations mean the question sets will form a subset of those which will have been used in previous works and therefore the task to produce valid measures will be a parsimonious task. As outlined in Chapter 2 these components fall into three broad categories (individual, family and 'higher') and this work shall attempt to produce accurate measures that 'borrow' from developmental and life course criminology (and, in Chapter 7, incorporate them into an 'atheoretical'[20] structure). The measurement process will be similar for each component. There now follows a brief review of the aims of the measurement with more methodological details presented in the relevant subsection later in the chapter.

Individual level

These are succinctly reviewed by Jolliffe and Farrington (2010) (for impulsivity, empathy and intelligence) and by Farrington and Welsh (2007), amongst others. Where no explicit reference is made the default is to those texts, but at an individual level risk factors should include, where possible:

- A measure of intelligence (or IQ, which can be defined as verbal or non-verbal) or (school) attainment, although previous work has typically measured intelligence at rather young ages (under 7), while some works focusing on attainment have looked at late childhood/early adolescent attainment.
- Personality - usually the three domains of extraversion, neuroticism and psychoticism, or the 'big five' personality types. Other work has touched on agreeableness and/or conscientiousness.
- Temperament - irritability, low amenability and adaptability
- Empathy - most often cognitive (understanding/appreciating) empathy, but also operationalised as 'emotional empathy'
- Impulsiveness and related behaviours - this is often seen as one of the strongest predictors, capturing a whole array of slightly differently named behaviours

[20] In the sense of Farrington's pre-ICAP days. To quote: "For many years I did not attempt to formulate a wide-ranging [DLC] theory of offending… I focused on identifying independently predictive risk-factors, testing specific hypotheses, and on investigating possible causal mechanisms intervening between risk factors and offending."(Farrington, 2005, p76)

from impulsiveness, hyperactivity, restlessness, clumsiness, no consideration of actions, poor ability to plan ahead, short time horizons, sensation seeking, risk-taking and poor ability to delay gratification.

- Self-control is characterised as a manifestation of impulsivity by Farrington and Welsh (2007) but deserves a separate and specific mention due to the centrality it takes in the work of Gottfredson and Hirschi (1990).
- Wikström (2012) centres his situational action theory around morality (along with self-control), although how separate this is from impulsiveness or self-control will be discussed in section 6.2.4.

Family level

Again, these are reviewed extensively elsewhere, notably by Blackburn (1993) and by Farrington and Welsh (2007):

- A measure of parental discipline is needed - not only has it consistently shown predictive power over offending behaviour, but it is also perceived by the public to be the most relevant (Farrington and Welsh, 2007). Erratic or inconsistent discipline is a related alternative.
- Parental reinforcement of behaviours (either poor or good, dependent on the normative framework employed)
- Parent/child relations - this often has a strong crossover with discipline, although is broader (including supervision, parental involvement and parental warmth).
- Child abuse and neglect
- Parental conflict or inter-parental violence
- Disrupted family - quantitative work to this has tended to focus on the loss of a parent (usually the father)
- Large family sizes
- Criminogenic families - i.e. families which have a history of criminal involvement

Higher level

Of greatest analytical importance to the current work are a set of measures of area level factors, and previous work has included:

- Collective efficacy - comprised of social cohesion/trust and informal social control
- 'Broken windows' or neighbourhood condition
- Fear of (local) crime

Other 'higher level' factors in previous work have been:

- School influences - broadly these are usually seen as between schools, although the relationship has been grouped to consider characteristics such as levels of trust, organisation structures/practices or discipline regimes within the school environment
- Peer influences
- Socioeconomic status (SES) of the subject, which is typically a measure of family situation but can also be the deprivation of the area

The peer and school higher level factors that are not directly related to area, and will not be the focus of this thesis, are left aside for future work. The evidence base is mixed for each of these higher level factors and to try to incorporate all in one step would muddy the waters too much to be able to pick out effects just of those of highest interest. In the current chapter only individual, family and the three area level factors noted above shall be defined. SES of the area and other structural characteristics of the area will be the focus of Chapter 9. The fourth domain that as yet has not been mentioned here is the dependent self-report offending. The validity of the measure was discussed in section 4.6 and further discussion will be returned to in the structural chapters.

The goal of this work is to try to model offending risk factors (with a focus on area level data) using an existing non-specific dataset, the OCJS, and the process for the measurement side of these components is to be a parsimonious, pragmatic exercise,

with close attention paid to previous work in these areas. The first step at each level is to review the available variables in the dataset and then take pragmatic steps to try to define factors based on the face validity of those available items. The majority of these will be via SEM although some correlates will be directly measurable (such as parental divorce or time spent in care homes). The fit of the latent concepts will be considered along with their substantive meanings, as will the reliability of any directly measureable components. These structures will then be considered in terms of previous operationalisations of concepts and their theoretical definitions with any similarities/differences highlighted. This will involve a good deal of assumption and often the components' validity will rest as much on face validity as on methodological steps. This too will be discussed. All work will take place on the constructed 2004 to 2006 longitudinal dataset, details of which can be found Chapter 4 (although checks as to the measurement equivalent across the years will be a corollary of Chapter 10 and will be reported in full there).

6.2 Defining Predictors at the Level of the Individual

6.2.1 The Available Data

"One of the primary difficulties in areas such as psychological measurement is that the empirical system being modelled is often not well defined. This means that there is a considerable freedom in choosing the pragmatic constraints: one can choose to include different questions in a test and to combine those questions in different ways" (Hand, 2004, p153). Measurement of these individual level factors is probably the most complex task that is attempted here. These are essentially psychological constructs and "in psychology... essentially all variables are related and the trick is to tease apart this complex tangle" (ibid, p152). Thus, the pragmatic will take a particularly key role in this task. After consideration of the available items in the OCJS the possible constructs will be considered alongside some of the extensive and often highly specific (in the sense that papers often concentrate on just one dimension and do not judge the correlation with other factors) previous work.

The questions that fit in with what has been referred to as the individual level risk factors are predominantly contained within the *health, lifestyle & risk factors* section of the questionnaire, with some questions asked in the *attitudes to the criminal justice system*. The *health, lifestyle & risk factors* section was asked via CASI (computer-assisted self-interviewing) whilst the *attitudes to the criminal justice system* was asked via CAPI (computer-assisted personal interviewing). The questions found to be relevant are displayed in Table 6.1 (all questions were recoded where necessary so that they all scale in the same direction).

Table 6.1 Individual level questions

Code	Question	Responses categories	Coding	Valid in 2004 sweep
i1	It is OK to steal something if you are very poor	5 pt Likert	1 = strongly disagree up to 5 = strongly agree, inc. neutral category	99.7%
i2	OK to steal from somebody rich who can afford to replace it	5 pt Likert	1 = strongly disagree up to 5 = strongly agree, inc. neutral category	99.8%
i3	OK to steal something from a shop that makes a lot of money	5 pt Likert	1 = strongly disagree up to 5 = strongly agree, inc. neutral category	99.9%
i4	Sometimes OK to break the law	5 pt Likert	1 = strongly disagree up to 5 = strongly agree, inc. neutral category	99.8%
i5	How wrong is it for your age to start a fight with someone	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	99.7%
i6	How wrong is it for your age to write or spray paint on building	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	99.7%
i7	What about someone your age playing truant, skipping school	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	96.8%
i8	What about someone of your age smoking cannabis	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	98.6%
i9	I usually do what I am told	4 pt Likert	1 = agree strongly, 2 = agree slightly, 3 disagree slightly, 4 = disagree strongly	98.4%
i10	I worry a lot	4 pt Likert	1 = disagree strongly, 2 = disagree slightly, 3 = agree slightly, 4 = agree strongly	99.6%

Code	Question	Responses categories	Coding	Valid in 2004 sweep
i11	I am usually helpful towards others	4 pt Likert	1 = agree strongly, 2 = agree slightly, 3 disagree slightly, 4 = disagree strongly	98.4%
i12	I am easily bored and find it hard to concentrate	4 pt Likert	1 = disagree strongly, 2 = disagree slightly, 3 = agree slightly, 4 = agree strongly	97.4%
i13	I get upset if I see people who are sad or hurt	4 pt Likert	1 = agree strongly, 2 = agree slightly, 3 disagree slightly, 4 = disagree strongly	95.4%
i14	I am often unhappy	4 pt Likert	1 = disagree strongly, 2 = disagree slightly, 3 = agree slightly, 4 = agree strongly	97.3%

A first glance at the questions gives some guidance as to what may be possible. There is a clear set of questions that seem to be examining some form of moral dimension - i1 to i4 are clearly related to the respondent's opinion as to breaking a general law, or to borrow from the lexicon of Wikström et al. (2012), "a moral rule is a rule of conduct that states what is the right or wrong thing to do... The law is a moral rule" (p12). i5 to i8 give a similar first impression, though in this case they are constrained into an age-specific realm. Questions i9 to i14 are less clear; i11 and i13 look to be related to some dimension of empathy, directed at the respondent's feelings towards others, whilst i10 and i14 appear to be akin to contentment, with i9 and i12 relating to some level of self-control. A different reading could imply i9, i11 and i13 as an empathetic construct, whilst i10, i12 and i14 could be some level of self-control *or* contentment combined. Overall, this initial appraisal of the variables leads to the expected discovery of between 3 and 5 latent variables, around morals, self-control, perhaps empathy and the possibility of life contentment/ability to control one's self. This tangle of possibilities will form the basis of the later theoretical discussion.

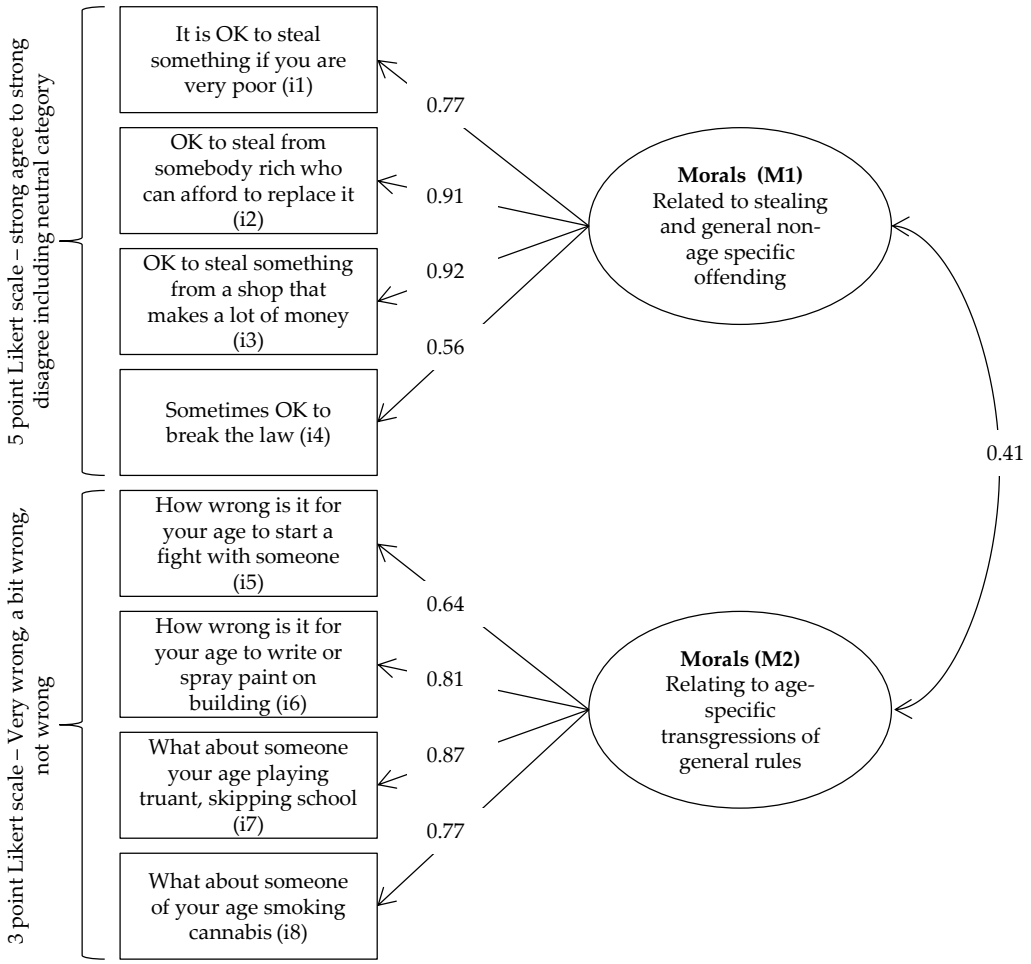
6.2.2 Morals

Building up from the bottom and working solely with the eight items that seem to relate to morals, i1 to i8, and running a 1 latent variable model (i.e. analysing whether all item response are determined from one psychological construct), gives a badly

fitting solution[21] (CFI = 0.891, TLI = 0.848, RMSEA = 0.204). Acknowledging the slightly different wording of the questions, namely that i1 to i4 appear to be specifically looking at opinions on breaking the law whilst i5 to i8 are age-specific transgressions, and fitting a two-factor solution as shown in Figure 6.1, leads to much better fit (CFI = 0.981, TLI = 0.973, RMSEA = 0.086 with 90% confidence interval 0.078 to 0.095 and WRMR = 2.148). This represents a good fit, with both the CFI and TLI of a suitable level (above 0.95). The RMSEA is slightly too high, with a 95% confidence interval not containing the hoped for level of 0.05, however as stated in section 5.5a level of 0.08 can be considered reasonably acceptable given suitable theoretical underpinnings. The WRMR, although slightly experimental, should be close to 1; at 2.1 this measure fails. Modification indices imply that allowing i4 *Sometimes OK to break the law* to load onto the second factor, the age-specific moral dimension, would improve the fit (as it does, CFI = 0.994, TLI = 0.990, RMSEA = 0.052 with 95% confidence interval 0.044 to 0.061 and WRMR = 1.269) but at this stage this step seems unnecessarily complicated given established good fit. The standardised factor loadings for each item are suitably high with the exception perhaps of i4, which at 0.56 is the lowest loading and reflects the slightly ambiguous nature of this variable and its correlation with M2. The covariance between the two moral dimensions, of 0.412, implies that the two factors have a moderately strong relationship with each other.

[21] To recap fit statistics and the appropriate levels, please see section 5.5.

Figure 6.1 Moral latent variable loading diagram



6.2.3 Empathy, Self-control and Contentment

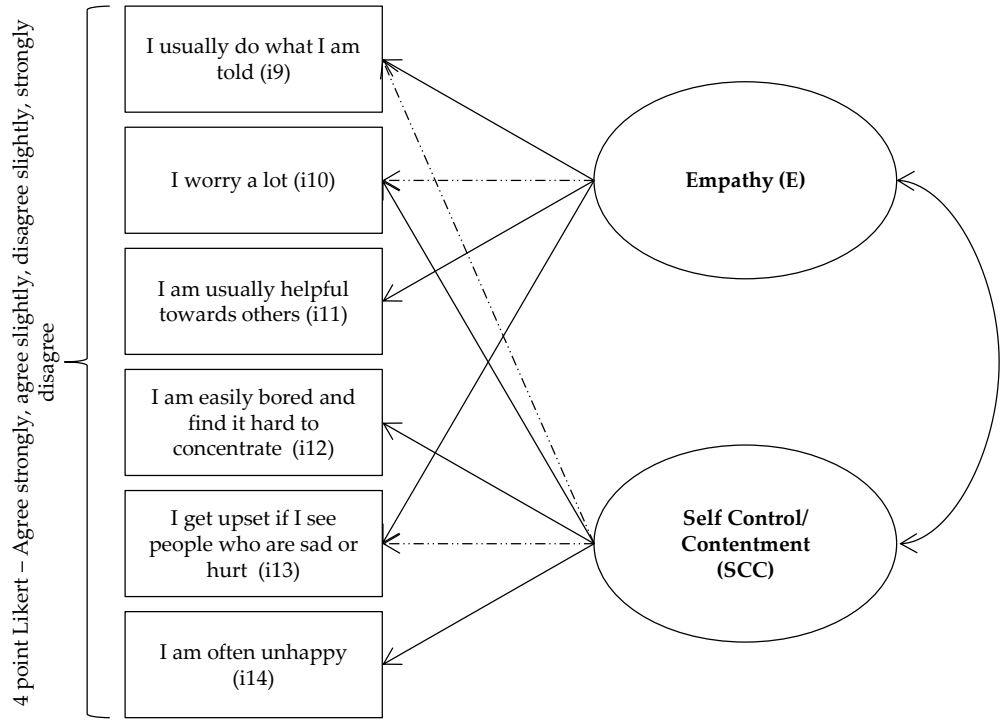
Approaching the remaining six items is difficult and perhaps the least clearly defined reduction in this work. The preliminary sense check seemed to point towards the possibility of a three factor solution. Unfortunately, with only six items this would not be identifiable. In this case the use of exploratory factor analysis, although slightly against the spirit of structural equation modelling and its theoretically based hypothesis testing methodology, gives a way to get an initial handle on the underlying structure (Asparouhov and Muthén, 2009). Using the same fit criteria as with a confirmatory approach, a 1 factor solution does not fit (CFI = 0.497, TLI = 0.161, RMSEA = 0.233). Looking instead at a two factor fit gives an excellent fit with CFI = 0.991, TLI = 0.967, RMSEA = 0.046 (giving a probability that the RMSEA is less than 0.05 of 60%) and a

WRMR of 0.018. The oblique rotated factor loadings are given in Table 6.2. The clear loading of i9, i11 and i13 onto one factor with i10, i12 and i14 onto the second lead to an interpretation closely fitting with the second scenario of an empathetic construct and some level of self-control or contentment. However, there are relatively significant cross-loadings, notably of i10 *I worry a lot* on factor 1 (in the 'opposite direction', indicated by the negative loading), i13 *I get upset if I see people who are sad or hurt* on the second and to a lesser extent i9 *I usually do what I am told*. The implied structure is displayed in Figure 6.2, where the dotted lines represent possible cross-loadings to be debated.

Table 6.2 Exploratory two factor rotated loadings for items i9 to i14

Item	Factor 1	Factor 2
i9	0.624	0.195
i10	-0.416	0.629
i11	0.810	0.002
i12	0.165	0.517
i13	0.512	-0.290
i14	0.010	0.659

Figure 6.2 Empathy and self-control possible model specifications



Turning now to a more deductive approach, there are four different models to be considered. The first is represented by the solid lines in Figure 6.2 and then three further models where each of the dotted lines is added in order given by the size of the loading in the exploratory factor analysis: that is i10 on empathy, i13 on SCC and finally i9 on SCC. In carrying out this analysis the process must be driven by empirical fit coupled with substantive meaning. Table 6.3 gives a summary of the fit statistics, the factor loadings and the covariance between the two latent concepts.

Table 6.3 Empathy and Self-control/contentment model possibility results

		Model 1	Model 2	Model 3	Model 4
			Fit indices		
CFI		0.696	0.866	0.952	0.975
TLI		0.429	0.714	0.880	0.924
RMSEA		0.192	0.136	0.088	0.07
Prob RMSEA less than 0.05		0.000	0.000	0.000	0.017
WRMR		4.284	2.77	1.581	1.134
			Correlation		
SCC with E		0.131	0.362	0.461	0.344
			Loadings		
i9	on E	0.588	0.649	0.671	0.588
i11	on E	0.947	0.83	0.808	0.821
i13	on E	0.369	0.421	0.586	0.554
i10	on SCC	0.400	0.697	0.754	0.662
i12	on SCC	0.527	0.579	0.568	0.576
i14	on SCC	0.781	0.628	0.642	0.638
i10	on E		-0.558	-0.563	-0.512
i13	on SCC			-0.369	-0.303
i9	on SCC				0.204

Taking both the exploratory approach and the model building results together, it is clear that although there is certainly an adequate two factor solution available from these data, there is considerable interrelation between these two latent constructs.

Model 1, where no items are allowed to load onto both factors, does not fit the data and it is only with the introduction of the cross-loadings that an adequate solution starts to emerge. Model 4, by the requirements outlined in section 5.5, is the candidate that fits best, indeed the *only* formulation that fits and even then only to a moderately good level. Reading through the substantive implications of the loadings makes intuitive sense and with fit indices such as seen here, this is a vital consideration for any formulation to be considered acceptable. In particular, i9 *I usually do what I am told* and i13 *I get upset if I see people who are sad or hurt* both load positively onto what has been termed the empathy factor, and load alternatively positively (i9) and negatively (i13) on the self-control/contentment item. So an adolescent who is more likely to do as they are told (i9) has a greater level of empathy, an understanding of the reason why they should do what is told of them, and a greater level of self-control (which in this case is expressing itself as obedience). With respect to i13 (*I get upset if I see people who are sad or hurt*), this item's loading implies that the higher the level of empathy the more an adolescent feels for those who are hurt or unhappy, whilst a respondent with higher self-control will appear more restrained in expressing their feelings when they see someone in that position, explaining the alternatively positive and negative loadings. The item i10 *I worry a lot* is similarly sensible - a person with higher self-control/contentment is less likely to report worrying (a natural finding) and those with higher empathy are more likely to worry about their own position, demonstrating an understanding that the positions that others find themselves in could be visited upon themselves, a particularly pertinent finding considering the direction of the empathy loadings. The remaining items, i10, i11 and i12 are left to load onto one latent factor each and it is clear that the relationships are sensible once the direction of the coding is taken into consideration.

Overall this formulation implies that a fictional adolescent who has a higher level of empathy is likely to:

- answer that they are more likely to do what they are told,
- be more likely to help others if needed,
- be likely to say that they get upset when they see others who are sad or hurt and

- be more likely to be worried about their own position.

With regards to the self-control/contentment, someone with higher levels of this latent construct will:

- worry less,
- find it easier to concentrate and be less likely to report getting bored,
- be less likely to report being unhappy,
- usually do as they're told and
- will be able to contain themselves when they see someone who is sad or hurt and be less likely to report being affected themselves.

And lastly there is a relatively strong relationship between the two concepts; someone with higher empathy is likely to have higher self-control or self-contentment (depending on the nomenclature attached to this latter factor). It should be noted that due to the way these items are coded and entered into the model, a higher manifest variable response is implied by a *lower* level of empathy or a *lower* level of self-control and this will be important in the structural work.

6.2.4 Individual Level Factor Theoretical Consideration

So these definitions have good face validity, with a seemingly sensible interpretation based on working through their implications with regards to a fictional respondent. How, though, do they compare to work that has previously been conducted on this topic? Unfortunately, the evidence base for an understanding of empathy is rather weak, as noted by Jolliffe and Farrington (2004). The paper acknowledges that there is widespread belief that low empathy is an important risk factor for offending behaviour, stating that "empathy and the acquisition of empathy are considered essential components of adequate moral development" (Jolliffe and Farrington, 2006, p591). They conducted a systematic review of various empathy scales, identifying five different scales that each contain over 20 items (and up to 64) that purportedly measure empathy (with three in general use). They are rather scathing about the utility of these scales, claiming that two equate sympathy with empathy and the third, the Hogan Empathy Scale (HES) (Hogan, 1969), was merely created by the convenience that it showed

discriminatory power for offenders versus non-offenders and face validity of the construct implied that it most probably did not measure an empathetic dimension at all.

Jolliffe and Farrington (2006) move on to develop a 20 item, two factor (affective and cognitive) test they call the 'Basic Empathy Scale' (BES) with examples such as *"I get caught up in other people's feelings easily"* (affective) and *"I can often understand how people are feeling even before they tell me"* (cognitive). The story they tell is that as a concept empathy cannot, and historically has not, been well defined. Nor is it separate from a range of other psychological components, harking back to the point made by Hand (2004) and referred to earlier, that "in psychology... essentially all variables are related". The 2006 paper makes a number of predictions as to the relationship between (their scale's definition of) empathy and other predictors of offending, such as that females should have higher empathy (and lower offending rates), a strong correlation with sympathy and perspective taking should be apparent, significant positive correlations with the personality clusters agreeableness, openness and conscientiousness, and negative correlations with neuroticism and extraversion. Finally, they expect to see a negative relationship between their definition of empathy with the 12-item UPPS Impulsivity Scale (Whiteside and Lynam, 2001) i.e. more empathy will be found in people with less impulsivity. Overall they find evidence in support of the majority of their predictions, although they find no relationship at all between impulsivity and empathy in males, attributing this to a methodological problem rather than a true symptom of their work and calling for further analysis. Miller et al. (2003) validated the Whiteside and Lynam scale with 20- and 21-year-olds, finding a correlation of 0.50 between the premeditation and perseverance dimensions and correlations with crime and delinquency, aggression, substance use and risky sexual activity across all four facets.

This expectation of a correlation between empathy and impulsivity is a natural position to take, given the rather inclusive nature of what is meant by 'impulsivity'. The Whiteside and Lynam (2001) work found impulsive-like behaviour manifested itself in four distinct but closely related dimensions which they labelled urgency, (lack of) premeditation, (lack of) perseverance and sensation seeking. The measurement problem

they describe when it comes to defining impulsivity is that it is often on the flip side of self-control, or at least certain components of it could be: "each of the four components is related to distinct aspects of personality ... (lack of) premeditation and (lack of) perseverance are related to conscientiousness, urgency is associated with neuroticism, and sensation seeking is a component of extraversion" (ibid, p686). The lack of premeditation and perseverance, they continue, results from a dysfunction in the decision-making process, a lack, therefore, of self-control: as Vollmer et al. (1999) put it, "impulsive behaviour occurs when responding produces the more immediate, relatively smaller reinforcers at the expense of delayed larger reinforcers. Self-control occurs when responding produces delayed larger reinforcers at the expense of immediate smaller reinforcers" (p451).

Although there are clear crossovers between impulsivity and self-control, they are not exactly the same and self-control deserves special attention as to its measurement, not least due to its central importance in the influential work of Gottfredson and Hirschi's (1990) *General Theory of Crime*. Their definition of self-control is dealt with via descriptions of criminal acts and according to Grasmick et al. (1993) these fall into six characteristics, only one of which is equated directly with impulsivity. The others, a lack of diligence, risk seeking, a preference for physical activity, self-centeredness and minimal tolerance for frustration, are seen as distinct. Grasmick et al. developed a question set of 24 items (four per dimension) including "*I'm not very sympathetic to other people when they are having problems*", "*I often act on the spur of the moment*", "*I lose my temper easily*" and "*when things get complicated, I tend to quit or withdraw*", and through factor analysis showed that the scale can be considered to be unidimensional, i.e., all six characteristics are indeed indicative of one overarching concept, self-control. This scale has been widely adopted with the meta-analysis conducted by Pratt and Cullen (2000) finding it to be "perhaps the most carefully designed and valid measure of self-control" (p943). The Peterborough Adolescent and Young Adult Development Study (PADS+) adopted a modified and limited version[22] of the scale and use it as a key indicator of

[22] Physical activity, simple tasks and self-centred components were omitted due to perceived lack of relevance to Situational Action Theory (Wikström, 2012, p 136, footnote 20).

crime propensity (Wikström et al., 2012). Despite the almost universal uptake of this scale as the measure of self-control, there was an early element of contention as to whether the Grasmick et al. scale is in fact unidimensional. Wood, Pfefferbaum, and Arneklev (1993) found that individual components predicted criminal action differently and that the constituent characteristics could form individual components. Cochran et al. (1998) similarly argue that the physicality element distracts from the unidimensionality. These criticisms are very much the exception and each of the works find some validity in the scale, especially the side most closely related to impulsivity. Overall, self-control should be considered a rather diffuse trait that can be accessed via a number of differing methodologies.

In addition to the self-control component of an individual's crime propensity, the Peterborough Adolescent and Young Adult Development Study (PADS+) research also utilises a range of 14 'morality' questions of progressively increasing seriousness, from *"How wrong is it for someone your age to... steal a pencil"* through to "…*paint graffiti on a wall or house"* and "…*steal from a shop"*. These form a cornerstone of Wikström's Situational Action Theory and are the basis of what is termed 'crime propensity'[23], containing both an element of morality and self-control which Wikström et al. carry forward to underpin their work. Perhaps this combination is not surprising; although Grasmick et al. identify six dimensions of low self-control, they acknowledge that they omit three statements originally made by Gottfredson and Hirschi (1990) which they consider to be consequences of low self-control, rather than definitional components. These statements are that people lacking self-control "need not possess or value cognitive or academic skills" (p89), "tend to have unstable marriages, friendships and job profiles" (p89) and crucially, "tend to pursue immediate pleasures that are not criminal: they will tend to smoke, drink, use drugs, gamble, have children out of wedlock and engage in illicit sex acts" (p90). It is this last statement in particular, the reckless behaviour, which connects to the moral dimension and means that Wikström's combination of the two measures forms a natural partnership.

---

[23] Section 2.6 noted criticism of the morality dimension of the theory by Felson (2014) but this centred around the morality of the place/situation and not on the morality measures included in the criminal propensity measure.

6.2.5 Bringing it together to Individual Attitude

It is at this point that the empathy, the self-control and the morality measures start to converge. It has already been noted that Jolliffe and Farrington (2004) saw empathy as a precursor to adequate moral development, and the cross between empathy and impulsivity is expected due to the very definition of impulsivity, and impulsivity and self-control have been shown to be if not the flip side of the same coin, at least different sides of the same die. Thus, there is a story of a closely related and often overlapping melee of individual psychological components that have been seen to be related and predict offending or deviancy (as Delisi and Vaughn, 2007 argue). As the goal of this work is to incorporate these individual constructs to allow an analytical control along with family and area level correlates of crime, it is therefore sensible to re-specify the modelling of the empirical components in the current work into a unified framework. This re-specification can be seen in Figure 6.3 and incorporates a second-order latent variable, named *Individual Attitude*, which is formed of the lower level latent factors. The prescription fits the data very well; the CFI is well above the 0.95 that signifies good fit at 0.978, the TLI likewise at 0.972. The RMSEA which should ideally be under, or at least as close to 0.05 as possible, is valued at 0.052, with 18.6% probability of being lower than 0.05 (the 90% confidence interval is narrow, between 0.048 to 0.057, indicating a stable solution). The slightly experimental WRMR, a supplementary measure, is the only judgement criteria that does not meet the level set out in section 5.5 at 1.782 (though this does not invalidate the identification of the solution as well fitting). This specification is more than just theoretically functional, it is analytically sensible too; the modification indices of the moral dimensions indicated that cross-loading would improve the fit and there are already numerous crosses with empathy and self-control/contentment dimension.

As is usual with SEM measurement models all loadings are significant at a very high level ($p < 0.001$) and are displayed on Figure 6.3. The interpretation of the first order latent variables and their manifest indicators are rather straightforward for the two moral items, the lower the moral level on either M1 or M2, the more likely the respondent is to answer the questions in a less-favourable manner. The relatively lower

loading of i4 *sometimes okay to break the law* onto the M1 factor is perhaps indicative of the wording of the alternative three questions associated with this latent factor, which all include *stealing* as part of their definition. Furthermore, the two items with the highest loading relate to theft where the victim is perceived to be able to replace the stolen item, perhaps relating to a more specific 'victim-less' offence than the broader 'all law' possibility. The second moral item, M2, has high loadings on all items. None of these depart significantly from the pattern seen in the previous section dealing specifically with the moral components separately. Similarly the loadings of the empathy and self-controls have changed little from the previous specification and the interpretation previously given in terms of a fictional respondent has not changed. This stability of the loading when combining the two previous disparate components adds weight to the assessment of good fit, as the combination of items are affecting only their latent variables, and not distorting across the structure.

The loadings and the relative pattern of the loadings of the first-order latent variables onto the higher second-order latent variable are important. M2, the second moral construct representing age-specific transgressions of general rules, has the highest loading and hence will be the dominant component of the *individual attitude* latent variable. This is reassuring on two levels; firstly, the manifest items of this component have arguably the highest face validity in terms of the individual's immediate (meaning age-specific) willingness to break rules. The items in M1 are defined in terms of more serious offences (stealing or breaking the law in general) and therefore likely subject to more of a social desirability bias (Mills et al., 2003). The M2 items, however, are age specific and more closely related to those used by Wikström and more likely to be answered impartially. The second reassurance, the high loading of a construct representing age-specific transgression, gives credence to the observation that this *individual attitude* factor is *not* merely a reflection of a pure propensity to crime, rather something more complex, and a truly latent propensity as alluded to in a myriad of previous works. The empathy loading (and it must be remembered that this is defined so that a higher score on this is indicative of *less* empathy) at two-thirds of the level of the age-specific moral construct reinforces this observation, as indeed does the relevant

size (at around half) of the loading of the self-control/contentment and the non-age-specific moral dimension.

To round off this discussion, let's return to the fictional person first introduced earlier. Considered this time from their individual attitude, an adolescent with a poor individual attitude, i.e. scoring highly on the scale, will:

- answer that they do not think transgressing rules such as smoking cannabis or playing truant is particularly bad;
- but be less likely to report that stealing from shops, or to an even lesser extent breaking the law in general, is an okay thing to do (although still more likely than a less criminogenic peer);
- are likely to have relatively low levels of empathy, manifesting as being uncooperative and unsympathetic to other others;
- will probably have lower levels of self-control than people with a better attitude;
- will often be relatively unhappy; and,
- will also worry about their unhappy state.

6.2.6 Missing Individual Domains

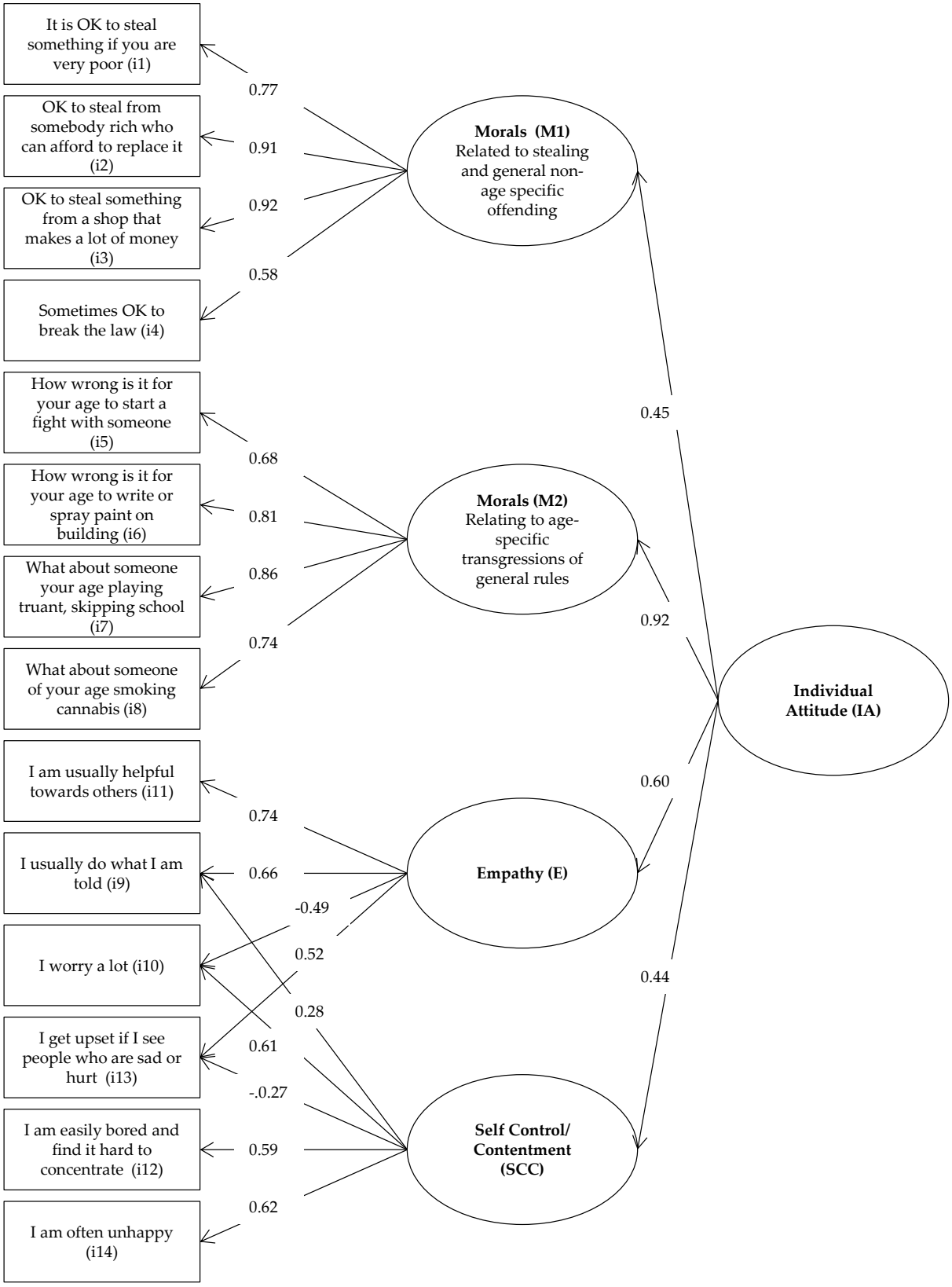
Finally for this section, consideration needs to be paid to what has been missed. Chiefly amongst these is a measure of intelligence and/or achievement - the survey holds no information on attainment or measures of intelligence for respondents aged 16 or under, and hence still at school. In previous works intelligence is typically assessed via an age-specific variation on the Stanford–Binet Intelligence Scales (Schweinhart, Barnes and Weikart, 1993) or the Wechsler Intelligence Scale for Children (WISC) (Kaslow et al., 1990). Latter life attainment is often taken to be school achievement (Farrington and West, 1993). None of these are possible.

'Personality' also lacks the relevant items. The big five personality traits of openness, conscientiousness, extraversion, agreeableness, and neuroticism have a range of specific questionnaires that are used in their measurement and although there is no single 'gold standard' (John and Srivastava, 1999), an approximation is not available here. Temperament is basically the childhood equivalent of personality, usually with more

emphasis put on constitutional predisposition and biological factors and usually assessed via parental judgement at a young age (Farrington and Welsh, 2007; Robinson et al., 1992). Due to data collection methods, i.e. a face to face questionnaire with the respondent and not the parents, these are also not possible, although strong correlations are likely with these and impulsivity/empathy items (Eysenck et al., 1985).

Overall, therefore, it appears that the current definition has strong face validity and incorporates a good deal of the correlates that have been seen with offending behaviour together 'under one roof'. This doesn't seem surprising, given the complicated interrelationship between what in many cases are rather arbitrary distinctions. The second-order specification of an *individual attitudes* latent construct shows good empirical fit to the data but also has a strong substantive interpretation with no counterintuitive implications. This definition of the second-order item should provide a strong, pragmatic and parsimonious item for further modelling work.

Figure 6.3 Final individual level latent variable construct



6.3 Defining Predictors at the Family Level

6.3.1 The Available Data

The majority of the variables to be used for the measurement of the family level correlates are again found in the *health, lifestyle & risk factors* section of the questionnaire, with a few additional items contained within the *household grid* section. Construction and identification of the family level factors is a little less problematic than their individual level counterparts, primarily due to the fact that the questions have a clearer face validity and the concepts being tapped are more concrete, less of the psychological world and hence easier to capture. The questions found to be relevant are contained in Table 6.4.

Table 6.4 Family level questions

Code	Question	Responses categories	Coding	Valid in 2004 sweep
f1	My parent(s) usually praise me when I have done well	yes/no	0 = yes, 1 = no	95.9%
f2	My parent(s) usually listen to me when I want to talk	yes/no	0 = yes, 1 = no	96.0%
f3	My parent(s) usually treat me fairly if I've done something wrong	yes/no	0 = yes, 1 = no	94.6%
f4	My parent(s) usually want to know where I am when not at home	yes/no	0 = yes, 1 = no	97.5%
f5	Would parents mind if found out you started fight with someone	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	95.7%
f6	Would parents mind if you wrote, sprayed paint on building	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	96.9%
f7	Would parents mind if you skipped school without permission	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	96.4%
f8	Would your parents mind if you smoked cannabis	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	97.4%
f9	My parent(s) often argue or fight with each other	yes/no and n/a	0 = yes, 1 = no, 2 = not together	94.6%

Code	Question	Responses categories	Coding	Valid in 2004 sweep
f10	Natural/adoptive mother in household	yes/no	0 = yes, 1 = no	100.0%
f11	Why doesn't your mother live here with you?	3 pt nominal	Routed from f10, 1 = divorce/separated, 2 = deceased, 3 = other	86.5%
f12	Do you still see your mother?	yes/no	Routed from f11, 0 =yes 2 = no	100.0%
f13	How often do you see your mother?	4 pt grouped	Routed on f12, 1 = once week or more, 2 = fortnight, 3 = month, 4 = less often	100.0%
f14	Natural/adoptive father in household	yes/no	0 = yes, 1 = no	100.0%
f15	Why doesn't your father live here with you?	3 pt nominal	Routed from f14, 1 = divorce/separated, 2 = deceased, 3 = other	94.9%
f16	Do you still see your father?	yes/no	Routed from f15, 0 = yes 2 = no	100.0%
f17	How often do you see your father?	4 pt grouped	Routed on f16, 1 = once week or more, 2 = fortnight, 3 = month, 4 = less often	99.7%
f18	You lived - foster family, children's home, young person's unit	yes/no	0 = yes, 1 = no	34.7%
f19	Siblings - Have any been in trouble with police in last 12 months	yes/no	0 = no, 1 = yes, Missing includes don't know and not applicable (no siblings)	95.7%
f20	Have your guardians ever been in trouble with police	yes/no	0 = no, 1 = yes	92..6%
f21	Have your guardians ever been sent to prison	yes/no	routed on f20, 0 = no, 1 = yes	95.2%
f22	How well do you get on with your (step) mum	4 pt Likert	1 = very well, 2 = fairly well, 3 = fairly badly, 4 = very badly	95.7%
f23	How well do you get on with your (step) dad	4 pt Likert	1 = very well, 2 = fairly well, 3 = fairly badly, 4 = very badly	77.5%
f24	How well do you get on with your guardians?	4 pt Likert	1 = very well, 2 = fairly well, 3 = fairly badly, 4 = very badly	1.2%

Questions f10 to f13 and f14 to f17 hold similar information around family disruption, the first set about the maternal positions, the second paternal. These questions were routed so that only respondents answering negative to the first of the set, f10 or f14

whether the mother/father are in the household, are asked the later set and these shall have to be combined into one ordinal scale. f20 and f21 similarly have to be combined, and along with f19 should provide a measure of criminogenic family status. Questions f22 to f24, how well the respondent gets on with their mother/father/guardians may fit with f1 to f4 or f5 to f8, or possibly form their own parental relations factor.

6.3.2 Understanding the Family Question Routing and Low Response Questions

From an end user point of view, the questions f10 to f13 and f14 to f17 involve some rather obtuse routing. This was obviously taken forward in the questionnaire design to lower the respondent burden (Bryman, 2004), but the dataset needs some significant recoding. This is made more difficult by the use of feed forward data taken from the 2003 wave (which is not used in this work - see section 4.4). However, pulling it all together gives the set of results for parental living situations displayed in Table 6.5 (missing values included cases where the respondent either did not specify why they did not see either parent or refused the section or any subsection).

Table 6.5 Derived parental living situation

		Mother	%	Father	%
Live with		2,086	95.3%	1,459	66.6%
Divorced/separated and see...	every week	37	1.7%	223	10.2%
	Fortnightly	9	0.4%	106	4.8%
	less frequently	11	0.5%	98	4.5%
	Doesn't see	3	0.1%	172	7.9%
Deceased		7	0.3%	26	1.2%
Missing		37	1.7%	106	4.8%
Total		2,190	100.0%	2,190	100.0%

As can be seen from Table 6.5, there are very few cases where the respondent did not see their mother, with the vast majority still living with the maternal figure (either natural or adopted). On the paternal side a greater number of respondents did not live with their father (just over a third), with the majority of those in contact with their father at least once a fortnight. A small number of mothers and fathers were reported as deceased. The small group who did not live with their mother, 67 respondents under

the age of 17 in total, are no doubt an interesting group but with a sample size of less than 3%, quantitative analysis is an inappropriate tool to approach the question. The movement of a mother out of the family home is typically a large scale disruption associated with a range of other lifestyle risk factors such as moving into care, reductions in income and a heightened risk of child abuse (Rutter, 1979), and will therefore manifest itself as a complex web of risk factors. The inability to access such an item in this work is a loss, but given the rarity could only sensibly be approached via a specifically designed survey or through other social research methods.

The movement of a father figure from a household, however, is more common in the sample and has received significantly more analytical attention (Farrington and Welsh, 2007). Again the problem of sample size is present when considering those whose father is deceased (just 26 respondents), and setting these unfortunate respondents to missing on this variable is a sensible step, thus preserving the ordinality of the question responses with the answers *"lives with father" "divorced/separated, see weekly", "divorced/separated, see fortnight", "divorced/separated, see infrequently"* and *"divorced/separated, no contact"*.

A similar cleaning exercise was conducted with whether the parent or guardian had ever been in contact with police or sent to prison. The results are displayed in Table 6.6. As can be seen, there is a large number of missing values, with 8.6% of the sample either refusing the first question (f20, *Have your guardians ever been in trouble with police*) or responding that they do not know. In the similar question but asked of siblings, only 115 respondents responded in the affirmative, representing 7% of the sample reporting that their sibling had been in trouble with the police, and this item is removed from further analysis. It should be noted, however, that a third of the respondents were only children. Parental criminal involvement is therefore of sufficient variability to carry forward, whilst the sibling equivalent is not.

Table 6.6 Derived parental criminality variable frequencies

	Frequency	%
No reported contact	1,803	82.3%
Trouble with the police	161	7.4%
Been sent to prison	38	1.7%
Missing	188	8.6%
Total	2,190	100.0%

Periods in care have been shown to increase delinquency and the question f18 *Have you lived - foster family, children's home, young person's unit* would provide a way to capture this. However, only 12 respondents answered yes to this question, leaving it unsuitable for further analysis in this work. Question f24 *How well do you get on with your guardians* is answered only by those who are adopted and has only 25 respondents, hence for similar reasons as f18 is not suitable for further analysis and is deleted.

The process of understanding the questions, their routing and removing certain items that do not have enough variation to allow further statistical analysis has reduced or refined the variables originally in the dataset. f9 *My parents often fight*, has been recoded so that it reads, *no fights, often fight, divorced,* on the understanding that divorce is usually the 'end point' of a dysfunctional (fight laden) relationship. Two new variables have been created and two sets of variables have been removed. A summary of the variables carried forward to the next phase is contained in Table 6.7.

Table 6.7 Family level final and derived variables

Code	Question	Responses categories	Coding	Valid in 2004 sweep
f1	My parent(s) usually praise me when I have done well	yes/no	0 = yes, 1 = no	95.9%
f2	My parent(s) usually listen to me when I want to talk	yes/no	0 = yes, 1 = no	96.0%
f3	My parent(s) usually treat me fairly if done something wrong	yes/no	0 = yes, 1 = no	94.6%
f4	My parent(s) usually want to know where I am when not at home	yes/no	0 = yes, 1 = no	97.5%
f5	Would parents mind if found	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	95.7%

Code	Question	Responses categories	Coding	Valid in 2004 sweep
	out you started fight with someone			
f6	Would parents mind if you wrote, sprayed paint on building	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	96.9%
f7	Would parents mind if you skipped school without permission	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	96.4%
f8	Would your parents mind if you smoked cannabis	3 pt Likert	1 = a lot, 2 = a little, 3 = not at all	97.4%
f9	My parent(s) often argue or fight with each other	yes/no and n/a	0 = yes, 1 = no, 2 = not together	94.6%
f22	How well do you get on with your (step) mum	4 pt Likert	1 = very well, 2 = fairly well, 3 = fairly badly, 4 = very badly	95.7%
f23	How well do you get on with your (step) dad	4 pt Likert	1 = very well, 2 = fairly well, 3 = fairly badly, 4 = very badly	77.5%
f25	Father's living situation	4pt ordinal	1 = living with, 2 = div/sep, see every week, 3 = d/s see fortnightly, 4 = d/s infrequent, 5 = don't see	94.0%
f26	Parental criminal involvement	3 pt ordinal	1 = no reported contact, 2 = trouble with police, 3 = been to prison	91.4%

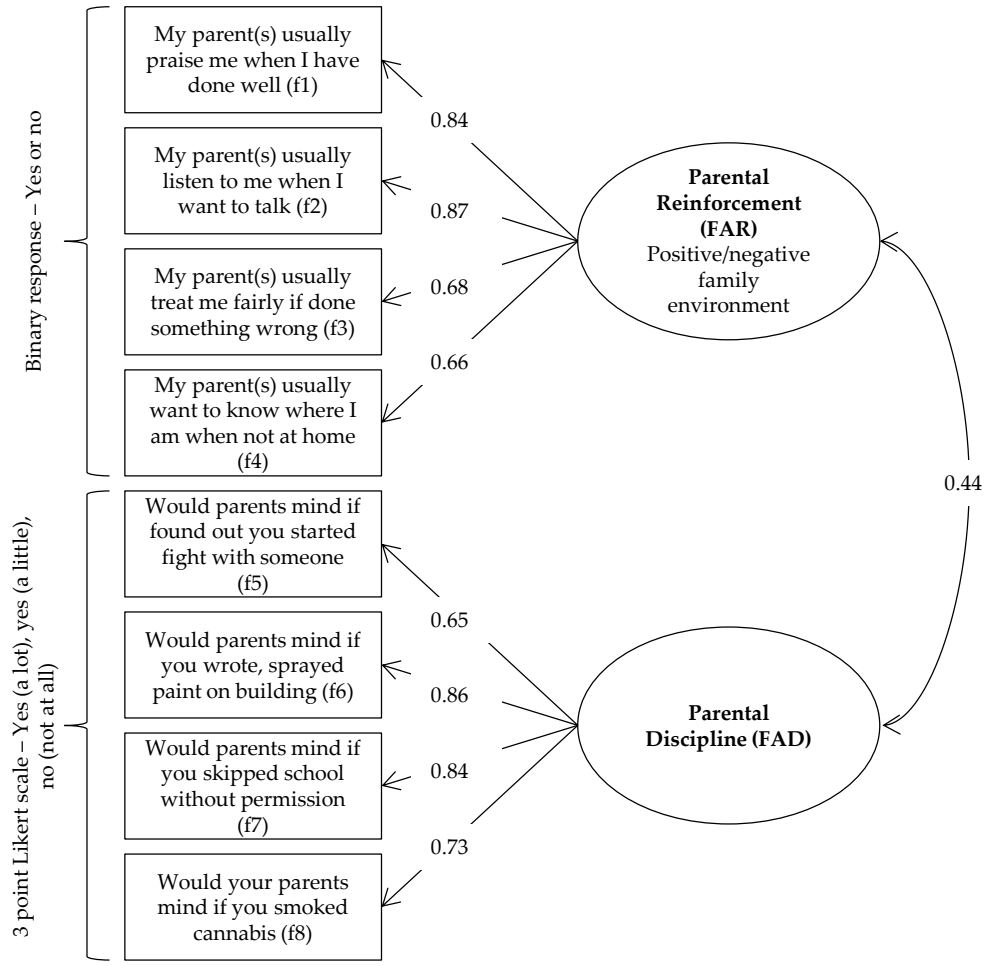
6.3.3 Family Level Latent Variables

As in the individual level section, an initial look at the variables present again provides an outline of how to proceed. It seems relatively clear that f1 to f4 are looking at a level of parental reinforcement of behaviours or parental praise. Items f5 to f8 appear to be closely related to discipline, and it is expected that these items will show a strong relationship with the first latent variable, possibly a one factor solution with the f1 to f4 factor as there is a crossover between parental praise and parental discipline. f9, *My parents often fight or argue with each other*, is clearly representative of parental conflict and may be a manifest variable that can be utilised as it stands without forming a latent variable, although checking whether it fits with other items will be carried out. Question f22 and f23 at first glance seem distinct from other variables, asking about the

slightly nebulous concept of 'how well' the respondent gets on with their parents, although they may load with f1 to f4 as parent/child relations, or f5 to f9 as a reflection of how discipline relates to relationships. This will also be tested. The newly formed f25 *Father's living situation*, and f26 *Parental criminal involvement*, are also not expected to load with any other factors, although again checks will be run.

A one factor solution for the items f1 to f8 does not fit the data (CFI = 0.835, TLI = 0.769, RMSEA = 0.090) but a two factor solution with, as hypothesised, f1 to f4 from the first and f5 to f8 from the second latent variable fits excellently on all criteria (CFI = 0.988, TLI = 0.983, RMSEA = 0.025 with 90% confidence interval from 0.015 to 0.034, and WRMR of 0.887). This model with standardised loadings is displayed in Figure 6.4. All loadings on each latent variable are significant and in the same direction. Due to the way the original manifest variables are coded, a high score on the manifest variables relates to a high score on the latent variables and this represents lower levels of reinforcement, or lower levels of parental discipline. The direction that the latent variables are coded will be a key consideration in structural work. The relationship between the two latent constructs is relatively strong at 0.44, showing that an increase in one factor is likely to see an increase in the other, though these two concepts are not collinear (reflecting the lack of fit onto a one factor solution).

Figure 6.4 Parental reinforcement and discipline model diagram and loadings



The results of the checks of loading f9 *parental fighting*, f22 *getting on well with mother*, f23 *getting on well with father*, f25 *father's living situation* and f26 *parental criminal involvement* are displayed in Table 6.8. Fit indices are displayed alongside standardised loading and their significance levels. The addition of one variable is unlikely to make much difference to the fit indices as the model will still be predominantly defined by the other items, and this can be seen in modification 1 to 5 in the below table where the CFI, TLI and RMSEA values are similarly across the specifications and comparable to the model in Figure 6.4. The main consideration, therefore, is the loadings and the significance of such. Question f9 has insignificant loadings at the 99% level when placed on either of the two latent variables, an unsuitable result in SEM. A similar story is evident for both f25 and f26, with the difference that although the loadings are analytically significant they are meaningfully much smaller than the estimates for the loadings of the original eight variables f1 to f8 (whose loadings change very little in the modified

specifications). This implies that their contribution, or relevance, to the model is minimal and the specification is not correct from a substantive viewpoint.

The only remaining points to be addressed are the addition of f22 or f23, how well the respondent reported 'getting on' with their parents. These had been hypothesised to show an association with either reinforcement or discipline and there is evidence for both of these scenarios, with good fit and loadings of a comparable size to the original items (between a half and two-thirds of the original item loadings). Although a chi-squared test between the two variables shows there was a significant difference between respondents' answers with regard to each parent ($p < 0.001$), one would expect that when these two substantively closely related items were loaded together onto parental discipline or reinforcement, then the model would still have good fit and that the loadings of each item would follow the pattern for each variable when added individually. This hypothesis was tested in modifications 6 and 7 also displayed in Table 6.8. As can be seen in both scenarios, especially when loading both items onto parental discipline, the model can no longer be considered to be a good fit. Furthermore, the loadings of these items significantly affect the size of the original four items on either latent construct, when both f22 and f23 are loaded onto reinforcement, for example, the loading of f5 *My parent(s) usually want to know where I am when not at home* was reduced by 14% (with other items affected similarly). The high loadings of these two items show that they are dominating the specification, changing the interpretation of the latent factors and when coupled with the lack of fit, represent an unwanted result. In summary, these observations lead to the alterative hypothesis, that these items are measuring a concept that is distinct from either discipline or reinforcement.

Table 6.8 Additional items with parental reinforcement and discipline models

	Loadings			Fit indices		
		Standardised loading	Sig.	CFI	TLI	RMSEA
Modification 1						
f9 with FAR		0.108	0.018	0.986	0.981	0.023
f9 with FAD		0.089	0.033	0.985	0.980	0.024
Modification 2						
f22 with FAR		0.565	0.000	0.986	0.980	0.024
f22 with FAD		0.470	0.000	0.939	0.916	0.050
Modification 3						
f23 with FAR		0.483	0.000	0.980	0.973	0.028
f23 with FAD		0.408	0.000	0.965	0.952	0.037
Modification 4						
f25 with FAR		0.153	0.001	0.989	0.984	0.021
f25 with FAD		0.119	0.003	0.986	0.980	0.023
Modification 5						
f26 with FAR		0.329	0.000	0.986	0.981	0.023
f26 with FAD		0.285	0.000	0.983	0.976	0.026
Modification 6						
f22 and f23 with FAR	f22	0.840	0.000	0.952	0.936	0.049
	f23	0.848	0.000			
Modification 7						
f22 and f23 with FAD	f22	0.813	0.000	0.894	0.860	0.073
	f23	0.833	0.000			

In conclusion for the analytical investigation for the family level measurement process, the model to be adopted is that displayed in Figure 6.4. This specification has extremely strong face validity and shows excellent fit to the data with statistically and substantively significant loadings for each item. Reports of frequent parental fights, family disruption in the form of the father's living situation and an increased level of familial criminality (as represented by f9, f25 and f26 respectively) cannot be sensibly forced to create latent variables, and with no other related items in the dataset these will have to be left as is, to be considered as measures in and of themselves. How well the respondents reported getting along with their parents (represented separately as f22

and f23) will also have to take a similar future role. Although there are two of these conceptually similar items with exploratory factor analysis consistently loading them highly onto the same factor, with only two responses any latent model would be unidentified.

6.3.4 Family Level Theoretical Considerations and Missing Domains

To start this discussion of what has been achieved and what has not, a quick review of what is missing is a sensible starting point. Firstly, there is no measure of child abuse. The OCJS was to measure offending patterns in the general population and was not a specific study aimed at any one particular domain. Child abuse is notoriously hard to measure requiring delicate questioning and specifically sensitive approaches. Even with these precautions there is little agreement as to what constitutes child abuse with different measures tapping different aspects and definitions (for a discussion see May-Chahal and Cawson, 2005). In these circumstances the omission is entirely reasonable. Brezina (1998) theoretically framed the offending/child abuse link in terms of social learning theory yet found mixed evidence for this position but the causal chain of child abuse has been found to interact with familial criminality (Simons et al., 1995). The lack of this item will have to be noted as an unfortunate but expected weakness. Another weakness, though this time less expected, is the lack of detail on the size of the families. This simply was not asked, although the dataset hold information on whether the respondent has brothers or sisters living with them (with around one-third of children answer that they live without a sibling). The lack of this is again to be noted but nothing more can be done.

Parental discipline, parental reinforcement of behaviours, parent/child relations, parental conflict, disrupted families and familial criminality (or criminogenic families) appear to be controllable using this dataset. Loeber and Stouthamer-Loeber's (1986) meta-analysis of family factors that correlate or predict juvenile conduct problems or delinquency finds that (in order of importance) poor parental supervision, parental rejection of child, large family size, low parental involvement with children, parental conflict and anti-social parents are the most relevant predictors. Despite the age of this work current perspectives have changed little and although there are a whole host of

measures that have been operationalised the clear face validity and broad correspondence between the items found here and those described elsewhere (Farrington and Welsh, 2007) leads to a strong position for further analysis.

6.4 Defining Predictors at the Area Level

6.4.1 What is meant by 'Area'?

Defining area is a non-trivial task and is often linked to the way that the area or neighbourhood is being studied. Lupton (2003) makes a distinction between the community study and studies of the effects that neighbourhood can have on social and economic outcomes. The former treats the space as the unit of enquiry and is traditionally the domain of social anthropologists, social geographers and sociologists usually using qualitative techniques to explore internal dynamics of a small number of places, often those that are 'deprived' in some manner. The latter, as in the case of this thesis, aims to investigate the difference between deprived and non-deprived areas, and the differences between them.

The division between the two types of investigation reflect data limitations when applying detailed qualitative evidence over larger areas, as it is impossible to collect enough of the fine grained qualitative evidence to apply and create quantitative measures of a large set of neighbourhoods. The quantitative therefore has to rely on administrative data to be compiled to create overall measures or employ exceedingly detailed observational/survey data over these larger areas. This latter approach is taken by the PHDCN and reported in Sampson's *Great American City* (2012) as having been an "intellectual treat that reached its apex in a marathon session" (p84) and consisted of a detailed survey of 8,782 people, systematic social observation whilst sitting in a Sports Utility Vehicle and driving down streets as well as interviews with around 2,800 community leaders. The PADS+ study deployed 12,681 questionnaires (receiving 6,651 responses) in the Peterborough area in its small area community survey to allow measurement of the characteristics of that town. These are particularly rare and expensive data collection processes, and the majority of quantitative area classifications

use existing administrative data and as such are usually constrained to some non-overlapping (Hipps and Boessen, 2013) administrative boundaries.

In the UK context these are often taken to be electoral wards and Lupton (2003) argues that these boundaries often do not match the lived experience, are of such varying size so as to make them unhelpful and can often change. The Office for National Statistics recognised this and in 2004 released a set of concentric "super output areas" (SOA) that are of comparable size with two levels; lower SOA has an average 1,200 households with the larger middle SOA averaging 6,000. From these basic units the Index of Multiple Deprivation was built, a multi-faceted measure of deprivation that will be discussed extensively in Chapter 9 but has been shown to have strong relationships to a number of economic outcomes (National Equality Panel, 2010). Even armed with such standardised special definitions, Lupton questions their usefulness and concludes that less significance should be attached to catch-all measures of neighbourhood and that multidisciplinary work is needed to develop measures specific to research programmes in any one field.

Kearns and Parkinson (2001) adapt Suttles' (1972) schema of neighbourhood and identify three general tendencies towards spatial classifications at which the neighbourhood can exist; the home area, the locality and the urban district or region. The urban district or region is the broadest definition and is primarily defined by an individual's employment and leisure interests. The locality is the level at which residential activities take place and is defined by local administrative and service provisions as well as physical structure. The smallest of these sets is the home area and is typically a five- to ten-minute walk from the individual's home and reflects the interactions with others in their area and (local) friendships. As such, these latter 'neighbourhoods' is where they "would expect the psycho-social purposes of neighbourhood to be… in terms of the quality of environment and perceptions of co-residents, is an important element in the derivation of psycho-social benefit from the home" (Kearns and Parkinson, 2001, p2103). In trying to capture the effect of the neighbourhood, allowing the respondent to self-define their area captures their understanding of their neighbourhood and with appropriate survey controls is an

appropriate and important method that avoids the pitfalls of using administrative blocks. This level of neighbourhood measurement is found in the OCJS where questions around attitudinal aspects of the respondents' area are prefixed by the clear instructions that *"By 'this area', we mean about 15-20 minutes' walk or 5-10 minutes' drive from your current home"* and it is these questions that shall be the main focus here.

6.4.2 The Available Data

There are two different types of area measures available in the OCJS dataset. Administrative structural characteristics (geo-demographics) such as index of multiple deprivation measures (and their subdomains), police force areas, health authority, ONS area classification of local authorities and CACI ACORN customer classifications will be considered separately. The important definitions for this section are to form self-perceived measures of areas, tapping the harder to access respondent opinions about their area circumstances. The task at hand, therefore, is to analyse and model the respondents' answers to questions about the area in which they live. These questions are found exclusively in the *neighbourhood* section of the questionnaire and those relevant to the current work are set out in Table 6.9.

Table 6.9 Area level variables

Code	Question	Responses categories	Coding	Valid in 2004 sweep
a1	This area is a friendly place to live	5 pt Likert	1 = strongly agree up to 5 = strongly disagree, inc. neutral category	99.7%
a2	I trust most people who live in this area	5 pt Likert	1 = strongly agree up to 5 = strongly disagree, inc. neutral category	13.7%
a3	You often see strangers in this area	5 pt Likert	1 = strongly disagree up to 5 = strongly agree, inc. neutral category	99.4%
a4	If children causing trouble, local people will tell them off	5 pt Likert	1 = strongly agree up to 5 = strongly disagree, inc. neutral category	98.4%
a5	People in this area pull together to improve the area	5 pt Likert	1 = strongly agree up to 5 = strongly disagree, inc. neutral category	13.3%

Code	Question	Responses categories	Coding	Valid in 2004 sweep
a6	This area is a place where people from different backgrounds get on well together	5 pt Likert	1 = strongly agree up to 5 = strongly disagree, inc. neutral category	96.8%
a7	People move in and out of my area a lot	5 pt Likert	1 = strongly disagree up to 5 = strongly agree, inc. neutral category	13.3%
a8	How safe you feel walking alone in this area after dark	4 pt Likert	1 = very safe, 2 = fairly safe, 3 = fairly unsafe, 4 = very unsafe	99.7%
a9	How safe you feel walking alone in this area during the day	4 pt Likert	1 = very safe, 2 = fairly safe, 3 = fairly unsafe, 4 = very unsafe	99.9%
a10	Drop wallet near where you live with name & address - get back	4 pt Likert	1 = very likely 2 = fairly likely, 3 = fairly unlikely, 4 = very unlikely	13.6%
a11	How worried are you about your home being broken into and something stolen	4 pt Likert	1 = not at all worried, 2 = not very worried, 3 = fairly worried, 4 = very worried	99.9%
a12	How worried are you about being mugged or robbed	4 pt Likert	1 = not at all worried, 2 = not very worried, 3 = fairly worried, 4 = very worried	99.4%
a13	How worried are you about being physically attacked by strangers	4 pt Likert	1 = not at all worried, 2 = not very worried, 3 = fairly worried, 4 = very worried	99.7%
a14	How common is litter or rubbish in the immediate area	4 pt Likert	1 = not at all common, 2 = not very common, 3 = fairly common, 4 = very common	96.5%
a15	How common is vandalism, graffiti or damage to property	4 pt Likert	1 = not at all common, 2 = not very common, 3 = fairly common, 4 = very common	96.7%
a16	How common are homes in poor condition	4 pt Likert	1 = not at all common, 2 = not very common, 3 = fairly common, 4 = very common	96.6%
a17	Common in your area - Noisy neighbours	yes/no	0 = no, 1 = yes	100%
a18	Common in your area - People sleeping rough in public places	yes/no	0 = no, 1 = yes	100%
a19	Common in your area - People using or selling drugs	yes/no	0 = no, 1 = yes	100%

Code	Question	Responses categories	Coding	Valid in 2004 sweep
a20	Common in your area - People being drunk or rowdy in public	yes/no	0 = no, 1 = yes	100%

The area variables are something of a change from both the individual and family level considerations in that there is a plethora to choose from. Other variable sets that are present include questions around the proximity of friends and family within a 15-20 minute walk or a 5-10 minute drive and the availability of local clubs (such as drama, youth and religious groups). The 20 above represent those that were chosen for their theoretical considerations more than as a census of those available, as in the two previous domains.

Questions a14 to a16 are something of an exception in the OCJS as they are reported by the interviewer on their approach to residency. These are analogous to what Brunton-Smith and Sturgis (2011) use based on the Crime Survey in England and Wales where they call their underlying factor "neighbourhood disorder". In this work however it felt that these are more akin to neighbourhood condition. As these are assessed by the interviewer these shall be discussed as control variables in Chapter 9.

The rest of the variables are selected specifically to create three distinct latent variables. Questions a1 to a10 (or some subset thereof) should collapse as a manifestation of collective efficacy, a11 to a13 should reflect fear of local crime and a17 to a20 should give a measure of what will be termed here *Neighbourhood disorder* (a full discussion of this shall be returned to). The questions a8 and a9 look as though they may be part of the fear of local crime dimension or a more general confidence issue, especially with regard to a9 *how safe the respondent feels walking alone at night*, whereas walking at day could be more of a reflection of social cohesion and trust. This will be discussed.

A major and rather troublesome problem with this part of the survey relates to item response rates. Table 6.9 shows that there are valid answers for just over 13% of the sample for some questions. This is because questions were asked differently within the survey based on age, with a2, a5, a7 and a10 only asked of those aged 16 or over. This

stands out against the majority of age related question changes, which for the main alter at age 17 and over (rather than 16 as here). This will form an integral part of the consideration of the collective efficacy dimension.

6.4.3 Defining a Collective Efficacy Dimension

Sampson et al. (1997) define collective efficacy by using questions present in the Project on Human Development in Chicago Neighbourhoods (PHDCN). It was originally conceived as having two dimensions, *informal social control* and a *social cohesion and trust* component. Questions were answered on a 5 point Likert scale with the following items:

- To measure the *informal social control* component whether their neighbours would intervene in various ways if:

    a) children were skipping school and hanging out on a street corner,
    b) children were spray-painting graffiti on a local building,
    c) children were showing disrespect to an adult,
    d) a fight broke out in front of their house,
    e) the fire station closest to their home was threatened with budget cuts.

- To measure the *social cohesion and trust* component whether:

    a) people around here are willing to help their neighbours,
    b) this is a close-knit neighbourhood,
    c) people in this neighbourhood can be trusted,
    d) people in this neighbourhood generally don't get along with each other (reverse scored),
    e) people in this neighbourhood do not share the same values (reverse scored).

For each of these two scales the score was the average of the responses, and the analysis included those respondents who answered one of the ten questions or more. Finding that these two measures were highly correlated across neighbourhoods ($r = 0.80$, $p < 0.001$) the two scales are combined and ultimately collapsed to a summary measure and labelled the underlying latent construct to be collective efficacy. The measure has received significant analytical and policy attention and although Odgers et al. (2009) identify the Sampson et al. (1997) definition of collective efficacy as "one of the most

widely used assessment tools in neighbourhood research and is currently considered the gold standard for measuring neighbourhood-level social processes" (p948) differing variable availability in other works has meant that other definitions can and have been used. For example:

- Lawrence and Heath (2005) use factor analysis, the 2005 Citizenship survey and the questions matching the informal social control component in Sampson's 1997 work to create a collective efficacy factor that they add to their multi-level model. They do not use a social cohesion and trust component, which they treat as a separate variable accessed through a single question, and take this measure of cohesion as dependent in their multi-level modelling.

- Bellair (2000) focuses on the informal social control elements and constructs this through an informal surveillance measure (do neighbours watch either others' properties?) present in a Seattle Police Department 1990 victimization survey and use structural equation modelling to analyse street crime.

- Cohen, Farley and Mason (2003), although also using the PHDCN, concentrate on the social cohesion and trust component (i.e. trustworthiness, friendliness, helpfulness) of Sampson's (1997) definition and conduct least-squares regression analysis on official mortality rates.

- Browning (2002), again using the PHDCN as a base, adjusts the measure of collective efficacy to account for the social composition of neighbourhoods to extend the coverage of the collective efficacy measure to the 1995-1997 Chicago Health and Social Lifestyle Survey and uses hierarchical linear modelling focused on the victimization of women by men.

- Xu, Fielder and Flaming (2005) use structural equation modelling and the Citizen Survey conducted by the Colorado Springs Police Department in 2001 in their study of the effectiveness of community policing. To represent collective efficacy they construct social cohesion from three questions about neighbourly interaction (do they talk, help and trust) and a shared expectation/social control

component out of questions about whether residents are willing to take responsibility, work effectively on local issues and work with the police.

- Wells et al. (2006), in their hierarchical modelling on the effects of collective efficacy on responses to neighbourhood problems review differing measures of collective efficacy before being confident with their definition covering 'looking out for each other' on the social control side and sharing items/having conversations with neighbours for the cohesion/trust component. A 'community' is defined by individual responses aggregated to small residential units (containing on average 61.75 respondents).

- Reisig and Cancino (2004) investigate incivilities in nonmetropolitan communities in the State of Michigan and combine ten questions similar to those used by Sampson but without accounting for the different cohesion/trust and social control aspects; they simply sum all ten responses and enter the measure into hierarchical Poisson regression models.

- Bradford and Myhill (2015) look at changes to public confidence in the criminal justice system and use an operationalisation of collective efficacy using just two items of how likely local people would do something about "*Children spray painting graffiti on a local building?*" and "*If there was a fight near your home and someone was being beaten up or threatened?*"

The story, therefore, is of a robust and versatile measure that can be defined in a number of slightly different ways dependent on item availability. The important point is that the manifest variables relate to either informal social control and/or social cohesion. Although ideally both of these domains will be captured by any measure, there are examples where only one or the other has been used successfully and produced broadly consistent results with alternative definitions.

The question subset of Table 6.9 displayed in Table 6.10 holds the items that are to be considered for inclusion in this analysis, along with their age-specific validity and whether they are best considered as social cohesion or as informal social control. Which

domain most questions belong to, although slightly irrelevant due to the single latent construct ultimately thought to be causing these response patterns, for the most part are rather clear. Four items deserve attention:

- Question a3 *You often see strangers in this area* has been defined as informal social control as it is felt that it represents watchfulness rather than a level of trust. The question asks whether a stranger is seen, not whether they cause problems or anything about a level of interaction. The 'seeing' seems to relate closely to respondents looking out, although an argument could likely be made in favour of either domain.
- Question a7 *People move in and out of my area a lot* is out of place, with exploratory factor analysis consistently loading it onto a separate factor. This is therefore removed. This is also only valid for 16-year-olds and is the only measure of residential instability in the area, and with its deletion so too goes the chance of controlling for a relatively important factor.
- a8 and a9 *How safe you feel walking alone in this area during the day/night* do not classically appear in the definition, and in the case of a8, walking alone at night is most probably tapping a slightly different dimension, a fear of the unknown or related to victimisation[24]. A chi-squared test between both a8 (night) and a9 (day) showed that although both have a statistically significant relationship with whether the respondent was a victim of a crime in the previous year, the association was strongest for night time (Goodman and Kruskal's gamma = 0.13) than day time (gamma = 0.08). It was felt better on balance to include only a9 (day time fear) in the model as it more clearly captures neighbourly watchfulness with the presence of light allowing neighbours to look out for one another.

[24] Interestingly, this item does not load with a11 to a13, which are directly asking about fear of crime, and this point shall be returned to in section 6.4.4.

Table 6.10 Collective efficacy variables

Code	Question	Domain	Respondents
a1	This area is a friendly place to live	Cohesion and trust	All
a2	I trust most people who live in this area	Cohesion and trust	16-year-olds
a3	You often see strangers in this area	Informal social control	All
a4	If children causing trouble, local people will tell them off	Informal social control	All
a5	People in this area pull together to improve the area	Cohesion and trust	16-year-olds
a6	This area is a place where people from different backgrounds get on well together	Cohesion and trust	All
a7	People move in and out of my area a lot	REMOVED	16-year-olds
a8	How safe you feel walking alone in this area after dark	REMOVED but to be checked	All
a9	How safe you feel walking alone in this area during the day	Cohesion and trust	All
a10	Drop wallet near you live with name & address - get back	Cohesion and trust	16-year-olds

The resultant models to be further analysed are displayed in Figure 6.5. These items are being collapsed straight onto the hypothesised underlying collective efficacy dimension, without paying heed to the intermediary items of social cohesion or informal control. This is the expected final structure and with only two informal social control items recognized in the range of variables, and with only two first order constructs, both the first order and the second order latent constructs would be unidentifiable. The dotted lines represent the components of the model that are only valid for 16-year-old respondents.

Two models were therefore estimated, one with the complete set of items and one with the reduced set, leaving out those variables that were only valid for the 16-year-olds. For added rigour the sample was also restricted to just those aged 16 years of age (301 respondents) and both models were run against this dataset as well. The fit indices and loading for all four of these possibilities (two models for two different groups) are presented in Table 6.11.

Figure 6.5 Possible collective efficacy model definitions

All model specifications fit the data well, with high CFI and TLI levels, and the RMSEA ranging across specifications from 0.029 to 0.080. The WRMR, although generally taken as a supplementary measure, also shows a good level of fit. In the reduced models the variables a1 *this area is a friendly place to live*, a6 *this is an area where people for different backgrounds get along well together* and a9 *how safe do you feel walking alone during the day* dominate in terms of the loadings for the latent variable, although a4 *if children are causing trouble, local people will tell them off* has a lower relative loading (at between a half and two-thirds of the other items' loadings); with a model such as this, very well fitting and with no overly dominant terms, is an acceptable result. The first three variables represent what has been identified as cohesion and trust items, but the lower loading of a4, an informal social control component, reassures that these models are indeed

tapping collective efficacy, though perhaps with a slight bias towards more of a cohesion latent construct.

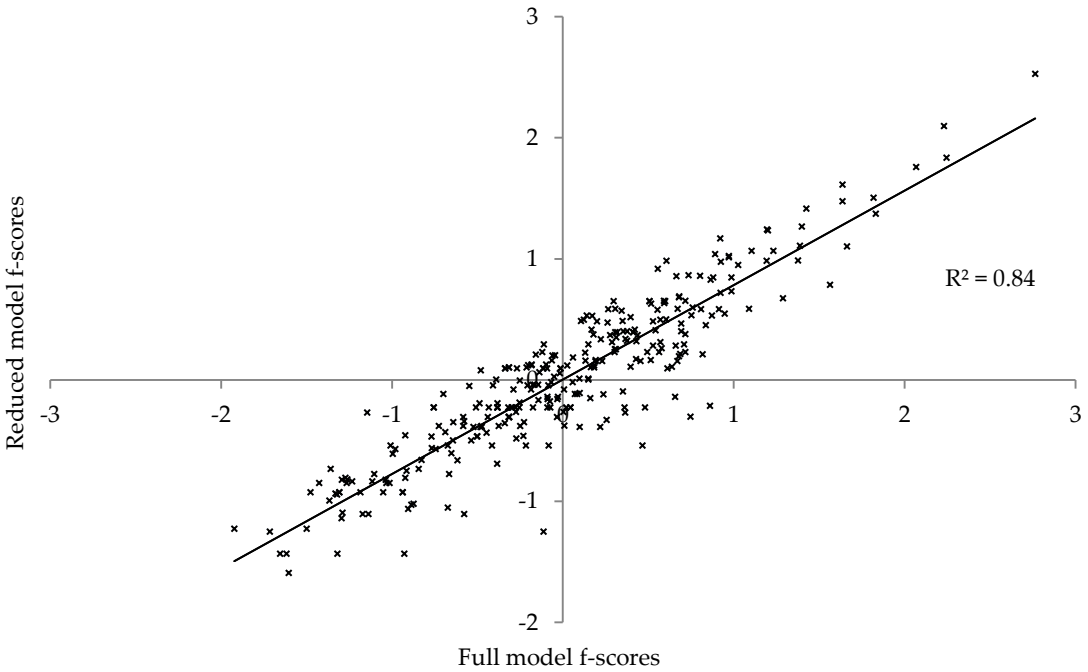
Table 6.11 Collective efficacy model fit and loadings

	All respondents		16-year-olds only	
	Reduced	All items	Reduced	All items
		Fit indices		
CFI	0.989	0.985	0.969	0.970
TLI	0.979	0.979	0.939	0.957
RMSEA	0.038	0.029	0.073	0.080
Prob RMSEA less than 0.05	0.827	1.000	0.180	0.019
WRMR	0.672	0.870	0.553	0.784
		Loadings		
a1	0.788	0.813	0.798	0.846
a2		0.862		0.848
a3	0.253	0.265	0.335	0.312
a4	0.358	0.367	0.444	0.420
a5		0.597		0.597
a6	0.620	0.596	0.578	0.590
a9	0.528	0.532	0.609	0.568
a10		0.585		0.583

When widening out the model to include all the items for both groups (i.e. allowing a large set of missing values for respondents under the age of 16 which can be handled adequately by the WLSMV estimator in Mplus), the loadings on the items common to both models vary little and are certainly all of a substantively similar magnitude. With these added items the definition of the latent factor as being collective efficacy is much clearer, and although it could be argued that there is still a bias towards the cohesion and trust component, this is as much a symptom of the available variables as the model results. With only one true informal social control component this will always be the case with this data (although using just 'one half' of the definition of collective efficacy is not unheard of, as the review earlier in this section demonstrated).

So, there are two well-fitting alternative structures, one based on a reduced set of items and, in the case of all respondents, one with a high level of systematically missing data. To assess whether the construct is the same in both cases, as this is fundamentally a confirmatory factor analysis, the factor scores (i.e., each individual's resultant score on the latent variable for each of the models) can be produced and tested for equivalence via a paired sample t-test, and visualised via a simple scatter plot. If there is no statistically significant difference between the two scores they can be considered equivalent. Results back up this assertion; checking between the f-scores for the reduced and all item models based on 16-year-olds (the models in the two right hand columns of Table 6.11) gives a p value of 0.89, meaning there is no evidence that the two models are giving different results on the latent factor. Looking instead at the models defined from all respondents, the overall p value of the difference between the two sets of latent scores is 0.59. In summary, there is no evidence of a difference between the model definitions of the collective efficacy dependent on which variables are included. This can clearly be seen in the linear relationship between the two scores, with an r-squared of 0.84 (Figure 6.6).

Figure 6.6 Scatter of f-scores from both definitions of collective efficacy, 16-year-olds



The brief review of collective efficacy measures has shown that there are considerable differences in the definition that has been applied in previous research, and the analysis presented has shown that the measures available in the OCJS can indeed be considered suitable measures despite missing values and that the question set is reduced for respondents aged under 15. Roughly speaking, more variables are better (Bentler, 1980) but with such a large number of missing values present for those aged under 16 years, and considering that the scores amount to the same under both definitions and that many different, less rigorous operationalisations have been used elsewhere, the reduced form of the definition shall be used throughout this analysis.

6.4.4 Adding Neighbourhood Disorder and Local Fear of Crime

Variables a17 to a20 were hypothesised to be indicative of a neighbourhood disorder latent variable and why this was named as such was left to the side. The reason why it has been decided to call this disorder, rather than giving that name to the interviewer-assessed set of questions is that it seems to be more closely aligned to the concept of disorder as reported by Wilson and Kelling's (1982) *Broken Windows* and match what Brunton-Smith, Jackson and Sutherland (2014) call *social disorder*. This latter paper also includes a measure of problems of litter and vandalism as *physical disorder* and finds that types of disorder are not separately perceived by respondents. The differentiation in nomenclature is more for sake of clarity; neighbourhood disorder as perceived by the respondent, and neighbourhood condition (to be fully defined in Chapter 9) as reported by the interviewer. It would have been equally possible to call them *perceived neighbourhood disorder* and *interviewer neighbourhood disorder*. The items a17 to a20 fit together very well (CFI = TLI = 1, RMSEA = 0.000 and WRMR = 0.143) when assessed independently with all loadings over 0.45.

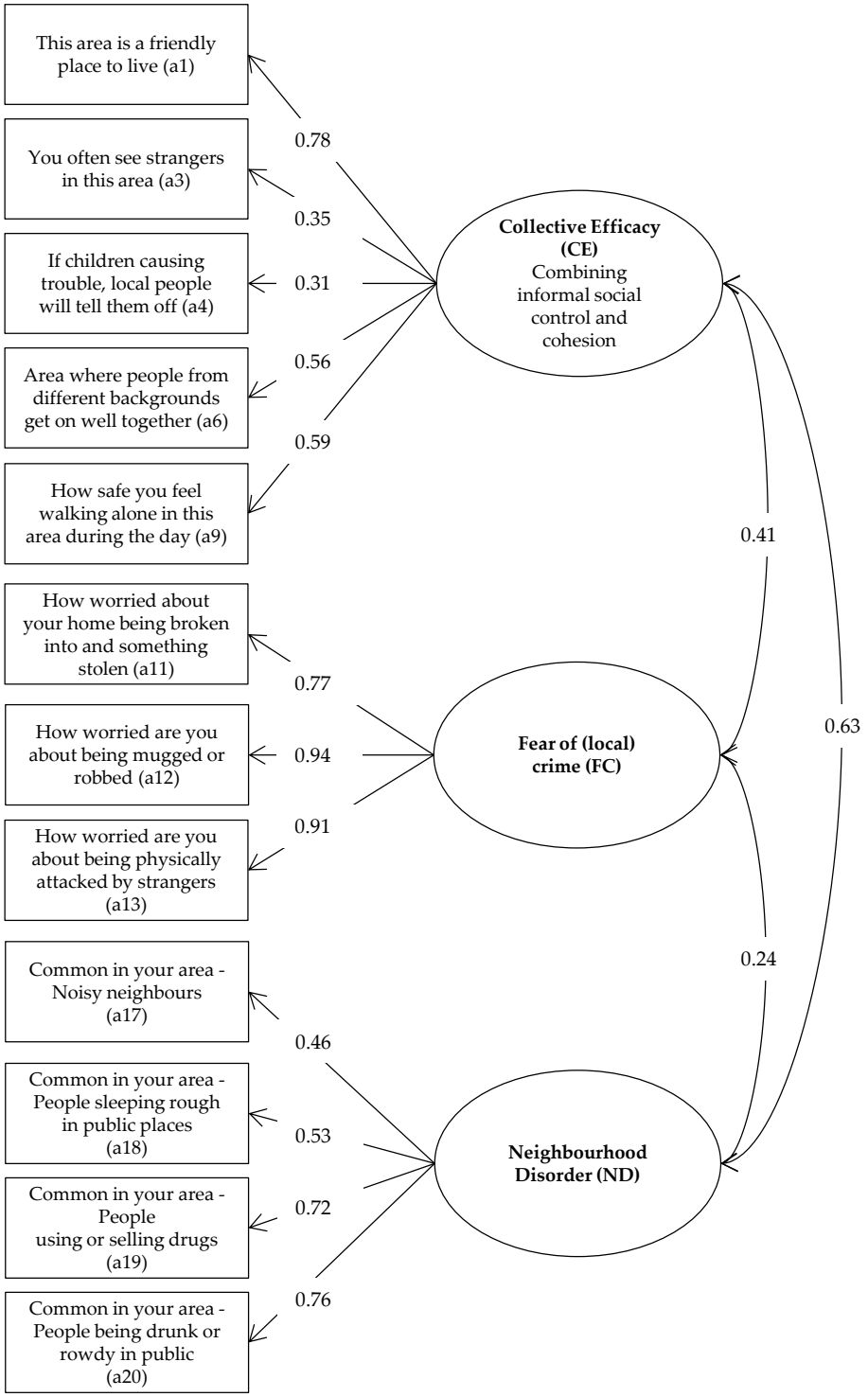
Similarly, the proposed local fear of crime (a11 to a13) latent construct fits the data extremely well (CFI = TLI = 1, RMSEA = 0.000 and WRMR = 0.001 in both cases), with all standardised loadings over 0.9 (bar a11 at 0.766). Adding a8 *how safe do you feel walking alone after dark* has a much lower loading at 0.4, less than half the loading of the other items. When looking bivariately, this variable (a8) with a12 *how worried are you about being mugged or robbed (in your local area)* or a13 *physically attacked by strangers* shows that

even of those who claim to feel very unsafe (10% of all respondents), around a third are not worried about being mugged or being attacked by strangers. Coupled with the weak association with victimisation, a8 thus appears to be measuring something closer to a general level of fearfulness which appears unconnected to other measured items.

To make sure that these latent variables are indeed three separate factors, an exploratory factor analysis against all of the variables was run with either two or three factor solutions allowed. A two factor solution showed poor to adequate fit (TLI = 0.941, CFI = 0.937 and RMSEA = 0.087), with fear of crime showing a clear separation from collective efficacy and neighbourhood disorder, which both loaded onto the same factor. This does not mean that these answers are actually derived from one latent variable, however, as the loadings of both the collective efficacy items and the hypothesised neighbourhood disorder items are lower (by around 0.2) than when analysed separately, showing that it is something 'in-between' the two. Given the clear expectation that these two should be caused by two separate but related perceptions, one about neighbourliness and the other about actual disorder (as demonstrated in numerous papers, Brunton-Smith, Jackson and Sutherland, 2014 being one clear example) and finding that a three factor solution fits the data very well (CFI = 0.989, TLI = 0.987, RMSEA = 0.021) with the loading of the items approximately equal to those in the individual models, then it appears that this three factor solution is the most appropriate. Collective efficacy items match the rotated factor 1, the local fear of crime loading onto factor 2 and neighbourhood disorder variables loading onto the third factor. Item a9 (safe walking in the day) loads onto the collective efficacy dimension with a value of 0.5 and onto factor 2, local fear of crime, with a value of 0.2 (compared to the other items' loadings of over 0.8). This shows that whilst there is a degree of relationship, fear of walking alone during the day is not measuring fear of crime and is rightfully placed with collective efficacy. Item a8 (walking at night) loads approximately equally across collective efficacy (0.3) and fear of crime (0.4). This further reiterates the confused concept to which *fear of walking alone at night* matches, and rather than battle with an unknown is hence removed from the analysis.

All that remains, therefore, is to specify the model from a confirmatory approach. The structure and loadings can be seen in Figure 6.7. The model fits very well (CFI = 0.989, TLI = 0.985, RMSEA = 0.050 and WRMR = 1.673). All questions are coded such that a higher score on the latent variable represents a lower level of collective efficacy, a higher level of fear of local crime and worse neighbourhood disorder. Loadings are sensibly high, except where discussed in terms of the collective efficacy components and each covariance between factors are in the same direction. There is a strong relationship between disorder and collective efficacy, as expected, but not too high as to be too much of a worry about collinearity in any analysis, and given the EFA results and their theoretical distinction, is allowable.

Figure 6.7 Final area level model diagram and loadings



6.5 Latent Variable Measurement Conclusions

This chapter has demonstrated that with appropriate consideration of theoretical positions, previous works and manifest items within the Offending, Crime and Justice Survey, it is possible to create well-formed measurements of various dimensions that have been shown to predict offending. Each model that has been defined fits the data well from an analytical perspective. Moreover the discussions around these definitions, coupled with their clear face-validity, have resulted in definitions that are substantively meaningful and should be considered to accurately represent the dimensions they were aiming to tap.

The individual attitudes dimension was probably the hardest to access and relied a good deal on the identification of a general mixture of often overlapping factors that have been discussed and analysed in other work. Nevertheless the definition arrived at with a second order individual attitude latent construct above the moral, self-control and empathy dimensions makes intuitive sense. The addition of this higher order factor could be contentious, despite the fact that there is no clearly exposed delineation between the concepts that go into its formulation and there being both analytical and theoretical overlaps between the constituent parts presented here. The parsimonious nature of the current conception will lead to more manageable structural work, allowing a clearer reading of how the individual's own psychological perspectives relating to their own behaviours can influence their offending.

The family level measurement was a little clearer than the individual level counterpart, though came with its own set of issues to be addressed. Whilst it was relatively straightforward to define the parental reinforcement and discipline factors other correlates that have been seen to be relevant were defined here and some were seen to be missing. Primary amongst these missing components was a measure of family size, and its absence can only be attributed to survey design. Parent/child relations, parental conflict, disrupted families and familial criminality can be measured but to a lower level of accuracy, as their measurement rests on individual manifest variables and therefore their reliability suffers. The good news is that these items represent the less

important end of the family level issues and will be worked into the analysis at a later stage.

Formulating the area level measures took an exceptional place in the analysis in that the variables used were taken for their theoretical importance rather than a census of all that were available in the dataset. Other domains such as density of social ties or availability of activities in the local area will also be measurable but as this realm of criminal risk factors is the least well developed only those that have been seen to be of most importance have been included thus far. Once the items to be carried forward were decided upon the deduction of the latent variable models followed relatively easily, especially in the case of neighbourhood disorder and local fear of crime, although a few cross-loadings with other items were considered and rejected on the grounds of fit and/or substantive meaning. The collective efficacy definition was complicated by missing data due to questionnaire design. However sensitivity analysis showed that the reduced form, with fewer variables, correlates extremely highly with the full definition, with no statistically significant difference between the two measures and hence using the sparser definition is suitable. Careful checks were made to ensure that disorder and efficacy were indeed analytically as well as theoretically separate.

Taken together, this work represents a satisfactory start to modelling criminal risk factors from this non-specific dataset. The next component of the work is to start to interact these latent variables with each other and the ultimate dependent, self-report offending.

Chapter 7

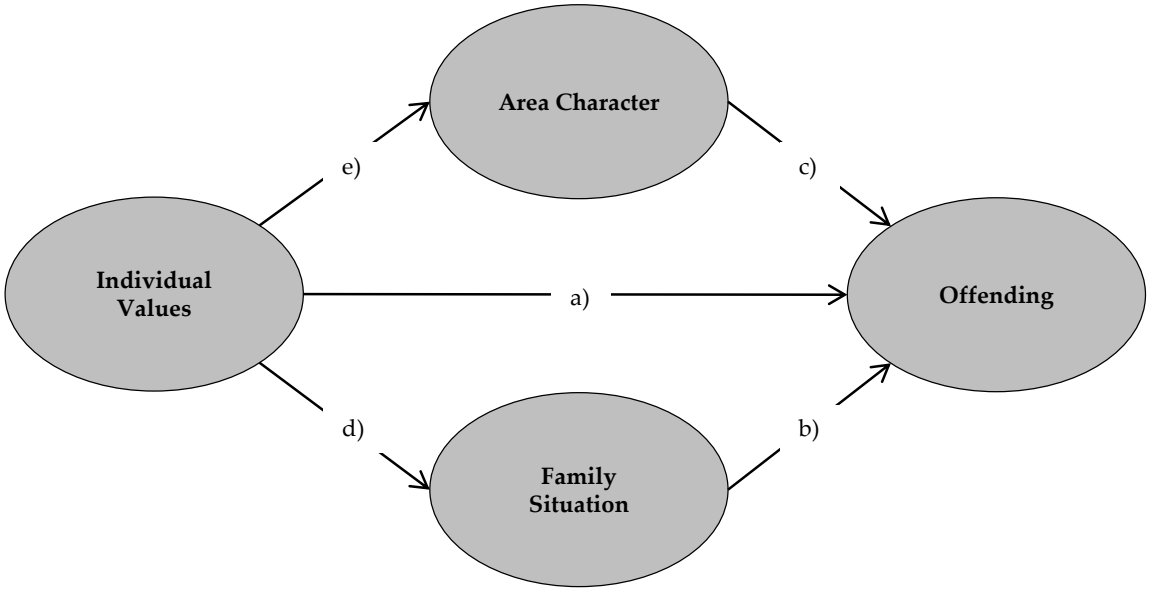
Building a Psychological Risk Factor Map of Interrelations

7.1 Process and Aims

This section will attempt to bring together a viable model based on the components that were shown to be measurable in the previous chapter. This structural model is to be built from the bottom up, following best available theoretical evidence with constant reference to expected structures. As such the best apparent approach is to add components sequentially, starting with the simplest model and working up a full model. Ideally, the building of a theoretical model should be grounded in other related theories and this approach will be taken here. This chapter will focus on the perceived risk factors and will create a 'psychological map of interrelations' of the latent variables as defined in the previous chapter. The 'external' control variables as mentioned briefly in the previous chapter will be considered in full in Chapter 9.

Figure 7.1 gives a sketched outline of that which is being aimed for, and will be used as a start point for the theoretical discussion. Each (grey) oval represents a theoretical domain of interest, and can comprise manifest variables, a latent variable or a number of latent variables (dependent on the model). Before moving on to represent this schematic as an operationalised analytical model it is first necessary to consider the interrelations from a theoretical viewpoint, after which it will be possible to analyse within each domain and then go on to combine them to form the final model.

Figure 7.1 Theoretical schematic for the psychological map of interrelations



There are four domains (individual values, family situation, area character and offending) to be discussed, alongside five pathways, labelled a) through to e) and ordered in Figure 7.1 as will become apparent. The first part to talk about is the offending component on the very right hand side, which has been discussed in section 4.6. This is the dependent variable of the modelling and this work is to look at whether an individual offends or not. There are numerous works that look at the offending rate of an area (Bottoms, 2007), what affects a person's fear of crime (for example, Brunton, Smith and Sturgis, 2011) and levels of victimisation (Cohen and Felson, 1979; Hindelang, Gottfredson and Garofalo, 1978). Although important these options are not the focus of the current work. In this chapter offending will be the simplest measure: whether they reported offending in the previous 12 months or not. This will be varied in Chapter 8, but will always be kept to offending behaviour, rather than, say, straying into drug use or anti-social behaviour more generally. This is again for reasons of focus.

The next two domains are family situation and area character. Their causal paths to offending, b) and c) respectively, are similarly rather simple to understand at a basic level. 'Worse' family situations, or more precisely more criminogenic family environments, would be expected to increase the reporting of offending for an individual. Similarly a less positive area character, i.e. one that has been theorised to be more conducive to criminal behaviours, is expected to increase the rate of reported

offending. These causal chains are not complicated, their analysis relatively simple and a discussion of them is rather reductive without talking about the specific latent variables that are being included in the model.

That leaves the individual values component to discuss and its relationship, path a), to offending along with the paths from individual values to family situation or area character, paths d) and e). Path a) is simply saying that those with higher criminogenic propensities (a phrase used synonymously with the operationalised version called individual attitudes) are more likely to self-report offending. This is backed up in numerous works and is a well-accepted criminological finding (as previously discussed in section 6.2). Paths d) and e) on the other hand are hypothesising a more complicated process and are looking at how perceptions of higher level domains are working as mediating effects, or social controls, on the individual propensity to offend. This exact relationship will be the subject of the next section. What this work is in essence doing is taking these individual values as the basis for the modelling procedure and by extension, with specific reference to the realist ontological position and the causality implication as discussed in section 5.4, these individual positions as defined in this foundational block affect the answers which a respondent gives whenever they are considering questions related to their criminogenic controls of the higher level components, and *not the other way round*. In other words, it needs to be theoretically argued that the individual attitude latent variable, which has been shown to correspond to the operationalisation of the criminal propensity in previous developmental criminological work, does not vary based on the answers that a respondent gives to their area or family circumstance, rather that the answers to these latter items are (in part at least) caused by variation in the individual attitudes. This seems at initial reading a bold claim to make, but a reading of the theories upon which a good deal of developmental criminology is based reveals that it is not unfounded, and in fact is a sensible position to take. Empirically this result is fundamental to allowing a causal chain to be constructed that corrects for perceptual response biases to these higher level components given an individual's criminogenic disposition.

7.2 The Centrality and Stability of Individual Attitude

As previously noted, self-control is the bedrock of Gottfredson and Hirschi's (1990) *General Theory of Crime*, and the work is clear: "Individual differences in self-control are established early in life (before differences in criminal behaviour, however the state defines it, are possible) and are reasonably stable thereafter" (p177) and are attributable predominantly to parenting behaviours, child-rearing practices and socialisation influences. A rather prevalent misconception is that Gottfredson and Hirschi meant that self-control does not change *at all.* This does not bear out, as the original work clearly implies that it is the *ranking* of an individual's self-control across a population that is stable, allowing for age-graded increases over time (as found by Sampson and Laub, 1993, to name one of many). Regardless, at any one moment in time, "individual differences in the tendency to commit criminal acts (within an overall tendency for crime to decline with age)… remain reasonably stable with change in the social locations of the individuals and change in their knowledge of the operations of the sanctions systems" (p87). This early development of self-control or, to put it more broadly, criminal propensity (Akers, 1991) can equally be seen in Sampson and Laub's (1993) *Crime in the Making*. Described pictorially on page 244, where "individual difference constructs" (comprised of "difficult temperament, persistent tantrums and early conduct disorder") form in childhood (ages 0 – 10) they subsequently interact with varying social control processes in adolescence and adulthood to produce persistence or desistance from crime. Although focused on the whole life course, one of the key observations at the very start of their book is "criminologists have largely ignored the link between social structural context and the mediating processes of informal social control" (ibid, p7). Le Blanc (2006) concludes similarly in his developmental psychology approach to the interaction of self-control with social controls; that social controls are mediating effects, rather than causal on self-control *at any one period of time* (though paint a much more complex picture over a longer time frame). Wright, Tibbets and Daigle (2008) would take this even further in their neuro-psychological approach to the development of offending behaviour, attributing the basis for criminal propensity to genetics and hence claiming that it colours the perceptions of all else.

In Moffitt's (1993) dual taxonomy, although often set as diametrically opposite to Sampson and Laub's sociogenic developmental theory (Sampson and Laub, 2005), the same is immediately obvious for the small group of 'life course persisters' whose offending, it is claimed, is driven by neuropsychological vulnerabilities developed and set in early age by parental failure (poor nutrition, parental substance abuse and the like). The 'adolescent limited group', however, seem to present a more complicated picture. They are generally well-socialised and adjusted, and their delinquency is adaptive to reinforce certain stages of the maturing process and expresses itself through mimicry of the life course persister group. What this does not mean is that their personalities change during this process, and although the 1993 work does not explicitly state that these personality traits are fixed at a younger age, the article makes tacit acknowledgements that this is indeed the case throughout: "In stark contrast with the earlier account of life-course-persistent offenders, personality disorder and cognitive deficits play no part in the delinquency of adolescence-limited offenders" (p685); "In general, these young adults have adequate social skills, they have a record of average or better academic achievement, their mental health is sturdy, they still possess the capacity to forge close attachment relationships, and they retain the good intelligence they had when they entered adolescence" (p685); and, "Instead of a biological basis in the nervous system, the origins of adolescence-limited delinquency lie in youngsters' best efforts to cope with the widening gap between biological and social maturity" (p686). Hence, as with alternative theories, when questioned about their criminogenic propensity (individual attitude) their answers should be stable. Their responses to questions surrounding family and area situation, on the other hand, should alter dependent on the maturing process.

Finally, Wikström's situation action theory (SAT) contends "that acts of crime are the outcome of a perception-choice process initiated by the interaction between a person's crime propensity (determined by a person's morality and ability to exercise self-control) and criminogenic exposure (determined by the setting's moral norms and their enforcement" (Wikström et al., 2012, p406). In an earlier work Wikström and Treiber (2007) state "Critically, this implies that stability and change in an individual's ability to exercise self-control depend not only on the stability and change in his/her executive

capabilities (an individual trait) but also on the stability and change in the environment(s) in which he/she operates" (p238). This is relating to the specific moment in which an offence is or is not committed; SAT is "concerned with causation only when an individual considers (deliberates) whether or not to engage in an act of crime" (ibid, p238). What this does not mean is that the responses to questions surrounding self-control or morality are affected by the place in which they live, more that the individual acts in accordance with their morality when faced with temptation or provocations and self-control acts as an inhibitor. The development of a person's crime propensity is not within the scope of their theory (as currently expressed, although the 2012 work hints that this will be the focus of a forthcoming, as yet unannounced publication), but it appears to be regarded as an executive capability and therefore a trait, and hence and again, stable.

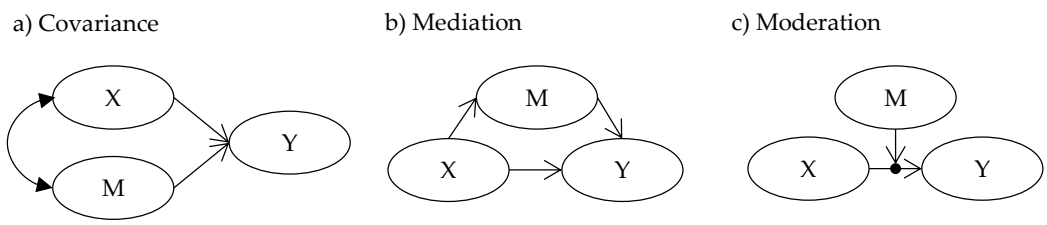
It appears, then, that the criminal propensity of an individual is widely regarded as being a relatively fixed trait (relative in the sense that it is fixed *relative* to others) across a number of competing theories in developmental criminology, something that develops in early through to late childhood (at the bare minimum before the age of 10), perhaps with a natural decrease across the life course due to developing maturity. Each of the theories discussed above explicitly acknowledges that environmental factors (such as locations, temptations and family settings) all interact with whether the individual would commit a crime at any specific time-point, but none imply that the basic propensity of the individual would alter based on these situational settings. When questioned in a situation devoid of temptation or provocation (a professionally conducted interview where the respondent is asked to carefully consider their answers to each question) then one should expect that the morality or the self-control to be expressed properly and answers to be reliable. If this is not the case then it is the survey instrumentation that is at fault, and a larger issue than that which is being addressed in the current work is called into question (namely questionnaire reliability).

7.3 Covariance, Moderation or Mediation?

Given that a wide range of literatures and evidence suggest that it is reasonable and justified to assume that individual criminogenic propensity develops and fixes at a

relatively early age, the specific question that now needs to be addressed is how it interacts with other variables that can be placed into the model. Although the components that were measured and defined in Chapter 6 were split into three different 'levels', individual, family and area, in the survey they were all asked of individual respondents (with the odd exception, such as the neighbourhood condition set). It is also necessary to split them into attitudinal and factual (or equivalently psychological and external) to the respondent. The distinction between the psychological and external is rather clear in most cases, the latter being those that are either reported from another source (such as the index of multiple deprivation of the household or the neighbourhood condition, as assessed by the interviewer) or reported by the respondents where they are dealing with a factual matter. Whether the person's parents have split up/divorced is a case in point – although answered by the respondent, there is less scope for psychological variation (though of course the respondent could deliberately mislead the surveyor or be confused) and this is best placed as an 'external' factor. The factual or external questions are not the subject of this chapter and will be considered as controls in Chapter 9; what is to be discussed here is the role that the internal attitudinal questions play and how to fit them into the analysis.

Figure 7.2 Covariance, mediation and moderation SEM path diagrams



To that end there are three possible ways as to how these internally perceived attitudinal questions fit together; covariance, mediation or moderation. SEM path diagrams representing each of these possibilities are displayed (respectively) in Figure 7.2a), b) and c) (assume the X and M are independent variables and Y a dependent variable – whether latent or continuous is not a necessary piece of information for this discussion). The familiar method of covariance is exactly the same as adding a variable into a 'standard' regression technique. As such it does not allow a distinction between

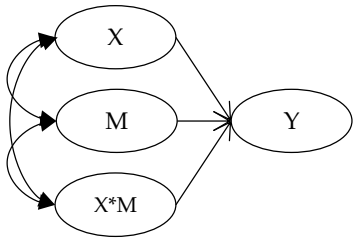
the psychological versus external, attitudinal versus factual, components as defined above. This also does not make sense theoretically when applied to the concepts in this work. Each of the theories in the previous section considered there to be some level of stability in individual attitude *and* some level of (varying) interaction of the individual and the environmental setting in which they find themselves at any one time. This interaction is explicitly seen in all the theories to be a dynamic process across a population. Wikström sees it as the absolute cornerstone of his SAT (though focuses much less on the attitudinal perceptions of the environment, rather on quantifiable external measures of it), Sampson and Laub's varying social controls demonstrate this change in perceptions, as does Moffitt's gap between biological and social maturity leave space for perceptions of environment to alter based this individual attitude. To define the interaction as a simple covariance would be inappropriate, as it would not allow the variation in the perception of the attitudinal measures of the familial and area level components, by the individual's criminogenic propensity, to be controlled.

Turning attention to the alternatives and specifically whether the interaction is a moderation effect, it is helpful to step away from the usual SEM path representation of a moderation effect and describe it as it usually referred to in other disciplines, namely an interaction effect of variable X with variable M. This can therefore be represented in an alternative manner displayed in Figure 7.3 (as used by Baron and Kenny, 1986). Using this visualisation it is clear that, given that with SEM a realist ontology is enforced (see section 5.4.1), this is in essence describing a situation where there is *another* latent variable that is within the individual that is a combination of both X and M[25]. Applied to the specific concepts being worked with here, it would mean that there is another component and the criminogenic propensity of the individual combines with their feelings about, say, their family environment and this is a thing in and of itself, its own latent construct separate from all others. This situation seems particularly unlikely:

[25] This is testable and can be achieved by creating a measurement model with three latent components and with all the manifest variables that are caused by X to associate only with X, those caused by M with M and the product of each and every one of the X manifest variables multiplied by each and every one of the M manifest variables to associate with the X*M term only and checking the fit of the model. This has not been reported in full in this thesis due to absurdity of the arguments needed for it to be viable, but unreported analysis found models such as this not to fit.

here it would mean that someone's answer to the question "sometimes it is okay to break the law" would be caused by one factor (their individual attitude), "My parents usually praise me when I have done well" is caused by another (the actions of their parents, as they perceive them) and the combination of the answers to both of these would be caused by something else entirely!

Figure 7.3 Alternative representation of a moderation effect



This leaves mediation as the only possible alternative but that is not to say that it is merely accepted as a default, through exclusions of others. It is possible to make this argument hold even without the treatment and rejection of the two possible alternatives. Baron and Kenny's (1986) oft-cited investigation into moderation and mediation effects is replete with observations that imply that mediation is the correct specification when dealing with psychological constructs such as those being discussed here under the banner of internal, or attitudinal, concepts. For example, "Mediators explain how external physical events take on internal psychological significance. Whereas moderator variables specify when certain effects will hold, mediators speak to how or why such effects occur" (p1176) and "For research oriented toward psychological levels of explanation (i.e., where the individual is the relevant unit of analysis), mediators represent properties of the person that transform the predictor or input variables in some way. In this regard the typical mediator in cognitive social psychology elaborates or constructs the various meanings that go beyond the information given" (ibid., p1178). The usage of area level perceptions even form a key analytical component of Newman's (1972) work on territorial constructs in his defensible space thesis, a fact highlighted by the Baron and Kenny work. Although the statistical maxims of the Baron and Kenny piece have been somewhat superseded (see for example Mackinnon 2008 and the discussion in Chapter 5), these basic assumptions about the nature of the difference between mediation and moderation have not.

Mediation, therefore, is the only way to treat the interactions of these perceived attitudinal responses to the familial and area (and indeed any other components that could be added) level questions. It seems, then, that placing the criminal propensity at the 'beginning of the causal chain' and for it to cause variations in the situational responses seems not only sensible but indeed necessary for a model of criminal behaviour to be built from this dataset.

7.4 Starting Steps: Testing Each Domain

Just as the first step of structural equation modelling is to form well-fitting measurement models, starting with the smallest components and building up so that each latent domain is well defined, it is similarly sensible to build the structural steps forward in component parts, analysing the relationship of each domain against the dependent and moving forward to add together theoretical components sequentially. This process of analysing each domain individually will also strengthen the claim that they are measuring that which the measurement models chapter purports they ought to be measuring, i.e. relevant predictors of offending involvement.

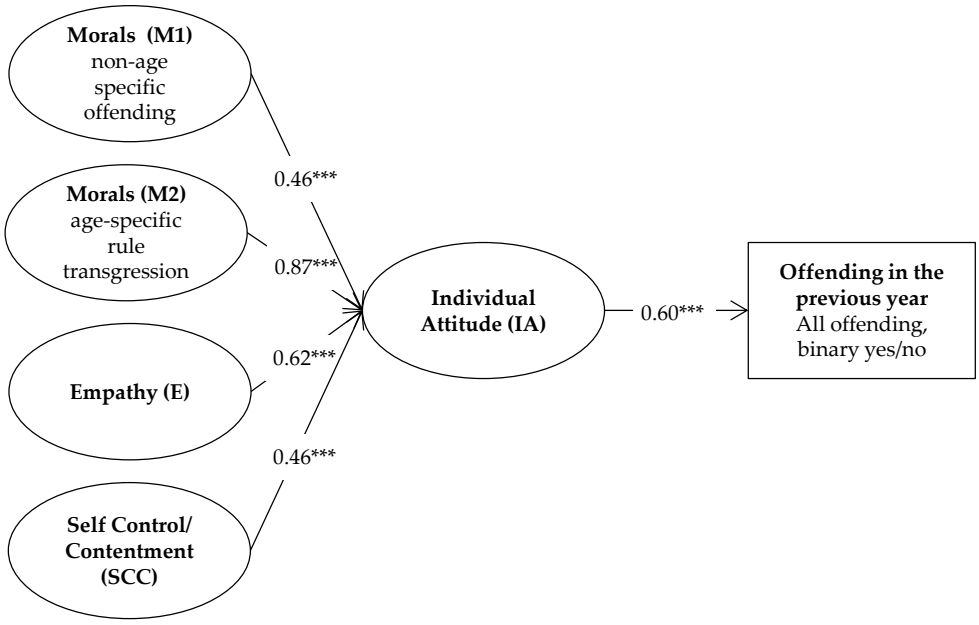
In that this process is cumulative and the measurement models have been specified in previous analysis, these will not be called into question and only the structural parts will be concentrated on (that the measurement models do not vary will be checked, but this is unlikely under the assumption that they are well defined, see Bryne, 2012). To aid presentation, therefore, the manifest variables that contribute to these measurement models will be left out of all the model diagrams in this chapter. Finally, throughout this chapter the statistical significance of any loadings/regression coefficients will be suffixed by asterisks relating to the level of significance as *** = Sig at 99.9%, ** = sig at 99%, * = sig at 95%.

7.4.1 Individual Attitude

The relationship between offending and individual attitude, defined in the previous chapter and taken to represent an individual's criminal propensity, is relatively straightforward to analyse. Simply incorporating a dependency to offending as pictured (with results) in Figure 7.4. This diagram shows that the loading from the first

order latent parts to the second order individual attitudes component is broadly the same as in the measurement models, implying that this a reliable model. Fit statistics, although less important in the modelling phase (see section 5.6), are more than satisfactory with a CFI of 0.975, TLI of 0.968 and RMSEA at 0.051 with a 90% confidence interval of 0.049 to 0.057.

Figure 7.4 Offending and individual attitude structural model



In a rough sense the coefficient of 0.60 between individual attitude and whether the respondent reported offending in the previous year (or not) shows that a more criminogenic individual (a higher score on individual attitude) is strongly related to an increase in self-reported offending prevalence. Were both individual attitude and the offending dependent variable continuous, normally distributed and standardised (rather than just the individual attitude variable) this would be simple to interpret – a one standard deviation rise in individual attitude (causing, as it would be, a 'higher score' in some combination of the answers to the morality, empathy or self-control suite of questions) would cause a 60% of a standard deviation rise in the number of offences committed. Similarly, if it were possible in Mplus to employ a logit link function with the WLSMV estimator and categorical data, then the 0.60 would be a log odds ratio, and the resultant odds ratio of 1.8 would imply that a 1 standard deviation rise in the individual attitude would imply an 80% increase in self-reporting offending.

With categorical data and the WLSMV estimator (as required for robust standard errors) Mplus uses a probit link function and although time consuming, it is possible to be more precise in terms of the predicted probabilities of having reported offending As the individual attitude variable is standardised and has no meaningful scale of its own, it is only possible to work in terms of standard deviations above and below the mean value (in this case a score of zero). Using the steps outlined in section 5.7, converting to probabilities gives the results shown in Table 7.1.

Table 7.1 Predicted probabilities of offending at various standard deviations from mean score of individual attitude

IA standard deviation change	**-2**	**-1**	**0 (mean)**	**1**	**2**
Predicted probability	3.9%	12.2%	28.5%	51.2%	73.5%

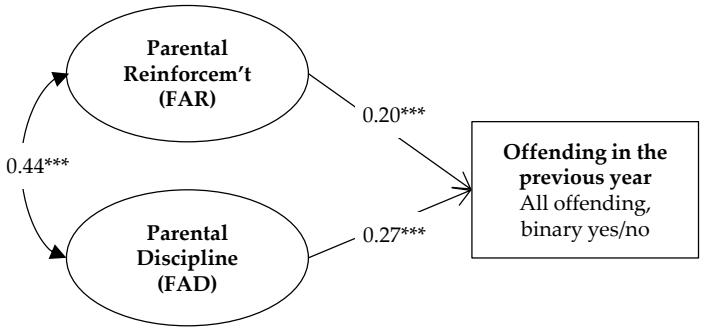
From this table it is clear to see the importance of the individual attitude in terms of the expected rate of reporting of offending. A zero score on individual attitude represents the mean value and equates to a predicted probability of reporting offending of 28.5%, the same as the average over the whole sample of 10- to 16-year-olds. Those with an individual attitude score two standard deviations below average have a predicted probability of reporting offending of just 3.9%, whilst of those two standard deviations above the mean the probability increases to 70%. For the group one standard deviation above the mean, around 51% are expected to report offending, a 23 percentage point, or an 80%, increase. In this latter predicted probability the interpretive link to the interpretation of a logit link (a logistic regression) can be seen. As noted previously, a coefficient of 0.6 would imply an odds ratio of 1.8 and this ties in with the results of a probit regression. This observation is due to the similarity between logit and probit models in the centre of the distribution (Liao, 1994) and can used as a rule of thumb approximation where the calculation of exact probabilities is unwarranted detail.

7.4.2 Family Situation

A first analysis of the family level variables is also straightforward and is displayed in Figure 7.5. Fit is excellent (CFI = 0.97, TLI = 0.96, RMSEA = 0.034) and there is a relatively strong correlation between levels of perceived parental discipline and levels reinforcement. These items are coded such that a higher score represents a lower level

of positive reinforcement or a more lax disciplinarian regime and hence the positive regression coefficients onto the offending dependent are expected.

Figure 7.5 Offending and family reinforcement/discipline structural model



A more specific understanding of the effects of these latent variables is possible using the techniques applied to the individual attitude latent variable. As the dependent is the same as previously, it is possible to predict that the coefficients of 0.20 and 0.27 (respectively for reinforcement and discipline) should have a lower effect on offending than the individual attitude model, and using the approximation of the logit/probit equivalence, one would expect a one standard deviation increase in poor reinforcement to increase the likelihood of reporting offending by around 22%, a similar rise in lax discipline to increase the offending rate by 30%.

The predicted probabilities, when computed exactly, can be seen in Table 7.2 and confirm the tentative conclusions drawn above. Lower levels of reinforcement do seem to be associated with higher offending, but the effect of a one standard deviation change is less than the effect for the same sized change in discipline. The correlation between the two factors, however, implies that on average a respondent reporting a lower level of reinforcement will also report a lower level of discipline.

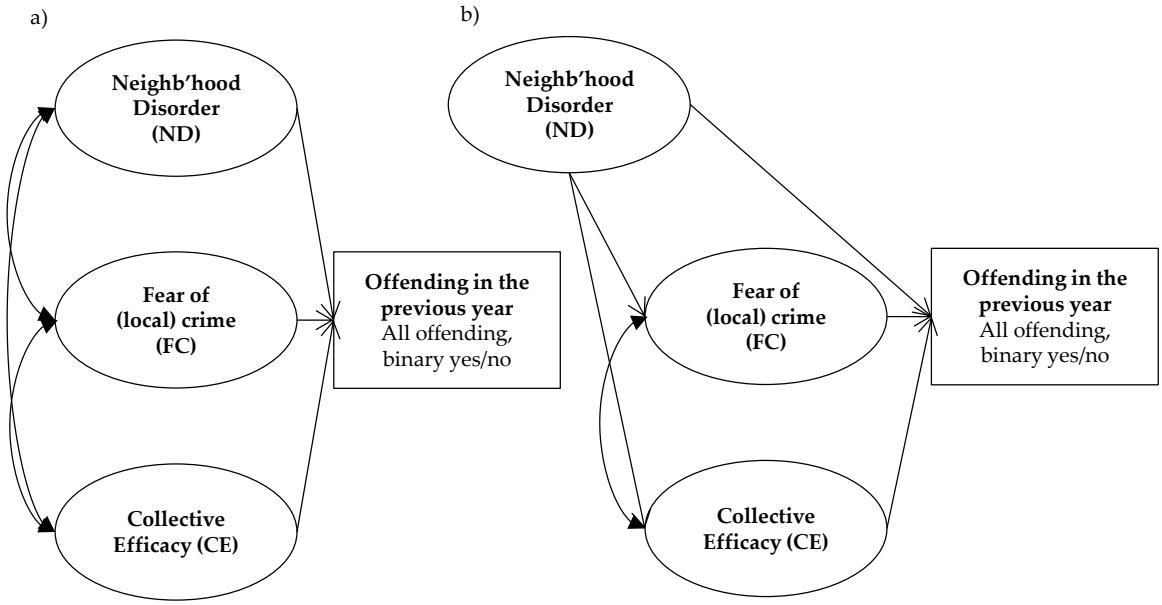
Table 7.2 Predicted probability for changes in parental reinforcement or discipline

Standard deviation change	**-2**	**-1**	**0 (mean)**	**1**	**2**
Parental reinforcement	16.7%	22.2%	28.5%	35.7%	43.4%
Parental discipline	13.6%	20.3%	28.5%	38.1%	48.5%

7.4.3 Area Character

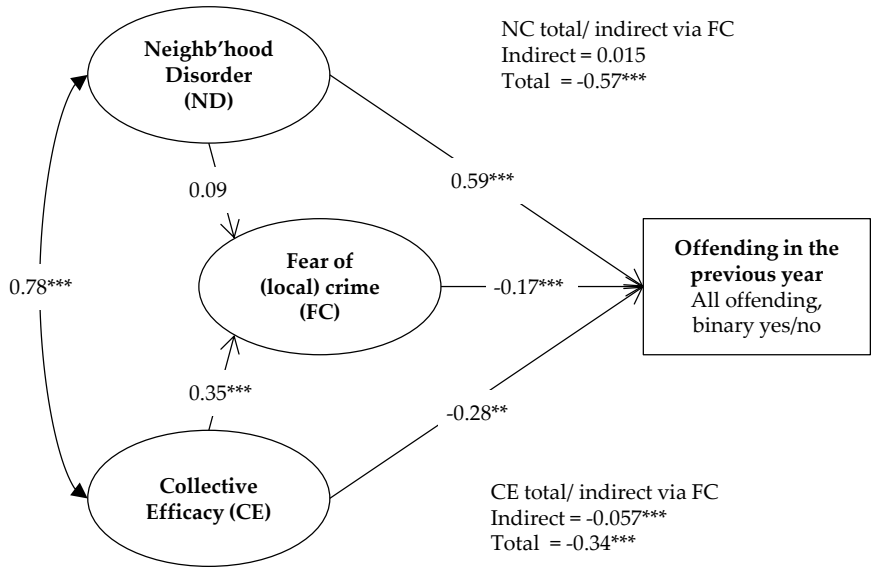
Testing the construction of the area character variables is more complex, and requires more careful consideration of underlying latent variables, as the manner in which the three components fit together is not self-evident. Taking as a starting point the three latent variables capturing self-reported perceptions of area as defined in Chapter 6, the first problem to address is whether the neighbourhood disorder latent construct co-varies with fear of crime and collective efficacy, or is causal on an individual's fear of local crime/collective efficacy, shown respectively as path diagram a) and b) in Figure 7.6. Wilson and Kelling's (1982) broken windows hypothesis claims that poor neighbourhood condition is associated with an increase in the likelihood of incivilities in an area and this in turn leads to poorer community relations, a higher fear of crime and a greater criminal involvement for those who live in such areas. Taking this as the basis for the dependencies would imply the configuration in diagram b). However, this makes two mistakes. The first is that the broken windows hypothesis operates at a neighbourhood level. It implies that a *neighbourhood* with high disorder would have associated higher average levels of fear of crime and worse average perceptions of community relationships at a *neighbourhood* level; it does not make a strong claim about how these mechanisms would work on an *individual* level. Although there is a natural expectation that if the average were lower, then the individual within the area would also have a lower value, this is not necessarily the case. The second issue, as highlighted by Sampson and Raudenbush (2001), is that "disorder does not directly promote crime, although the two phenomena are related, and collective efficacy is a significant factor in explaining levels of crime and disorder" (p2).

Figure 7.6 Alternative causal paths for neighbourhood level relationships



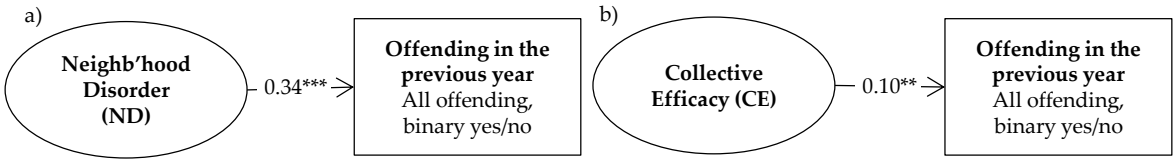
Overall, therefore, it seems that if one of the two definitions were to be chosen, the second is clearly incorrect. The first diagram (option a) in Figure 7.6 could be considered more accurate and, with all variables co-varying, places fear of local crime on the same footing as neighbourhood disorder and collective efficacy. However, this goes against Sampson and Raudenbush's formulation and is inconsistent with recent work looking at neighbourhood level processes (Brunton-Smith et al., 2014). These place neighbourhood disorder and collective efficacy as co-varying and together predicting beliefs and worry about crime. The correct specification therefore would most likely look like that displayed in Figure 7.7 and when estimated fits the data well (CFI = 0.99, TLI = 0.98 and RMSEA = 0.04). The effects of neighbourhood disorder and collective efficacy on fear of crime are partly consistent with the Brunton-Smith et al. (2014) paper: consistent in that a lower level of collective efficacy increases fear of crime (demonstrated by the positive regression of fear of crime onto collective efficacy, which is defined so a higher score means a more criminogenic setting, i.e., less trust and informal social control) and this effect is stronger than that seen for neighbourhood disorder, inconsistent as the neighbourhood disorder falls just outside standard significance levels, at 92%, and is significant in the Brunton-Smith paper. It should be remembered though that the Brunton-Smith paper uses a multi-level approach and the results are again at the neighbourhood (lower super output area) level.

Figure 7.7 Revised neighbourhood level relationships



Moving beyond looking just at the effects on fear of crime, and looking to see how fear of crime impacts offending behaviours, an interesting association appears. A lower level of fear of crime appears to increase offending, perhaps due to the respondent's increased confidence with their local area but it is in this implication that the work runs into difficulty. One could certainly make an argument for a mechanism from lower fear to higher offending (increased local confidence based on an understanding of the criminal dynamics in the area due to involvement in such behaviour?), but there does not appear to be any well tested and hypothesised mechanism in the literature that can be applied here. Fear of crime is almost universally treated as a (often multi-faceted) dependent variable in and of itself and a great body of work has investigated its behaviour (Hollway and Jefferson, 1997; Jackson and Gouseti, 2014). Unfortunately though, its connection to offending behaviour is predominantly missing. Although an interesting avenue for further work, this current piece is not the place to investigate such a relationship. SEM should be used for hypothesis testing, not generation, hence despite this current configuration giving credence to the measurement models defined in the previous chapter (through its basic consistency with fear of crime analysis), it is not the best setting to investigate fear of crime further. The definition of a fear of crime latent construct certainly took little effort and could be fruitfully investigated in another work, but it is best left as an open analytical question for the future and not carried forward any further.

Figure 7.8 Offending regressed on area perceptions



Deletion of fear of crime from the model leaves just two perceptual area components to be considered, neighbourhood disorder and collective efficacy. Regressing self-reported offending as a dichotomous variable onto each of these individually gives the results as displayed in Figure 7.8. In both cases the area characteristic behaves as would be expected, lower levels of collective efficacy (i.e. a higher score on the variable) and higher disorder both predict higher offending. The relationship between a rise of one standard deviation in perceptions of neighbourhood disorder is considerably stronger than that seen with a similar one standard deviation in more criminogenic collective efficacy perception (a lower reported level), as demonstrated by the predicted probabilities shown in Table 7.3.

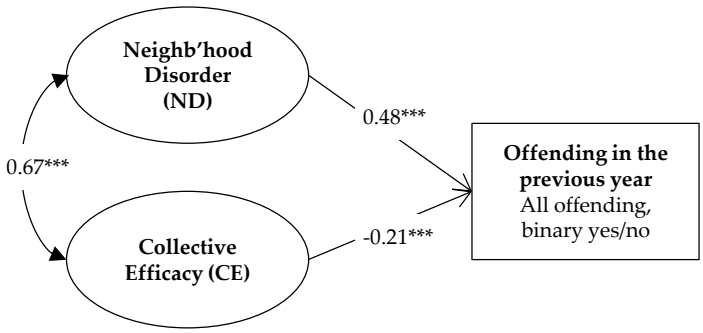
Table 7.3 Predicted probability of offending for changes in neighbourhood disorder and collective efficacy (modelled separately)

Standard deviation change	**-2**	**-1**	**0 (mean)**	**1**	**2**
Neighbourhood disorder	10.7%	18.2%	28.5%	41.0%	54.4%
Collective efficacy	21.9%	25.1%	28.5%	32.2%	36.0%

Combining the analysis and adding both of these latent variables into a combined regression and hence allowing neighbourhood condition to co-vary with collective efficacy leads to a rather simple structure being estimated, very much akin to Figure 7.5, with collective efficacy and neighbourhood condition in place of discipline and reinforcement (Figure 7.9). The fit is good (CFI = 0.98, TLI = 0.99, RMSEA = 0.025) but a confusing pattern emerges (as was first seen in Figure 7.7, but was left unmentioned due to more pressing concerns). Perceptions of neighbourhood disorder are strongly linked to collective efficacy via a strong correlation, with an individual reporting higher disorder very likely to report low collective efficacy (as seen in the previously defined and discussed measurement models). Higher neighbourhood disorder still appears to increase self-reported offending, but lower perceived collective efficacy appears to *decrease* offending! Using the logit rule of thumb method, neighbourhood disorder

appears to add approximately 60% to the offending rate, whilst a lowering of collective efficacy of one standard deviation appears to cut offending by approximately 20%. This is not only counter-intuitive from a theoretical perspective (lower collective efficacy at the very least has been shown to have no effect on offending rates), but it is also at odds with the mechanisms implied in the regression with just collective efficacy as an independent variable (in Figure 7.8 and Table 7.3).

Figure 7.9 Reduced neighbourhood level relationships



Speaking from a statistical perspective, this is a case of a suppression effect (Maassen and Bakker, 2001; Conger, 1974), a continuous equivalent of the better known categorical equivalents of Simpson's (1951) and Lord's (1967) paradoxes. Tu, Gunnell and Gilthorpe (2008) investigate these apparently disparate paradoxes, note parallels between all three and coin the umbrella term *reversal paradox*, before moving on to look at the implications and how to deal with them. The culprit most basically charged for this is multicollinearity between the independent variables and the standard methodology is to analyse their dimensionality using a factor analytical technique and, depending on the results, either combine the two or delete one or the other offending independent variables (Tabachnick and Fidell, 2007). Neither of these approaches is suitable here, the first because by design (in the previous chapter) these two concepts are not reducible to a composite, the second as deletion would go against that which is being aimed for, as Freidman and Wall (2005) nicely explain:

> "Moreover, it is reasonable to consider highly correlated independent variables. One might choose a measure of arm strength in testing mountain-climbing skills. A measure of leg strength is an obvious further predictor. Arm strength and leg strength are probably highly correlated: however, it is hard to believe that either one alone would

> predict mountain climbing skills. The researcher needs to consider the substantive interplay of all variables with each other" (p135).

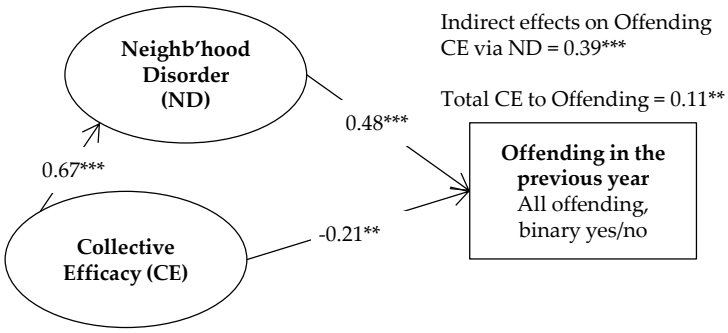
Maassen and Bakker (2001) analyse various different forms of the suppression effect; 'classical' (where a variable uncorrelated with the dependent increases the predictivity of another independent), negative (as here, where two independent variables have positive zero-order correlations with the dependent and correlate positively with each other, but one receives a negative regression coefficient when modelled together) and reciprocal (where two good predictors cancel out each other's effect through a negative correlation), and show that the probability of their occurrence is relatively high when working with latent variables. Interpretational problems go hand in hand with the occurrence of these effects, but "suppressor variables… usually unintentionally, can indeed play an important role. It is important to acknowledge a suppressor structure… [and]… after recognizing the phenomenon, one cannot discard suppressors when interpreting the results" (ibid., p243).

Once it has been accepted that the suppressor effect is a true symptom and not reducible via collinearity, Tu, Gunnell and Gilthorpe (2008) state that "these paradoxes [their reversal paradox umbrella term] show that prior knowledge and underlying causal theory play an important role in the statistical modelling… where incorrect use of statistical models might produce consistent, replicable, yet erroneous results" (p1). Arah (2008) followed up the Tu et al. analysis and asks that if "the statistical phenomena they [the paradoxes] purport to represent are in fact causal in nature, requiring a causal language not a statistical one, and that the problem can be resolved only with causal reasoning… why bother with the statistics of these paradoxes, much less their equivalence… if both the correct language and resolution lie elsewhere?" (p2). Although there are many statistical techniques that can be used to understand suppression effects, as listed by the Maassen and Bakker (2001) piece, these are useful for identification of the type of effect (classical, negative, reciprocal etc.) in various different settings (regressions, path analyses, repeat longitudinal measures). With neighbourhood disorder and collective efficacy the effect is clearly a negative suppression manifestation and therefore as Arah (2008) states, and in keeping with the

major theme of this work, "explanations and solutions lie in causal reasoning which relies on background knowledge, not statistical criteria" (p5).

The substantive implication of the discussion is that either there is a common cause of the neighbourhood disorder and collective efficacy conflation or that the causal ordering of the two components should be revisited. Limiting the discussion for the moment to the second option and constraining the debate to just area level responses (the first option will be discussed in section 7.5.2), when revisited in this constrained manner the relationship between neighbourhood disorder and collective efficacy is clear and diametrically opposite to the mechanism discussed and rejected in Figure 7.6b), that collective efficacy is affected by perceptions of neighbourhood disorder. Perhaps one of the principal findings of recent collective efficacy work at an area level to date has been that both structural features *and* collective efficacy will contribute to neighbourhood crime *and disorder* (Sampson and Raudenbush, 1999; Reisig and Cancino, 2004). The constituent components of the neighbourhood disorder latent variable (people being rowdy and drunk, people selling drugs, people sleeping rough and noisy neighbours) fit the definition of this second consequence (disorder) of poor collective efficacy perfectly: the representation of disorder should be considered an effect of the varying levels of collective efficacy (amongst other components) and not simply to co-vary. From an even more basic survey administration perspective, the question set for the disorder component was asked directly after the collective efficacy set, so the respondent would have been primed to think of the disorder in terms of the neighbourhood characteristics they would moments earlier have been considering. It is therefore more than sensible to re-specify the causation between collective efficacy and neighbourhood disorder so that neighbourhood disorder is regressed *onto* collective efficacy and collective efficacy therefore operates on an individual's self-report offending both directly and indirectly through neighbourhood disorder. This configuration is shown in Figure 7.10.

Figure 7.10 Adjusted neighbourhood level relationships



When considered in such an alternative system the fit is equivalent to the specification in Figure 7.10 as the only difference is that the covariance between the two independents has been changed to a regression. This system introduces an indirect effect into this work for the first time, from collective efficacy via neighbourhood disorder and is computed by taking the product of the regression of collective efficacy on neighbourhood disorder with the neighbourhood disorder on offending path. With an indirect effect, the total effect of collective efficacy must also be computed and is simply the sum of the direct and the indirect. Mplus holds the functionality to compute this automatically, with standard errors and hence significance levels and the resultant values are displayed on Figure 7.10 as text (and labelled appropriately). The direct effect of reduced collective efficacy still implies a lowering of offending, but the indirect effect operating through neighbourhood disorder is greater, meaning that the total effect is that lower perceptions of collective efficacy imply higher offending with a probit regression coefficient of 0.11, significant at the 99% level, and similar in magnitude to the coefficient assigned to collective efficacy in the second diagram in Figure 7.8 (the model with just collective efficacy). Neighbourhood disorder has an r-squared of 44.7% showing that in this modelling framework collective efficacy is explaining just under half of the variability of the latter component, a strong relationship.

The specific interpretation of these loadings proceeds in much the same way as it would without the indirect effects. For neighbourhood disorder the predicted probabilities are the same as those that would have been produced from the model outlined in Figure 7.9 (via its direct effect) and for collective efficacy, as the

dependencies of the reported offending work in both a direct and indirect manner, the coefficient of interest is the combination of the two – the total effect. The results in Table 7.4 are consistent with those that would be expected from the regression coefficients and show neighbourhood disorder displaying an effect a little lower than when modelled separately and with collective efficacy showing a broadly similar effect size (comparisons made to Table 7.3).

Table 7.4 Predicted probability of offending for changes in neighbourhood disorder and collective efficacy (modelled separately)

Standard deviation change	**-2**	**-1**	**0 (mean)**	**1**	**2**
Neighbourhood disorder	6.4%	20.3%	28.5%	38.1%	48.5%
Collective efficacy (total effect)	21.7%	25.0%	28.5%	32.3%	36.2%

7.5 Combining Two Domains

The effect of each domain's measurable components on offending separately does not address the central problem that different types of people are found in different settings. To rephrase this statement to be in line with the data available in the present study more specifically, it does not address the issues of whether different types of people, as defined by their individual attitude (criminogenic propensity), *report* their settings in a different way. To assess this, the relationship between individual attitude and offending needs to be analysed with mediating effects for the family level latent variables and those representing the area (as explained in section 7.3). More complex models need to be specified and consideration of direct and indirect effects need to be in place from the very beginning. This section, however, will focus on defining the appropriate interrelations of the components and less on the substantive implication, which will be left until the model bringing all three domains together.

From this stage on, the lower level first order latent variables that are behind the individual attitude variable are to be left out of any diagrams as they do not play a part in the model beyond being influenced by the second order component.

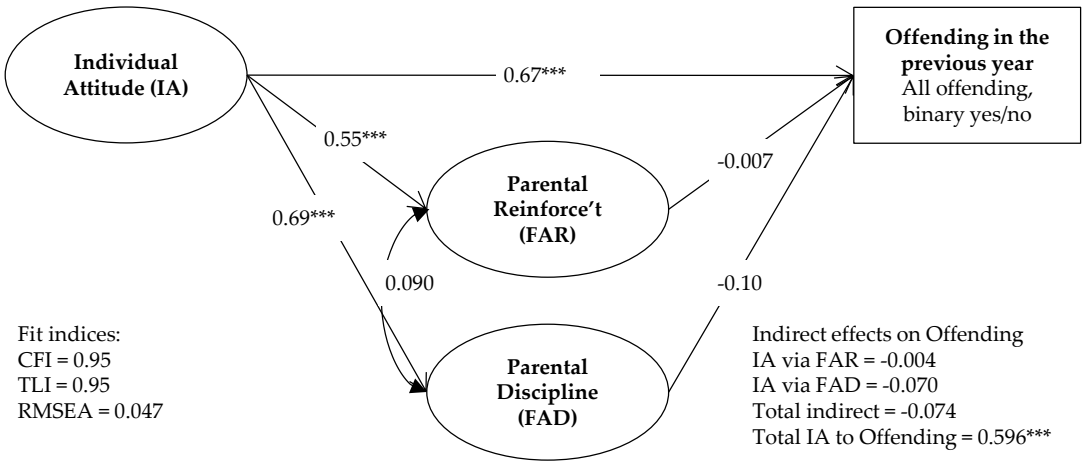
7.5.1 Family Level with Controls for Individual Attitude

Adding the family environment as a mediating effect between individual attitudes is shown, with estimated loadings, in Figure 7.11. The regressions of parental discipline and parental reinforcement control for any potential difference in response based on the individual attitude of the respondent as well as allowing the mediation of these components to be assessed. This could express itself in two ways: the first, simply that discipline/reinforcement is in actuality different dependent on the individual attitude latent variable, which has been discussed and taken to be stable. The second is one of whether the individual *perceives* their family environment differently dependent on their individual attitude. The first scenario is the strongest, and will give more concrete results. The second is weaker, implying only a difference of individual view, but there is no way to separate between the two positions in the current work. If the first were to be seriously called into question then this would have rather far-reaching consequences, particularly for surveys within psychology. With adequate survey design it must be hoped, at least, that this would not be the case. Regardless, to couch any conclusion in the most hesitant language is suitable and even under the conditions of the second possibility clear implications can still be arrived at and in all probability, the reality is somewhere between the two and it is likely the individual attitude influences both the behaviour of the parents and the individual's perceptions of their behaviour.

With that in mind, the results of the model in Figure 7.11 produce a mixed bag of results. It is clear that the responses given to questions about reinforcement and discipline are heavily influenced by individual attitude. Whether this is due to perception or actual difference in behaviour is unknown, but it is certainly a possibility that those with more criminogenic personalities report lower levels of parental reinforcement. In terms of disciplinary regimes, the more criminogenic the individual the lower the discipline appears to be. Again this is a logical finding, as there is likely a strong year-on-year continuity in parental discipline and a poor environment in childhood is seen as an important driver in creating the individual attitude in the first place (this will be important later). It is to be noted that this does not imply that the

relationship is reciprocal, or that the causal path be reversed, merely that the family environment is consistent at the current time with that experienced during the development of the (stable) individual attitude. The correlation between the two factors has reduced to an insignificant level compared to the uncontrolled model in Figure 7.5, showing that the relationship between these latent variables is entirely explained by the individual's personality.

Figure 7.11 Individual attitude mediated by family reinforcement and discipline



In terms of predicted probabilities from the model (Table 7.5) and when varying individual attitude, these are not statistically different to the values predicted in the uncontrolled model. This is a rather redundant finding as the effects of reinforcement and discipline are now insignificant and hence no change should be apparent. Similar analysis of a change in the reinforcement or discipline level, although showing slight differences (not reported) are not of interest as, again, these differences are statistically probable.

Table 7.5 Predicted probability of offending for changes of individual attitude, controlling for family discipline and reinforcement

IA standard deviation change	-2	-1	0 (mean)	1	2
Predicted probability	3.9%	12.2%	28.5%	51.2%	73.4%

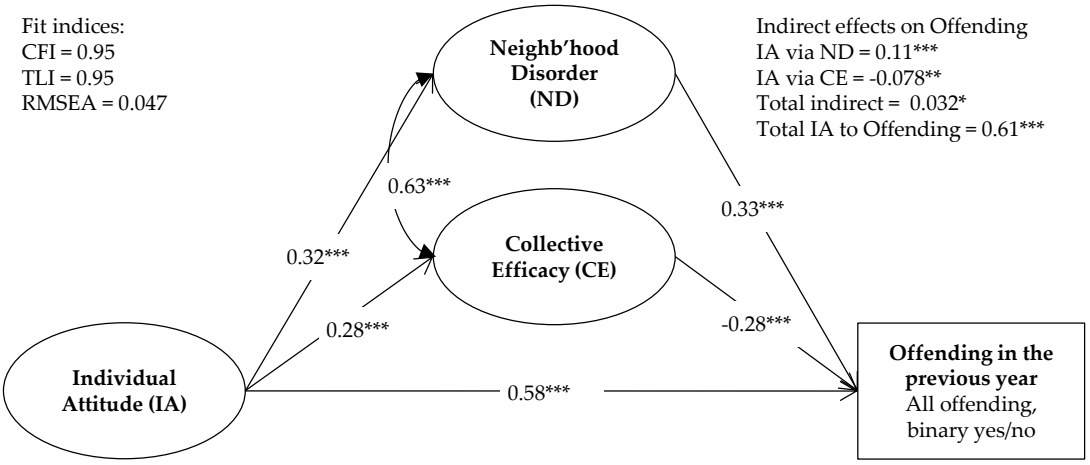
Although not statistically significant, the regression coefficients of offending on reinforcement and discipline appear in the opposite direction to those given in the analysis section 7.4.2 and therefore represent a (non-significant) reversal effect. This time, however, the reason for this lies in the common cause of the levels of

reinforcement, and discipline is found to be the individual attitude of the respondent, as demonstrated by the reduction to zero of the covariance between the two components. In short this implies that the apparent reversal of the effect of the two mediators is caused by the fact that the level of perceived reinforcement/discipline is highly dependent on the individuals' attitude. This point shall be returned to in depth in Chapter 10.

7.5.2 Area Variables with Controls for Individual Attitude

Adding individual attitude as a control for the effects of collective efficacy and neighbourhood disorder on offending with respect to both of the specifications discussed in the previous section (shown in Figure 7.12 and Figure 7.13) allows an analysis as to whether the individual's attitude can be considered as a common cause of both of the response variables. If this is the case then the covariance between the two components should be reduced and, more importantly, the reversal of the collective efficacy regression coefficient onto the offending variable should disappear. This specification can be seen, with estimated path coefficients in Figure 7.12, and it is immediately clear that this is not the case. Although the estimated model does shows that the response to the neighbourhood level questions do vary significantly by individual attitude the reversal of the collective efficacy component is still present. This of course does not rule out the possibility that there is another common cause that is unmeasured and inaccessible, and is causing the reversal but this shall have to remain a 'known unknown'. Coupled with the argument in section 7.4.3 of the seemingly sensible causal chain, with perceptions of disorder as being dependent on the respondents' understanding of the collective efficacy latent variable, then this specification again is out of place and probably best left as 'incorrect'.

Figure 7.12 Individual attitude mediated by collective efficacy and neighbourhood disorder



Nevertheless, calculation of the predicted probabilities in this model, with mediating effects between the individual attitude components and the offending variable can be carried out, though slightly more complicated than in the simple cases when only modelling one domain. Computing the predicted probabilities for movement on the individual attitude latent variable still uses the total effect, but it is with the mediating variables that differences arise. To simplify the discussion, consider just neighbourhood disorder and its role in the model. There are two possible ways to proceed; either calculating the approximate proportion of mediation as outlined in section 5.7 of the methodology chapter, but it is also possible to conduct something of a thought experiment.

Consider just the neighbourhood disorder latent variable. As this is regressed onto individual attitude it is not free to be fixed at a different value, it is dependent. Nevertheless, as the variability of this latent component that is explained by individual attitude is given by its r-squared value, reported in Mplus as 11%, it is still possible to imagine two sets of individuals who are one standard deviation apart on their perceptions of neighbourhood disorder, *with all other scores the same*. The neighbourhood disorder score therefore is the only component of the model that can impact the predicted offending rate and will depend only on the regression coefficient from neighbourhood disorder to the self-report offending in the previous year (ultimate) dependent. Calculations of predicted probabilities can thus proceed on this basis. Analysing the implications of this specification in terms of predicted probabilities

on this basis is displayed in Figure 7.6. It gives results for the movement of either individual attitude (calculated via total effects) or a change in either collective efficacy or neighbourhood disorder, *when holding all other components constant*.

Table 7.6 Predicted probability of offending for changes of individual attitude, collective efficacy or neighbourhood disorder

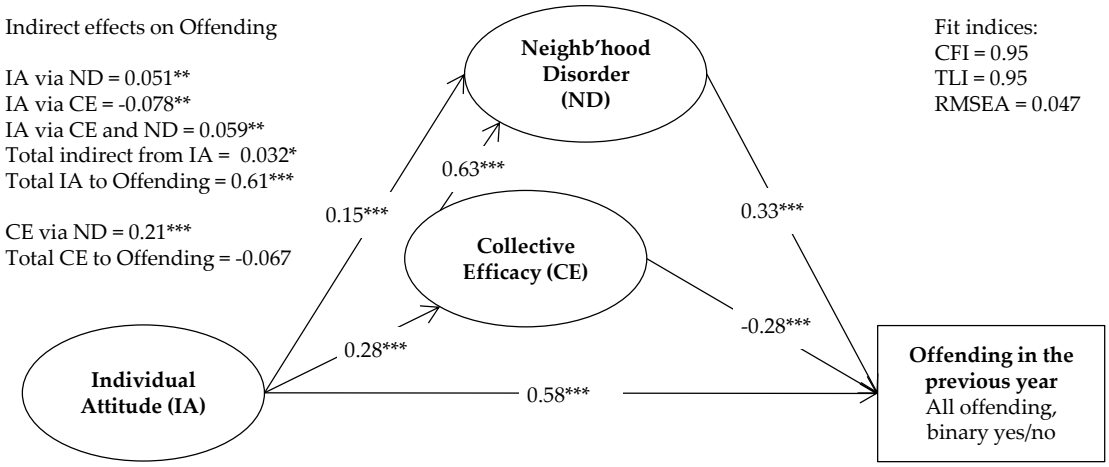
Standard deviation change	-2	-1	0 (mean)	1	2
Individual attitude	3.6%	11.9%	28.5%	51.8%	74.5%
Neighbourhood disorder	10.8%	18.4%	28.5%	40.8%	54.0%
Collective efficacy	49.3%	38.5%	28.5%	20.0%	13.2%

The effect of individual attitude is slightly amplified from those seen in Table 7.1, with the predicted probabilities of offending of those at the various standard deviations being more dispersed from the mean than previously. It is also clear that an increase on the collective efficacy components (so to a lower level) is working in the opposite direction to that which the bivariate model in Figure 7.8b) lead to expect.

Given the confluence of reasons that the Figure 7.12 specification is inappropriate, regressing the neighbourhood disorder latent variable onto collective efficacy is a sensible option and the model can be re-specified. This is displayed in Figure 7.13. Despite the coefficients attached to the arrows appearing to be the same as those displayed in Figure 7.12 it is in the direct and indirect effects that the true difference is seen. The highest coefficient is between individual attitude and offending, representing the centrality that this component of the model plays in determining whether the respondent offends or not, but (as with Figure 7.12) the answers given to the neighbourhood level questions vary significantly by this component, with more criminogenic individuals reporting that their areas have lower collective efficacy and higher levels of disorder. This latter dependency is split into two parts, some direct and some operating through decreased perception of collective efficacy (about equally). The direct effect of collective efficacy on offending is strongly negative, but this is cancelled out (as in the model with just the area components) by the causal path through neighbourhood disorder. Unlike in the model without criminogenic propensity the total effect of collective efficacy on offending is insignificantly negative, meaning that in and of themselves perceptions of collective efficacy do not alter self-report offending, all

other things being equal. The effect of individual attitude operating through the perceived neighbourhood disorder variable is positive, meaning that although those with a worse individual attitude report living in areas with worse disorder, this negative view of the area increases their offending propensity. The neighbourhood perceptions give a total indirect effect which is significant at the 95% level but it is in the fact that they seem to work in conflicting directions that will be returned to later.

Figure 7.13 Individual attitude mediated by revised collective efficacy and neighbourhood disorder specification



Again, computing the probabilities for people at different positions on the latent variables (using a similar type of thought experiment as that explained in the Figure 7.12 discussion) gives more specifically interpretable results which are displayed in Table 7.7 with the only difference from those displayed in Table 7.6 being that the collective efficacy variable set are reduced to being statistically insignificant from the mean value.

Table 7.7 Predicted probability of offending for changes of individual attitude, collective efficacy or neighbourhood disorder with revised causal structure

Standard deviation change	**-2**	**-1**	**0 (mean)**	**1**	**2**
Individual attitude	3.6%	11.9%	28.5%	51.8%	74.5%
Neighbourhood disorder	10.8%	18.4%	28.5%	40.8%	54.0%
Collective efficacy (total effect)	33.3% (ns)	30.9% (ns)	28.5%	26.3% (ns)	24.2% (ns)

(ns indicates that the results are not significantly different from the mean value)

The takeaway message from this discussion is not the substantive findings, these will come in the next section but that the rejection of the covariance between neighbourhood disorder and collective efficacy (as shown in Figure 7.12) is complete. The suppression effect noted in section 7.4.3 is not accounted for by the individual's attitude and the model specified in Figure 7.13 tells a story where the collective efficacy of the area has an insignificant direct effect on offending and it is only the perception of neighbourhood disorder that matters, with part of that perception being controlled by perceptions of collective efficacy. This second specification not only fits with the uncontrolled results more harmoniously, it is also theoretically sensible and the fitted probabilities have a greater face value.

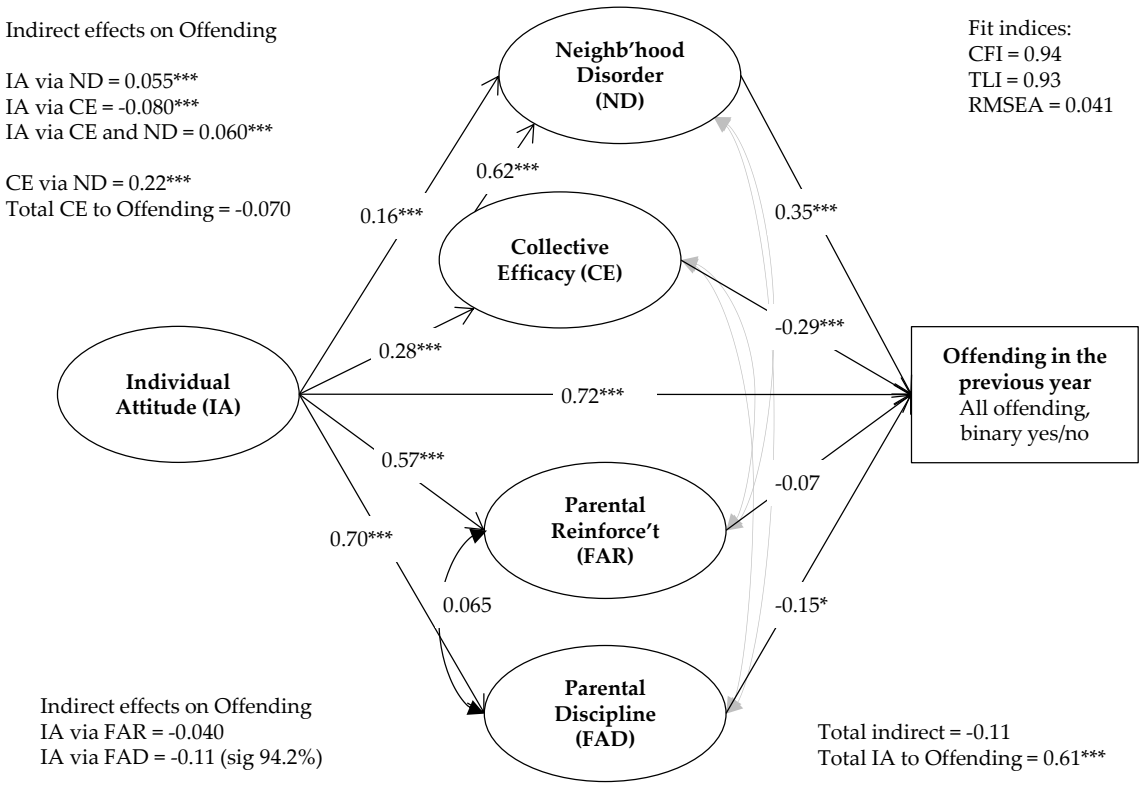
7.6 Modelling All the Domains Together

The final step of this first cut of the structural modelling procedure is to pull together all of the domains that have been analysed either individually or in pairs to create a model of the effect of individual attitude on (self-report) offending, with relevant mediating effects. This is not a final model in the sense that it is the pinnacle of the work, indeed there are another three analytical chapters to come, but it should be considered final in terms of modelling together the perceptual components that have been defined and included to this point. To recap, all of the latent variables in this model are at an individual level, with regressions from the individual attitude accounting for dependencies in response to the higher level latent parts, all of which were answered by the same individual. In that respect then, this should be considered as a *psychological map* of the social controls for/against offending as perceived by the individual. The model takes the form of a combination of Figure 7.11 and Figure 7.13 and is displayed along with estimated model parameters and fit statistics (good fit) in Figure 7.14.

There are four covariance values that are not displayed on this diagram (shown in light grey), which the Mplus defaults would naturally leave out. These were referred to in section 5.6 of the methodological chapter and are the covariance between the endogenous variables neighbourhood disorder/collective efficacy and the two familial components. These components happen to be insignificant in the current specification

and therefore their omission from the diagram does not matter as they do not impact upon the final interpretation of the estimated model. Their inclusion, however, does alter some of the other loadings when compared to alternative specifications where they are constrained to zero (which is not presented in this work). Preacher and Hayes (2008) are unequivocal when addressing the subject of multiple mediators; these covariances must be left in, as "collinearity plays a role in multiple mediation models in much the same way as in ordinary multiple regression" (p882). They continue stating that the fact that some mediators are insignificant, or that the total indirect effect across all mediators is not significant, does not necessitate their removal or invalidate the model. These types of findings can be substantively important.

Figure 7.14 Individual attitude modelled by perceptions of both area and family situation

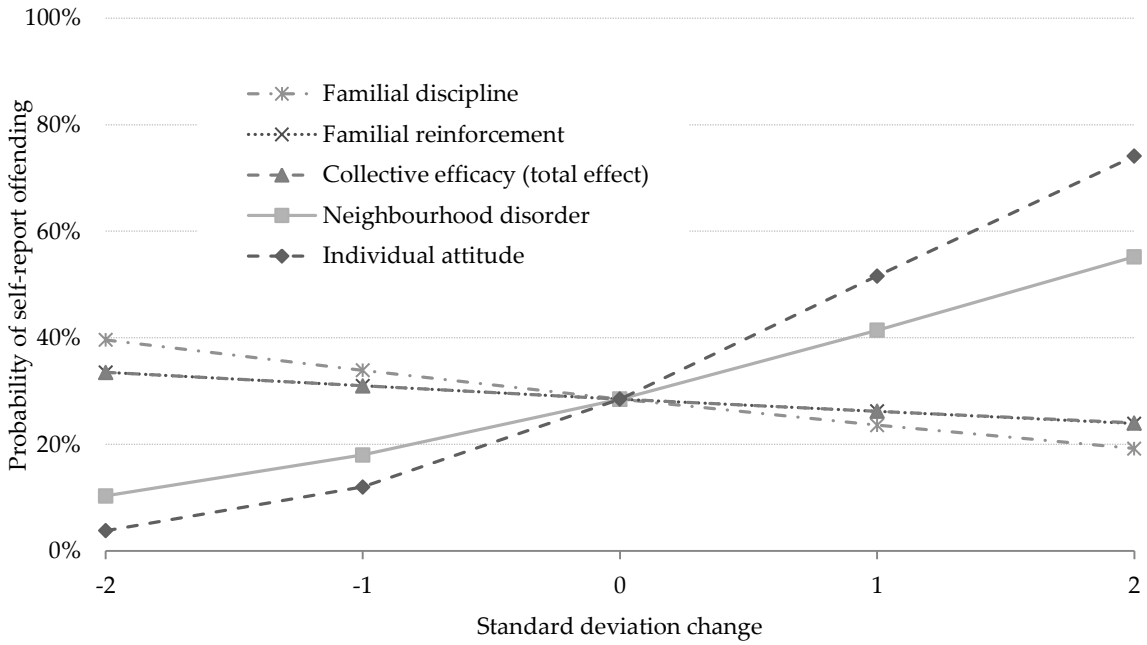


Each of the regressions of the mediating perceptions onto individual attitude are roughly the same as in the previous models (Figure 7.11 and Figure 7.13). Perception of parental discipline and reinforcement are most heavily affected (standardised linear regression coefficients of 0.7 and 0.57 respectively), with neighbourhood disorder (including the indirect effect through collective efficacy) and collective efficacy increasing by around a third of standard deviation for every standard deviation change

on individual attitude. The similarity of these coefficients is to be expected given that the covariances between the components of the two domains are insignificant (showing that the answers to these two sets of questions are relatively free from dependencies once individual attitude is a control).

As in the previous section, there are two ways to conceive of the impact of the mediating effects. One is to calculate the proportions of mediation as explained in the methodology section, the alternative captures the effect that would occur were it possible to compare respondents who are at different standard deviation levels on only the one component and all the others at their mean value (previously discussed in terms of a thought experiment, though for individual attitude also includes the indirect effects of the mediators). For consistency with the flow from the previous sections it is this latter option that shall be presented first and the results can be seen in Figure 7.15.

Figure 7.15 Predicted probability of offending in final model, at differing standard deviation levels of characteristic



The results are broadly the same as in the previous sections: the total effect of individual attitude still has the greatest effect on offending, with those two standard deviations below the average score on this latent variable only 4% are likely to report offending, whilst for those two standard deviations above, around three-quarters

would report offending. Neighbourhood disorder also plays a strong role in altering the predicted probabilities, whilst the total effects of collective efficacy and familial reinforcement are non-significantly different from the predicted probability of those with a mean value on the latent construct. Lower familial discipline, however, whereas previously insignificant in the model, now has a relatively weak direct effect on the propensity to self-report offending. The direction in which this works appears to be counterintuitive, and goes against the findings in section 7.4.2 in that those living in family environments with lower levels of perceived discipline are less likely to offend than those where the respondent reports that the parents often do not discipline them. The r-squared of parental discipline in the model is 49%, meaning that almost half its variation is explained by the individual attitude variable and this observation is an example of where the so-called thought experiment, which is considering respondents who are standard deviations different on just one component of the model, is stretched too far.

Calculating the proportions of mediation is therefore not only an alternative but the preferred method to understand the effects of these mediations, despite the acknowledged approximation that these are taking (section 5.7). These are displayed in Table 7.8.

Table 7.8 Proportion of mediation effect sizes in final model

Mediation via…		Proportion of mediation	Significance level
Area			
(1)	Collective efficacy only	-13.2%	***
(2)	Neighbourhood disorder only	9.1%	***
(3)	Neighbourhood disorder and collective efficacy pathway	9.9%	***
(4)	Any collective efficacy path	-3.3%	89.7%
(5)	Any neighbourhood disorder path	19.0%	***
(6)	Any area pathway	5.8%	n/a
Family			
(7)	Familial reinforcement	-6.6%	65.7%
(8)	Familial discipline	-17.7%	94.2%

The results of Table 7.8, coupled with the effects displayed on Figure 7.14, paint an interesting and complicated picture of the relationships that have been found. Firstly, the insignificance of the relationship (i.e., there being no significant covariance) between the answers to area-type questions and those related to the family environment show that, in this questionnaire and at cross-national UK level and using these measures, there doesn't appear to be a relationship between (perceptions of) area and (perceptions of) family environment. Of course the caveats of generalisability here are rather numerous and this is dealing only with a specific and small set of family and neighbourhood level concepts, but this nonetheless takes a stance against the commonly held perception that the reason that different types of area do not seem to affect the offending of an individual is due to family characteristics being sorted into neighbourhood via socioeconomic forces.

Rather, it appears that it is the tangled interrelations of the very areas themselves that means there is no apparent effect on the self-reported offending of the individuals. Moving from an individual's attitude via lower collective efficacy appears to actually reduce the propensity to offend and rather substantially at just over 13% (mediation effect 1 in Table 7.8), whilst the pathway through a neighbourhood with higher perceived disorder only seems to increase offending by around 9% (mediation effect 2). But, as was discussed in the previous section, the perceived lower collective efficacy has been shown to cause lower perceptions of disorder (as shown by the single-headed regression arrow in Figure 7.14) and this together means that collective efficacy has an insignificant mediating effect (so mediation effect 1 in Table 7.8 combines with mediation effect 3 to give effect 4). Conversely, the same inter-linkage, operating in conjunction with neighbourhood disorder (mediation 2 and 3, giving 5) means that the perception of neighbourhood disorder increases offending by around a fifth in comparison to just the direct effect of individual attitude on offending. These conflicting effects eventually almost cancel each other out, leaving only a small effect through either of the neighbourhood mechanisms (effect 6) and it is this overall result that could explain the elusiveness of this finding in other work.

At the family level the interrelations are simpler and hence easier to interpret. As previously seen in section 7.5.1, the relationship between the variables of perceptions of reinforcement and discipline is not significant once individual attitude is controlled for, showing that the bivariate interrelation is in fact explained by the individual in the setting. Reinforcement is far from significant and thus does not have an effect on offending (mediation 7). Parental discipline (in Figure 7.14) however is significantly regressed onto individual attitude, with a one standard deviation change on individual attitude associated with a 70% of a standard deviation increase in perceptions of negative discipline, discipline in turn is regressed significantly onto self-report offending but the total mediation effect (the eighth in the table) is found to fall just outside the level required for standard significance, at 94.2% (this type of relationship, two significant regressions and then an insignificant mediation, is not considered an anomaly by MacKinnon, 2008). The proximity of this to significance means that the result cannot be waved away without further investigation. It would appear from the negative sign on the mediating effect, as with the thought experiment method that was deemed to be slightly stretching the method, that lower family discipline actually reduces offending propensity slightly. This appears to be counterintuitive, as one would expect lower discipline or more lax rules to increase offending, though it could be explained by those with increased individual attitudes reacting badly to an overly prescriptive household environment, which is made worse by their very personality or that the discipline variable could be tapping a slightly different concept: one of permissiveness and a permissive or trusting environment is actually a protective effect. Or it could be a reflection and continuation of the poor environment that helped develop the poor individual attitude in the first place (as many DLC theories contend) or indeed it could, in the current work, have merely have occurred by chance. This shall be returned to as a central discussion in Chapter 10.

To summarise briefly and returning to the fictional respondent first introduced in the measurement chapter, an individual with a higher criminogenic propensity (synonymously a worse individual attitude) is likely to:

- report that they live in a household where they receive a lower level of parental reinforcement;
- and that they live in a home environment with lower levels of discipline (or a more permissive home environment);
- but these differential levels of perception of the family environment cause no significant difference on whether the respondent self-reports offending.

The more criminogenic individual is also likely to:

- perceive the area in which they live as having lower levels of collective efficacy (cohesions and trust amongst neighbours);
- and are likely to rate the area as having a higher level of disorder and anti-social behaviour, with this being increased by the perceived low collective efficacy;
- the perception of the cohesion and trust (collective efficacy) appears to have no significant effect on the rate of offending;
- but the perception of the disorder in the neighbourhood (or the level of neighbourhood incivilities) seems to significantly and substantially increase the rate at which the person would self-report offending.
- This effect is amplified by the perceptions of lower collective efficacy, meaning that the effect of worse perceptions of disorder is operating in part (around half) through the collective efficacy perceptions.

7.7 Structural Conclusions

This chapter very much straddles both the methodological and the substantive goals of this thesis. By observing that criminogenic propensity is seen as stable and develops before the age of 10 in all of the developmental and life course criminology theories reviewed, and that each of these theories acknowledges that the environment can also affect people's offending, this work was able to define what was referred to as a psychological map of interrelations (Figure 7.1) between the individual's criminogenic propensity, the family situation and area character. This is a predominantly theoretical step but leads to the methodological derivation of an empirical model that matches against this theoretical schematic. The strength of the model, once the measurement

system has been accepted as accurate, is that it relies on only the observation that individual attitude is stable (relative to others), that mediation is the only way to accurately capture the relationships amongst a set of latent variables that are all reported by one individual and that perceptions of neighbourhood disorder are affected by perceptions of collective efficacy. This was achieved by working through each of the domains and the relationships in a methodical manner and looking not just to theory or to statistical output, but in weaving both together.

From this methodological approach, some interesting and fresh findings have started to emerge. Firstly, it appears that perceived neighbourhood disorder does play a rather important role in determining the level of offending that an individual reports, with those living in areas that they perceive to have one standard deviation higher disorder offending at around the same rate as an individual with a fifth of a standard deviation higher score on their individual attitude (i.e. neighbourhood disorder mediates the relationship by just under 20%) and about half this operates through perceptions of worse collective efficacy making the perceptions of disorder worse. This tangle of area level effects does not seem to have been incorporated into analysis before, and by separating out the two mechanisms and correctly incorporating their interdependence this shines a light not only on how area affects self-report offending, but also why these results have gone un-discovered before; without the interdependency the two factors seem to combine to almost cancel each other out!

Of equal interest is the finding that the family level measures, particularly the measure of family discipline, doesn't seem to have any effect on self-report offending, with a mediation pathway that verges just outside significance. Although not *statistically* significant, the negative coefficient and the relatively large point estimate (a proportion of mediation of around 18%, similar to that of the area level variable but falling just the 'wrong side of the line') are *substantively* important and this means that this requires further attention. A hypothesis can be put forward regarding stability over time of the disciplinary regime with extrapolation back into earlier childhood where the criminogenic propensity is first developed, but other mechanisms can also be posited and with this limited evidence more needs to be done to explain this result. This will be

a key part of the analysis in Chapters 9 and 10. This analysis is far from complete and these findings will be revisited and refined with the introduction of other risk factors in Chapter 8, but it certainly provides a good basis for further exploration.

By adopting an exhaustive approach to this model derivation, it was necessary to remove fear of local crime from the analysis, as there was not the theoretical underpinning to include this in the current analysis. The emergent findings here did appear to be of interest and future work could try to address this by incorporating the theoretical argument. For the current work, however, this was deemed outside the scope of the thesis.

Chapter 8

Attempting to Model Different Crime Types

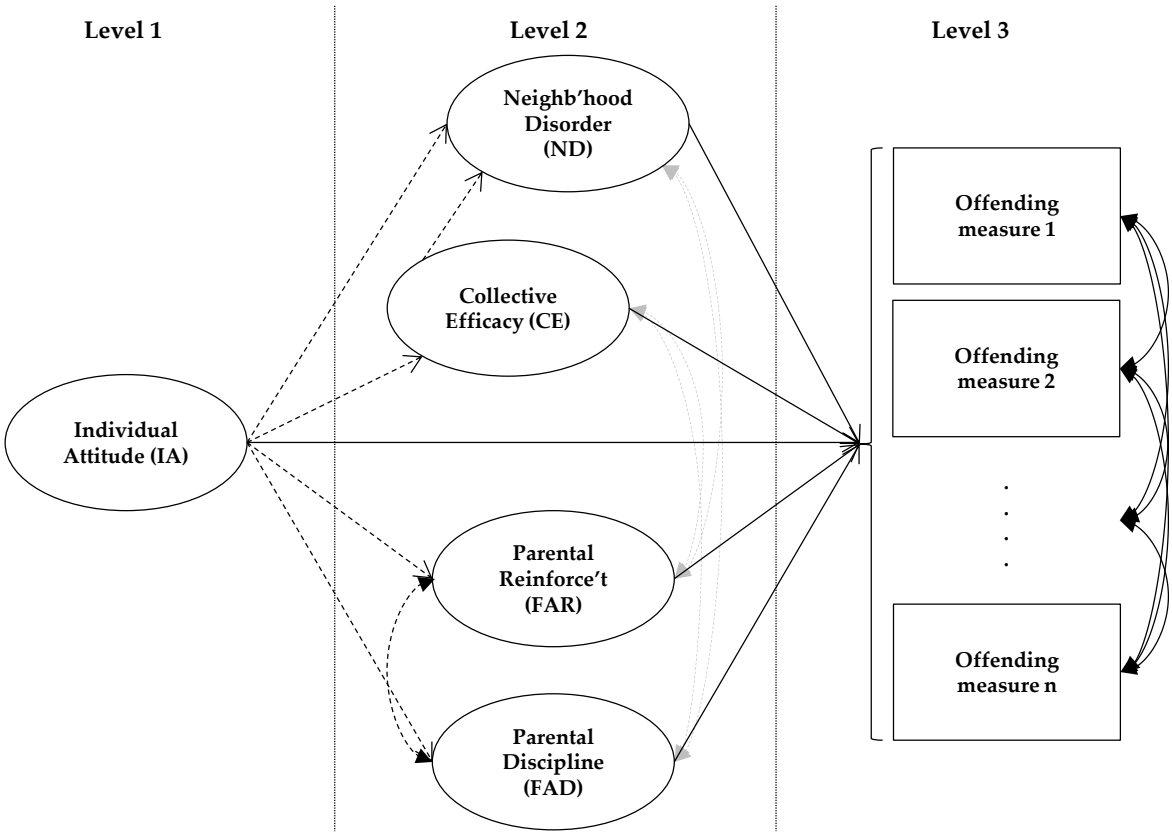
8.1 Aims: A Deepening of Understanding?

The focus of the previous chapter was to create a viable model and to define the interrelations between the concepts that were measured in Chapter 6. Now that this has been established, it is possible to change parts of it to get a better understanding of exactly how (self-report) offending is driven by the components present. As discussed in section 4.6, the Offending, Crime and Justice Survey (OCJS) uses a methodologically sophisticated and reliable technique to measure self-report offending. As such, the survey holds information on specific subdomains of offending, as well as a grouped frequency variable for each. Although the utility of some of these sub-divisions and counts is limited (see section 3.6 and later in this chapter) over 12% of respondents reported committing at least one offence for what are termed within the survey as violent offending, property offending and serious offending. Alongside this a 'frequent offender' dichotomous variable can be created by choosing a cut point in the grouped count offending item. Where this cut is put in the data needs discussion but if appropriately defined, this definition should be able to mirror, at least in some way, the highest rate offenders (similar to a prolific and priority offender (PPO) – see Hopkins and Wickson, 2012).

In order to aid discussion, Figure 7.13 from the previous chapter can be reformulated into a more general format, as shown in Figure 8.1. This representation also splits the domains within the model into three different 'levels': level 1 represent 'exogenous' variables, level 2 are endogenous mediators and level 3 are dependent variables. In the current chapter, the only level 1 variable is individual attitude, as discussed in sections 7.2 and 7.3. The dashed arrows on the left of the diagram are either regressions or covariances that are not affected by the dependent, and therefore will be the same for all models in this section of the work (barring estimation differences), whilst the solid

arrows are regressions onto the dependent and are the crucial part of this model for this stage of the investigation. The diagram is also intended to convey the fact that within a SEM framework the analysis is not limited to just one dependent and this is represented by the multiple different offending measures (up to n) present on the right hand side in level 3, all of which have possible covariances with one another. For the majority of this work, however, there is only one dependent and the models are run separately to give clearer results.

Figure 8.1 Empirical model specification when altering the dependent variable



The work need not only concentrate on dichotomous measures of offending either. As mentioned, there is a grouped version of the offending variable available and section 5.7 discussed how ordinal categories can be included in the modelling procedure with a probit link function. As the groupings are zero, one, two, three, four, five and six or more offences, the variable can also be considered as ratio level information and hence a linear link can be employed. With such high concentration of people reporting no offences (around 75% even for the most common all offending category) the linear

model is expected to be slightly inappropriate. Nonetheless, this lack of fit can be assessed and is worth carrying forward (particularly given the recent rise of linear probability modelling in economics – see Wooldridge, 2013). Other link functions could theoretically also be applied but these are not available in Mplus. This analysis can be conducted for all of the different types of crime available and of sufficient size in the dataset (i.e. alongside all offending, the aforementioned violent, property and serious).

As the arguments in Chapter 6 make clear, before beginning with any analysis it is of paramount importance to have a full understanding of exactly what it is that is being measured. In the previous chapters this warning has related mainly to the latent variables and the reliability of the dependent has only been approached from a survey methodological perspective (finding that the procedures in place were adequate to give a good measure). Although touched up on in section 4.7, there has been relatively little discussion about the offending variables and now deeper consideration and analysis is needed.

8.2 Offending Definitions and Statistical Power

In Chapter 4 it was seen that the OCJS used a methodologically robust approach to capturing self-report offending and a range of 20 offences were asked about by first questioning whether the respondent had ever committed the offence, whether in the last 12 months and then the last four weeks (although the four-week questions have significant missing data problems and appear as though they escaped the NatCen cleaning process). For each offence the (approximate) number of offences was also requested. As shown in Table 4.5 the vast majority of offences were committed and reported very rarely. For instance, only 16 respondents in 2004 reported stealing a vehicle and once the range of respondents is limited to those aged 16 years or under, only eight respondents (0.4% of the total) who were in each wave of the survey reported this offence type. This pattern carries on – of the 20 raw offence types, only assault without injury (14.3%), assault with injury (12.9%) and theft from school (9.5%) are reported in any significant numbers. Damaging a vehicle, theft from a shop or theft of anything else (i.e. not from school, work, another person or a shop) are reported as having been committed by less than 3% of respondents, with the remaining offence

types reported by less than 30 respondents. In short: the majority of offence types are rare events and certainly not enough to be adequately modelled. The impact of this is that the finest subtype analysis is not possible.

Chapter 4 briefly introduced the grouping of offences that are present in the OCJS dataset which offer a large enough sample to be modelled. These offence types of violent, property and serious have rather attractive names in terms of their policy significance: being able to make distinctions about the risk factors associated with offending in these different crime types would be an exciting and useful piece of knowledge. However, as Table 8.1 reiterates, these definitions need to be critically examined. For the grouping category of property offences, of the 314 respondents classified as having self-reported in this grouping, 64% of these people had answered that they had committed theft from school and 21% theft from a shop (with 42 respondents reporting having done both) with the other offence types in the grouping being reported less than half as frequently. Serious offences are similarly dominated by one survey category with the vast majority, over 90%, being assault with injury. Violent offences are almost entirely some form of assault, either with or without injury (with 146 respondents reporting both). These distinctions are important and need to be borne in mind during the analysis. Of particular note is that the difference between violent and serious offending is fundamentally the removal of injury from an assault. It is possible to imagine this distinction, especially given that the respondents are all aged 10 to 16; it becomes a difference between a school fight on the yard that is broken up and violent situation that gets rather out of hand resulting in a definite injury - a substantive difference. Nevertheless, these categories are worth sticking with, their policy relevance is clear to see and the differences in their definitions suitably clear as to be able to be understood later. They are the only sensible divisions that can be applied, due to low responses (an argument could be made for dividing violence into with/without injury, but it is felt that it is the presence of violence that is important, not the severity, and this will not be pursued).

Table 8.1 Grouped offences by base survey offence types, 2004 sweep of longitudinal dataset, 16 and under

Grouping	Offence survey item	Respondents reporting	Percentage of total in 2004 sample
Violent offences	Assaulted without injury	305	13.9%
Reported by:	Assaulted with injury	275	12.6%
434 (19.8%)	Commercial robbery	5	0.2%
	Personal robbery	3	0.1%
Property offence	Theft from school	200	9.1%
Reported by:	Theft from shop	67	3.1%
314 (14.3%)	Stolen parts off outside vehicle	30	1.4%
	Commercial burglary	28	1.3%
	Theft from person	27	1.2%
	Theft from work	20	0.9%
	Stolen a vehicle	8	0.4%
	Stolen parts from inside vehicle	8	0.4%
	Tried to steal a vehicle	7	0.3%
	Domestic burglary	4	0.2%
Serious offence	Assaulted with injury	275	12.6%
Reported by:	Commercial burglary	28	1.3%
297 (13.6%)	Theft from person	27	1.2%
	Stolen a vehicle	8	0.4%
	Commercial robbery	5	0.2%
	Sold Class A drugs	5	0.2%
	Domestic burglary	4	0.2%
	Personal robbery	3	0.1%

Alongside the three offence types groupings in Table 8.1 it is interesting to define and look at the most frequent offenders. As touched upon in the introduction to this chapter these have a clear theoretical importance but the question is how to define this group, where to put the cut in the count variable. The obvious choice is the very top, capturing those who reported the highest rate of offending. This would be six or more offences

and would result in 135, or 7%, of the respondents being included in the definition. This certainly represents the highest available here and the percentage of the population approximately matches the predicted rates for the high rate offenders in other works (such as PPOs and matches Moffitt's estimated percentage of life-course persisters). Moving to a lower threshold, say four or more offences, would still result in capturing people who are reporting offending more than "a couple of times" and would define around 10% of the sample as high frequency. Capaldi and Patterson (1996) define a frequent offender as someone who has three or more criminal sanctions, but this definition operates under the logic that to receive three full disposals in one year represents a lot more un-convicted offending behaviour[26]. The OCJS itself actually defines a frequent offender as someone with six or more offences in any one year, and this matches the Narrowing the Justice Gap[27] initiative definition (although this too was phrased in terms of convicted offences). There seems, therefore, to be little choice and despite the attraction of boosting the size of the group to increase precision of the estimate, using a basis of six or more offences is the more sensible definition of a frequent offender.

Table 8.2 shows some descriptive statistics on the rates of self-report offending for the 2004 wave of the main dataset used in this work and to those knowledgeable of Youth Justice Statistics may seem slight surprising. Specifically, the 'violent' offence rate is five percentage points higher than the property rate, whereas in Youth Justice Statistics these are seen to be of comparable levels. Once the nature of the violent category is taken into consideration, i.e. more or less equally split between with and without injury, then it becomes clear that this difference is entirely to be expected. Violence without injury, especially amongst 10- to 16-year-olds, would very rarely be reported to the police and a more apt comparator to the property offending is in reality the serious grouping (predominantly violence *with* injury) and the levels are roughly equal. These

---

[26] as the gap between total Crime Survey in England and Wales offences (7.3m) total recorded crime (3.7m) and the number of offenders convicted (1.2m) adequately demonstrates (Home Office, 2014 and Ministry of Justice, 2014)

[27] Available online at https://www.cps.gov.uk/publications/prosecution/justicegap.html [July, 2015]

reported levels are broadly in line with the levels reported in the Wikström et al. (2012) Peterborough setting.

Table 8.2 Number of self-report offences in last year by offence type, 2004 sweep of longitudinal dataset, 16 and under

	All		Violent		Property		Serious	
	Freq.	% valid	Freq.	% valid	Freq.	% valid	Freq.	% valid
Binary yes/no								
Not in last 12 months	1,456	71.5%	1,673	79.4%	1,746	84.8%	1,825	86.0%
Offended in last year	581	28.5%	434	20.6%	314	15.2%	297	14.0%
Total valid (% of grand total)	**2,037**	**93.0%**	**2,107**	**96.2%**	**2,060**	**94.1%**	**2,122**	**96.9%**
Missing	153	7.0%	83	3.8%	130	5.9%	68	3.1%
Count								
Never in last year	1,456	74.6%	1,673	85.7%	1,746	89.4%	1,825	93.5%
Once in last year	151	7.7%	163	8.4%	95	4.9%	163	8.4%
Twice	89	4.6%	80	4.1%	63	3.2%	39	2.0%
Three times	54	2.8%	42	2.2%	21	1.1%	34	1.7%
Four times	27	1.4%	24	1.2%	18	0.9%	13	0.7%
Five times	40	2.0%	24	1.2%	14	0.7%	6	0.3%
Six or more times	135	6.9%	84	4.3%	66	3.4%	34	1.7%
Total valid (% of grand total)	**1,952**	**89.1%**	**2,090**	**95.4%**	**2,023**	**92.4%**	**2,114**	**96.5%**
Missing	238	10.9%	100	4.6%	167	7.6%	76	3.5%
Grand total	**2,190**		**2,190**		**2,190**		**2,190**	

The problem of missing data needs to be cleared up quickly. It may seem peculiar that the all offending category has a greater number of missing values than the three sub-offence types, and that the binary categorisation has a lower number of missing values than the count. This is due to the survey design. The respondent would be asked whether they committed a certain offence and then the CAPI software would go on to ask for details, one of which would be how frequently the offence was committed in the

last year. This latter question could be missed, resulting in a valid yes/no, but a missing count for a specific offending definition. The all offending category would be set to missing if the respondent had missed any of the sub-category questions.

In summary, there are some issues that need to be kept in mind when reaching a conclusion based on these categorisations of offending. Firstly, the majority offence types in each division need to be understood and results need to be phrased in terms of these specific compositions. Second, the limited sample size of some of the groupings (especially frequent offenders) may mean that the precision of any models may be low and there may be a higher probability of type 2 error. A usual way to combat this is to increase sample size (Tabachnick and Fidell, 2005), although this shouldn't be thought of as a silver bullet, as the difference between statistically significant and statistically insignificant results can often be substantively insignificant. With this in mind, an aspect of the OCJS that has thus far been underutilised in this research can be called upon to strengthen results where needed: namely either including the 567 respondents in the 2004 wave who were excluded due to later year attrition, or in extreme cases all of the years' data may be pooled. These steps will be discussed where appropriate and a conclusion will be made based on an overall impression of fit across a number of models rather than solely on one set of results.

8.3 Modelling Different Versions of All Offending

This first step will alter the dependent from a binary yes/no in the previous 12 months to a count variable and treat it alternatively as ordinal and continuous. Although the primary distinction in self-report offending surveys is between those who admit having offended and those who do not, the level at which people report offending can also be of interest (Thornberry and Krohn, 2000) and this specification should capture this. Using the finer grained count as either a (truncated at six or more offences) continuous variable or as an ordinal dependent will therefore increase reliability and any consistencies or difference between the models will strengthen any results. This may seem like something of a trivial step and it could be argued that these results will supersede those with a binary dependent variable, and to some extent this is true; it is undeniable that the previous chapter imposed a rather arbitrary cut point at reported

versus did not report when a greater granulation is offered, whilst the following modelling framework captures that knowledge and more. The focus of the previous chapter, however, was on the formulation of the model in the simplest sense, and now that this has been established, results from further analysis can be taken in conjunction with those previously presented. Beyond that, the primary distinction to make is between those who report and those who do not, and the reliability of the self-report is called into increasing question the higher the level of granularity.

Predicted probabilities as calculated via the 'thought experiment' (as examined in section 7.4.3) are not as interesting here as the proportions of mediation and their significance. It has been established that both are indicators of the impact of the higher level (level 2) latent components and the proportions of mediation are a more succinct way of expressing this. In the case where the dependent variable is treated as continuous, the caution over a possible approximation in estimating this effect that was described in section 5.7 is no longer the case, and the proportion of mediation is exact as the link function is linear.

The results, along with a restatement of those from the previous chapter (the binary results), can be found in Table 8.3. This table presents only the most relevant parts of Table 7.8, the total effects of neighbourhood disorder or collective efficacy once the other neighbourhood characteristic has been included as in indirect pathway, along with the family reinforcement and discipline variables. All of the regressions/covariances represented by dotted lines in Figure 8.1 are the same in each model (to within estimation error, i.e., three decimal places).

Table 8.3 Proportion of mediation for various measures of all offending

	Binary		Ordinal		Continuous	
Mediation via…	**Prop'tion mediation**	**Sig level**	**Prop'tion mediation**	**Sig level**	**Prop'tion mediation**	**Sig level**
Area						
Neighbourhood disorder (total)	19.0%	***	18.2%	***	15.0%	***
Collective efficacy (total)	-3.3%	89.7%	-2.9%	87.9%	-2.5%	78.5%
Family						
Familial reinforcement	-6.6%	65.7%	-8.5%	78.2%	-9.2%	78.2%
Familial discipline	-17.7%	94.2%	-19.5%	*	-22.8%	*

Before trying to interpret the results, it is worth asking which model is most appropriate. In Chapter 6, and to a lesser extent Chapter 7, the TLI, CFI and RMSEA statistics were used to assess fit. In the current case as only a small component of the model is changing, just the single dependent variable, these fit statistics are to all intents and purposes the same across the specifications. Attention needs to be turned to how well the dependent is modelled and this can be done with reference to the r-squared. For the binary and the ordinal probit regressions the r-squared is 0.45 and 0.46 respectively whereas for the linear regression where the dependent is taken as continuous, the r-squared falls to 0.23. This reiterates the expectation that the continuous version is not appropriate and less weight should be paid to this result.

Despite the relatively bad fit of the continuous version the results in Table 8.1 demonstrate the need to conduct analysis such as this. Whereas in the previous model (the binary case) family discipline had been found not to have a significant effect on offending, in the ordinal and the continuous variant the effect falls the other side of 95% level. Neighbourhood disorder retains its significance at the 99.9% level but the size of the effect is reduced, from 19% to 15% in the continuous version. Set against the increase in the size of the family discipline effect point estimate (i.e. of comparable substantive level, if not of comparable significance) and it can be seen that whereas the results from the previous chapter implied a rejection of the importance of family, this component is firmly back in the picture. The implication and meaning of this will be discussed at the end of the chapter when set within a context of the other crime type dependents.

8.4 Modelling Alternative Offending Types

The following four subsections will present a set of results for violent, property, serious and frequent offending with little commentary as to the implications, before pulling this all together and drawing conclusions capturing all of the results in one narrative.

8.4.1 Violent Offending

When setting the dependent to the suite of measures (binary, ordinal and continuous) of violent offending, the r-squared is lower than for all offending. For the binary and ordinal probit link functions the values are 0.36 and 0.35 respectively, whilst the continuous indicator falls to 0.14. The lack of fit of the count is to be expected for exactly the same reasons as discussed for the all offending and warning as to the validity of these results holds. The regression coefficients represented by the solid lines in the diagram of Figure 8.1 have broadly the same rank order as those given when modelling the all offending dependent, although they are slightly lower and in keeping with the previous section are not presented as little can be seen from this observation that is not also reflected in the mediation effects. Table 8.4 thus presents for violent offending the analogous data as Table 8.3.

The neighbourhood disorder total mediation effect is the only component that is consistently significant and although the effect is roughly the same order of magnitude as with the all offending, as the predictive power of the model on the dependent is lower, the significance level is reduced. Family discipline is significant only for the continuous case, but due to the lack of fit less attention should be paid to this finding and on balance the role of discipline seems to be lower than for all offending. As in all previous models, collective efficacy and familial reinforcement are still insignificant across all specifications.

Table 8.4 Proportion of mediation for various measures of violent offending

	Binary		Ordinal		Continuous	
Mediation via…	**Prop'tion mediation**	**Sig level**	**Prop'tion mediation**	**Sig level**	**Prop'tion mediation**	**Sig level**
Area						
Neighbourhood disorder (total)	19.1%	***	16.7%	***	12.4%	*
Collective efficacy (total)	-3.7%	88.6%	-2.7%	79.1%	-1.1%	40.0%
Family						
Familial reinforcement	-9.5%	78.2%	-10.1%	79.9%	-11.5%	83.0%
Familial discipline	-15.8%	87.2%	-17.2%	91.5%	-21.7%	*

8.4.2 Property Offending

Table 8.5 presents the same results as Table 8.3, this time for property offending. The r-squared here is marginally higher than the all offending at 0.47 and 0.44 for binary and ordinal respectively, but falls 0.16 for the continuous measure, a large reduction and warrants an extra warning on the reliability of these findings. The results show the neighbourhood disorder total effect is again the most consistent mediator, but the relative effect size is reduced from around 19% to 14%. Family discipline again borders significance, but the ordinal version where the r-squared is highest does not show a significant result, indicating the effect of discipline here could be diminished from the all offending model.

Table 8.5 Proportion of mediation for various measures of property offending

	Binary		Ordinal		Continuous	
Mediation via…	**Prop'tion mediation**	**Sig level**	**Prop'tion mediation**	**Sig level**	**Prop'tion mediation**	**Sig level**
Area						
Neighbourhood disorder (total)	14.6%	***	14.3%	**	9.6%	*
Collective efficacy (total)	-2.6%	81.7%	-2.9%	81.8%	-2.8%	78.2%
Family						
Familial reinforcement	2.0%	24.2%	-0.5%	5.6%	-3.9%	44.5%
Familial discipline	-18.8%	*	-16.2%	90.2%	-20.7%	*

8.4.3 Serious Offending

Serious offending has the lowest reporting rate of the main offence types (excluding the frequent definition) and hence it is unsurprising that the r-squared is lowest, at 0.33 for both binary and ordinal and just 0.11 for the continuous version. Neighbourhood disorder's total effect still mediates the direct effect of individual attitude, but its magnitude and significance level has been reduced. Family discipline is no longer significant in any specification. Table 8.6 refers.

Table 8.6 Proportion of mediation for various measures of serious offending

	Binary		Ordinal		Continuous	
Mediation via…	**Prop'tion mediation**	**Sig level**	**Prop'tion mediation**	**Sig level**	**Prop'tion mediation**	**Sig level**
Area						
Neighbourhood disorder (total)	14.0%	**	9.7%	*	6.7%	75.1%
Collective efficacy (total)	0.2%	8.5%	-1.2%	43.8%	-2.5%	64.6%
Family						
Familial reinforcement	6.0%	54.6%	4.8%	45.3%	0.4%	4.2%
Familial discipline	-16.1%	85.0%	-11.3%	70.6%	-10.3%	59.2%

8.4.4 Frequent Offending

Incorporating high frequency offenders into the model (Table 8.7) sits apart from the analysis that has gone previously as the data only allows this to be a yes/no distinction. The significance of neighbourhood disorder is further lowered, the point estimate at 11% is lower than any of the previous models and it is the only component that can be seen as a mediator. This leads to a tentative statement that the mediating effects are lower (almost not present) for this type of high-level offender.

Table 8.7 Proportion of mediation for various high frequency offenders

Mediation via…	Binary	
	Prop'tion mediation	Sig level
Area		
Neighbourhood disorder (total)	11.1%	*
Collective efficacy (total)	-4.1%	87.9%
Family		
Familial reinforcement	-2.5%	23.4%
Familial discipline	-16.0%	84.3%

8.5 Summary So Far

This chapter so far has presented something of a blizzard of statistical output with very little commentary on what these results mean, hence it is worth taking a moment to take stock of the results and to attempt to pull them together before moving on to testing these conclusions with further analysis. With so many tests the chances of a type 1 (rejection of a true null hypothesis) or type 2 (non-rejection of a false null hypothesis) error increases, and the worry of the data miner becomes non-trivial[28]; therefore, some level of scepticism about the importance of the significance stars in the tables above is needed.

What appears to be coming though, however, is that the effects of the mediating variables seem to vary with the crime type being used as a dependent. For all offending the range of results over the models seems to be consistent with those from the previous chapter; neighbourhood disorder mediates offending, with those living in areas that they perceive to have higher disorder offending more than an individual who is placed similarly on the criminogenic propensity scale. There is some ambiguity about the role of family discipline which is insignificant for the binary but 'over the line' with the ordinal model and the (less reliable) continuous model. For violent and serious offending, the effect of family discipline seems to disappear and for the latter of the

[28] For a rather humorous representation of the dangers of focusing solely on an occasional significant result see http://xkcd.com/882/

two, serious offending, the importance of neighbourhood disorder is reduced with a lower mediation effect point estimate (from 19% to around 14%) and a lower significance level. For property offending, the mediation effect size of neighbourhood disorder is lower though still significant at most conventional levels but with the family discipline effect size increasing and again bordering significance. Finally, for frequent offending, there is evidence that there is a much reduced set of mediations with only neighbourhood disorder presenting a significant result and at just the 95% level and with a reduced point estimate of 11%, down from the all offending level of 19%.

But are these results to be believed? The results are certainly attractive from a theoretical perspective, with all offending and violent offending (and around seven out of ten who report any offending report violent offending) being mediated by structural perceptions, as would be expected given their spontaneous nature. Property offending (requiring more forethought) and serious offending (including more extreme violence) are less mediated and hence the individual propensity plays a more important role as the deciding factor. For frequent offending this mediation is analytically even less apparent, indicating that the respondent is behaving a little more like the life-course persister model of Moffitt's theory, the 'worst', least controllable individuals. A more sceptical view, however, could claim this is merely a symptom of the data that are being recorded, and specifically the sample sizes available. Section 8.2 of this chapter has already shown that the actual differences between these offence types (especially for violence and serious) amount to a much smaller distinction than the names suggest: namely whether any form of injury was sustained. And both of the frequent and, to a lesser extent, the serious definitions suffer from having relatively low numbers of people reporting having committed these offences. Further analysis can therefore be carried out, making use of a slightly different modelling framework and drawing on the extra years of data and those lost through attrition, to stretch and test these tentative conclusions.

8.6 Testing Conclusions with Further Analysis

8.6.1 Focusing on the Dependents

Structural equation modelling is not bound like normal regression techniques to have just one dependent variable and its versatility means that each of the models that were introduced in section 8.4 (with the exception of all offending, as it is a super-group of the others) can be combined into one modelling framework. This does not affect the point estimates or the standard errors that were estimated for each of the models presented separately in the previous section but it does allow the conditional correlations between the dependents to be assessed and these results are presented in Table 8.8 for both the dichotomous and ordinal version of the model.

Table 8.8 Conditional correlations for the 2004 offending dependents

	Binary				Ordinal		
	Violent	**Property**	**Serious**	**Frequent**	**Violent**	**Property**	**Serious**
Violent	1				1		
Property	0.396	1			0.387	1	
Serious	0.962	0.479	1		0.890	0.489	1
Frequent	0.719	0.65	0.668	1	n/a	n/a	n/a

The results presented are not overly surprising given the discussed overlap between the measures during the descriptive work, but what is important is the impact that these covariances will have on the model, and the manner in which this will affect interpretation. The most problematic overlap is between serious and violent offending and the value of 0.96 (binary) or 0.89 (ordinal) means that these measures are basically the same and the difference in the model results is most likely due to limited statistical power rather than something substantive. The same is true, though at decreasingly lower levels for frequent offending with violent, frequent with serious and frequent with property offending. The only distinction that can be easily made is between property and violent and property and serious offending. This lack of analytical separation when the dependents are modelled together is troublesome and were these being entered into a regression model as independents there would be a strong case to remove one from the analysis due to collinearity. Analogously, is this attempt to

distinguish these offence types bound to failure and purely a statistical power phenomena?

8.6.2 Focusing on the Mediation and Sample Size

The final piece of analysis for the current chapter is to run the models on different datasets, making use of the extra information that for most of this thesis is not needed. Table 8.9 holds the point estimates and significance levels when the analysis was conducted on three different datasets for each of the offending types using the binary measure of offending. The datasets are the 2004 to 2006 longitudinal version that has been used for the majority of the analysis within the work thus far, the 2004 dataset without excluding the 567 respondents who were not subsequently included in later waves and finally on a 7,666 respondent pooled set of all years' data where person level fixed effects were added to the model to account for possible intra-person response bias.

The results paint an interesting picture demonstrating the complexity and lack of clarity that the models are actually allowing. Firstly, across all the datasets and within each offending type there is relative stability for the mediating effect of the role of neighbourhood disorder and as sample size increases the standard errors associated with each pathway decreases and consequently the significance level increases. The point estimate of the disorder mediation effect decreases from all and violent offending (around 19%) to property and serious (between 14% and 17%) to frequent (between 10 and 13%), implying that the influence of the mediation effect through disorder (and including the effect through collective efficacy) reduces from all and violent through property and more serious offending to the most troublesome respondents, the frequent offender, where the mediation is lowest.

The conclusion for total mediating effect of collective efficacy and of family reinforcement must be that these are *not* significant mediators for any offending type. Although collective efficacy appears with a single star, representing 95% significance, for property offending when the analysis is run against the largest dataset, and intuitively this seems to make sense, little attention should be paid to this result. Not

least because it appears just once and the mediation effect is substantively small (2.8%) but also because the intuition that strong local trust and cohesion should protect against property offences melts away once it is recalled that the majority of these offences are theft from school or from shops, which works such as Wikström's et al. (2012) show happens predominately in urban centres and not in the local area (and hence occur away from the area that the respondent would be referring to when assessing collective efficacy).

Family discipline is once again a troublesome variable and does not follow the patterns demonstrated by the other components, and conclusions about its behaviour are hard to make. The point estimate of the mediation effect alters significantly across the datasets. For the any offending dependent the estimate goes from just under 18% to almost 30% as the sample size increases, and a similar pattern is seen for violent and property offending, and not too dissimilar for frequent offending. The only definition of offending for which the point estimate stays relatively stable is serious offending, and here it is significant only when using the all waves dataset.

Table 8.9 Proportion of mediation for binary measures of different offending types on alternative datasets

	In longitudinal dataset		All 2004 wave		All waves combined (fixed effects)	
	Prop'tion mediation	**Sig level**	**Prop'tion mediation**	**Sig level**	**Prop'tion mediation**	**Sig level**
All Offending						
Neighbourhood disorder (total)	19.0%	***	20.3%	***	19.5%	***
Collective efficacy (total)	-3.3%	0.103	-2.1%	0.198	-1.4%	0.196
Familial reinforcement	-6.6%	0.343	-5.6%	0.339	1.9%	0.6
Familial discipline	-17.7%	0.058	-25.2%	**	-29.5%	***
Violent Offending						
Neighbourhood disorder (total)	19.1%	***	21.1%	***	20.0%	***
Collective efficacy (total)	-3.7%	0.114	-2.3%	0.256	-0.4%	0.755
Familial reinforcement	-9.5%	0.218	-8.6%	0.214	0.7%	0.908

	In longitudinal dataset		All 2004 wave		All waves combined (fixed effects)	
	Prop'tion mediation	**Sig level**	**Prop'tion mediation**	**Sig level**	**Prop'tion mediation**	**Sig level**
Familial discipline	-15.8%	0.128	-21.9%	*	-30.3%	***
Property Offending						
Neighbourhood disorder (total)	14.6%	***	17.2%	***	16.0%	***
Collective efficacy (total)	-2.6%	0.183	-1.9%	0.293	-2.8%	*
Familial reinforcement	2.0%	0.758	4.3%	0.444	6.5%	0.066
Familial discipline	-18.8%	*	-28.8%	**	-27.4%	***
Serious Offending						
Neighbourhood disorder (total)	14.0%	**	16.3%	***	17.6%	***
Collective efficacy (total)	0.2%	0.915	0.9%	0.625	1.5%	0.224
Familial reinforcement	6.0%	0.454	4.3%	0.52	7.9%	0.059
Familial discipline	-16.1%	0.15	-18.3%	0.07	-17.9%	*
Frequent Offending						
Neighbourhood disorder (total)	11.1%	*	12.6%	**	9.7%	***
Collective efficacy (total)	-4.1%	0.121	-4.4%	0.07	-2.0%	0.142
Familial reinforcement	-2.5%	0.766	-1.6%	0.818	1.2%	0.815
Familial discipline	-16.0%	0.157	-26.7%	*	-25.2%	**

8.7 Implications and Discussion

After considering what is held by the OCJS in terms of different measures of offending, this chapter has presented a range of slightly different specifications of the model initially defined in the previous chapter. This has included using four different sub-definitions of offending (violent, property, serious and frequent) alongside the original all offending dependent. Along with these different definitions, there are different methods of measurement which have utilized the offending count variable present in the dataset for each offence type. To test further the results, each of the four sub-offence

(three when using an ordinal version) types were modelled as dependents in a combined model which allowed the relationship between them to be reiterated. Finally, the datasets upon which the models were run was altered to inspect whether the lack of significant results was a symptom of the low number reporting offences or due to an actual lack of effect.

Initially the results from the analysis pointed towards some substantive differences between the mediating effects for different types of offending. It appeared that neighbourhood disorder was a consistently strong mediator for all and violent offending, with a slightly reduced role for property, a little less still for serious and an even lower level for frequent offending. Neither collective efficacy nor familial reinforcement mediate in any model specification, for any offending type. Familial discipline presents a mixed picture, often bordering a significant effect, and with a point estimate of approximately the same magnitude as the neighbourhood disorder total effect across most offence types and most specifications, meaning that its interpretation is less clear. The negative sign of this mediation is the same problem as highlighted in Chapter 7, and a further discussion of this is again reserved for Chapter 10.

In reality, however, there are some issues with the measurement. The introductory analysis of the offence definitions showed that violent and serious offending are extremely similar, despite their nomenclature, with serious offending a subset of the violent definition. Once this relationship is revisited when the definitions are combined into one model the closeness between serious and violent is again seen to be problematic. This criticism also becomes apparent, though to a lesser extent, for frequent offending and each of the other offending types present. The only seemingly distinct definition is property offending, and this, from the introductory descriptive statistics, is mainly theft from school. This means that it is, at least in part, the statistical power of the tests that are driving the differences in the results.

The number of people reporting serious offending is also rather low, as is the case for the frequent definition and for this reason the models were run on a number of larger datasets (with checks made for the other offending types as well). However, this did

not provide any robust results beyond confirming the importance of neighbourhood disorder. Specifically, the variation in the point estimate of the mediation effect for discipline gave less confidence in the results, but confirmed the irksome nature of trying to pin down this mediation effect.

Overall, this chapter produced some interesting and theoretically attractive results, but conclusions have to be tentative. Given the low response rates to some of the offence types, and the low separation between the different definitions, this analysis will not be carried forward into the controls and longitudinal sections, which will be left to concentrate on the family discipline factor.

Chapter 9

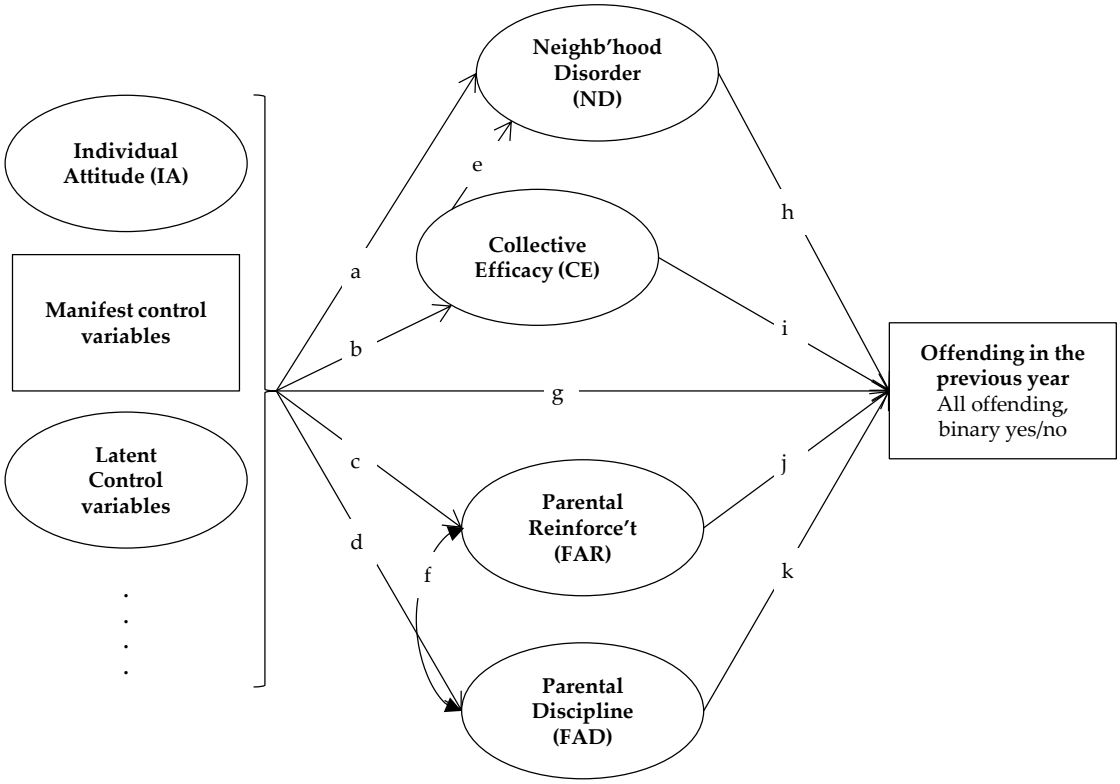
Expanding the Model to Account for Structural Characteristics

9.1 Overview of the Process

This analytical chapter will add exogenous controls to the model. This has not been attempted before, as previous chapters were concerned with creating the measures (Chapter 6), defining the model and the interrelationships between the measured concepts (Chapter 7) and testing the limits of what can be used as a dependent variable (Chapter 8). It is now time to rectify this missing piece of the puzzle.

Section 6.2 demonstrated that an individual's criminogenic propensity can be accurately measured using survey items from the OCJS and section 7.2 argued that this criminogenic propensity develops before the age of 10, is seen to be stable relative to others thereafter and therefore used this component as an exogenous latent factor within the modelling framework to control for the mediating effects of different known risk factors as reported by the individual. These were referred to as attitudinal or psychological mediators. The other set of possible risk factors that are available in the OCJS were called factual or external measures and were set aside to be dealt with later. It is these that will now be worked into the analysis. As with previous chapters this specification can be represented as a path diagram, and this can be seen in Figure 9.1. The left hand side of the diagram represents that a range of controls can be added to the model, both as manifest survey items and as latent variables.

Figure 9.1 Model specification with control variables



In keeping with the terminology used to distinguish the mediating risk factors these control variables can also be split into three levels: the individual, the family and higher level factors, and after some discussion, a list of those that are available and suitable can be found in Table 9.2.The key distinguishing feature of an exogenous variable to be added as a control is that it affects the model without being affected by it. Socio-demographics such as age, gender and ethnicity are examples at an individual level, whereas parental divorce, whether the parents have been in trouble with the police and household income would be at the family level. At the higher level there are items such as the neighbourhood index of multiple deprivation (or the sub-domains), region or length of time living in the area. The controls can be continuous, dichotomous, ordinal or latent and are interpreted in the same way as in a normal regression: if they are binary or categorical they would be in terms of the comparisons to a reference category; if continuous, in terms of a one unit change. This one unit change can be standardised in the case where the original scale of the variable is of no meaning (such as the neighbourhood condition) or in terms of a real value (such as for changes in age). As throughout this work the modelling will be done in a stepwise manner, as opposed to

identification of a range of items and then simply adding them all at once. Starting with individual level socio-demographics, it will consider what is possible to measure and add, before incorporating them and then moving to the higher levels.

Referring to Figure 9.1, in the previous chapter pathways a through to g were of little importance as they were invariant under the specification where only the dependent variable was changing. In the current chapter this is rather different and for each of these pathways there will be a different coefficient for each control variable. These will be of substantive importance as they will indicate how the perceptions of the mediators vary due to each control. Similarly, the coefficients h through k will alter depending on the controls added, and hence the indirect effects will also vary. Therefore all pathways will be considered, with proportions of mediation presented for each control. Predicted probabilities could be computed but their presentation would be cumbersome, especially for continuous or controls with many categories, and add little over the approximations that can be assessed via inspection of raw coefficients or via the mediation metric.

9.2 Control Variables for Inclusion

9.2.1 Individual Level Controls

There are two stand-out controls that absolutely need to be added to the model: age and gender. The first of these, age, is one of the defining characteristics of offending, with the age crime curve perhaps the most well-established and consistent findings in criminology (Loeber and Farrington, 2014), though it can be misunderstood. Although the percentage of offenders in the population at a certain age forms an asymmetric bell shape, increasing from late childhood to peak in the teenage years and declining from around the age of 20 onwards, this represents a combination both of individual incidences of offending and the population level prevalence of offenders, with there being some debate around which takes precedence in defining that shape. Different types of offending have been seen to peak at different ages despite there being little specialisation in offending (McVie, 2005; Farrington, 1986; Sampson and Laub, 2003). Although there is some contention as to the nuances of the role of age on offending,

what is undeniable is that it plays a role and therefore must be included. Age, somewhat obviously, is present in the dataset and has no missing values.

The second, gender, is also clearly seen to be associated with changes in the offending rate, though not necessarily as clearly as would be expected looking only at official arrest data. Whereas Youth Justice Statistics reports that male young arrestees are over four times more numerous than female, and the male prison population dwarfs that of females, self-report rates are much closer. Canter (1982) used the 1977 wave of the US National Survey and found significant differences in the proportion and mean number of behaviours between the two genders for the majority of offence types that were available, though the differences were less than the one to four ratio implied by arrest statistics. Gender is present in the dataset and has no missing values.

Beyond these two covariates, there is little else that can be added as a control that would be classified as individual. Ethnicity has played a central part of criminological theory as far back as the early Chicago School (Shaw and Mackay, 1942), with evidence suggesting that observed differences between ethnic groups disappear once items such as socioeconomic status, education levels of both the individual and the family, living standards and early childhood environmental factors are considered (Fergusson et al., 1993), and it has been argued that it is the criminal justice system's differential reactions to those of different ethnic origin that is the real crux of the issue (Webster, 2015). Unfortunately, the OCJS suffers along with many studies that attempt to investigate ethnic differences in offending through a self-report methodology (Hawkins et al., 2000) and that being of sample size. Although there are headline statistics produced from the OCJS and published by the Ministry of Justice (Sharp and Budd, 2005) in the reduced longitudinal dataset there are only 182 (8.2%) non-white respondents (in the entire 2004 dataset just 250) under 17 years of age. This low number is split across a set of disparate ethnic groups (with extremely low sample size even at the lowest granulation of Asian or Asian British, Black and Black British, mixed and other) and to combine into one super Black Minority Ethnic (BME) group would do little to get over this sample size issue and forcing this wide range of ethnicities into one grouping does not seem

appropriate. Unfortunately, therefore, ethnicity cannot be included, the implication being that this work essentially ignores any ethnic differences.

9.2.2 Family Level Controls

In addition to the two family level mediating variables that are present in the model already a number of covariates can be added to account for a number of the known risk factors as outlined in Chapter 2 and revisited in Chapter 6. Of particular interest are parental divorce, parental criminal involvement, family socioeconomic class and household income, and a measure of each of these can be derived from the data (here family and household are used synonymously).

Parental divorce and parental criminal involvement relies on the recall of the respondent, however it is fair to assume that the reporting of these, given accurate survey administration, will be less prone to misreporting than the endogenous variables. That is not to say that there is no reporting error in these items, especially the police contact question; whether a parent gets in trouble with the police could very well be invisible to a child or indeed the child could protect the parents by deliberate misreporting, but if it is reported one must assume that this reflects an increase in criminal involvement of the parent under the assumption that if the child reports it, it is likely to have happened. These measures are not considered as mediating effects as they are (or will be assumed to be) unaffected by the respondents themselves, i.e. the parents' criminal involvement or any incidences of divorce/separation is not caused by the children and their reporting is not caused by a psychological state. Problems with children are a rare cause of separation (Cleek and Pearson, 1985; Amato and Previti, 2003) and it is assumed to be similar and clearer for criminal involvement.

Parental criminal involvement (not including driving fines) is recorded into three categories with 90% reporting no trouble, 8% trouble but no report of prison and 2% that their parents had been arrested and subsequently jailed. There are 188 missing values caused by a mix of missing data (non-response, refusal and error) at different stages of the routing and is poorly documented and labelled in the documentation. It is therefore too mixed to include in the analysis as a 'non-response' category.

Parental separation is not specifically asked of all respondents but can be measured by questions asking the respondent whether they still have contact with their parents, and if not how often they see the parent in question. In the longitudinal dataset, over 95% still live with their mother whereas 33% (599) do not live with their father. Of those that don't live with their father, 28% (172 – 8% of the total) reported that they do not see him with a further 26 reporting their father was deceased. This can be used as a proxy for separation, and grouped into three categories (together, separate and still see, separate but don't see). There are 106 missing values and for the same reason as the parental criminal involvement (an unknown mix of reasons for the non-response) cannot be included in the analysis.

Two more closely related family level variables can be included: namely household socioeconomic classification and household income. A cross-tabulation of the two is presented in Table 9.1 and it contains some alarming figures. Although the split of the population (the column titled 'prop of sample') into the socioeconomic classification is broadly similar to that found in other sources (see, for instance, Hills et al., 2015 supplemental tables[29]) the annual income of these groups is, to put it bluntly, unbelievable. That 14% of employers or high professionals earn less than £10,000 per year is so completely outside the range of possibility that the whole income variable must be rejected. Similarly, that 19% of routine workers earn more than £30,000. Measuring income is a tough challenge and, although it seems remarkably lax, in the OCJS it was simply asked as one question to the respondent who was given a show card and asked to point to their household income. The respondents could have been any age (down to ten years old). This quite clearly has not given accurate results. The household income then shall not be considered further. The socioeconomic classification variable on the other hand, looks more reliable as it matches with other estimates, was specifically asked of the household reference person and has been shown to be highly correlated with household income (Hills, 2010). This measure, therefore, shall be used as a proxy. It has 195 missing values.

[29] Particularly the wealth data, which holds population estimates by socioeconomic classification available at casedata.org.uk [accessed March 2015]

Table 9.1 Percentage of OCJS sample in each socioeconomic classification, and percentage of each status by household income

	Prop of sample	Percentage within each annual income band						
		Under £10k	£10k - £15k	£15k - £20k	£20k - £30k	£30k or more	Don't know	Refused
Employers or high professionals	14%	14%	8%	36%	6%	19%	13%	6%
Lower professionals or higher technicians	23%	25%	8%	16%	10%	25%	13%	3%
Intermediate occupations	6%	39%	1%	4%	21%	18%	11%	5%
Small Organisations and own account worker	13%	27%	7%	9%	13%	21%	17%	6%
Lower supervisory	12%	35%	2%	5%	16%	24%	15%	2%
Semi-routine	13%	34%	2%	1%	23%	19%	19%	1%
Routine	11%	33%	0%	2%	26%	19%	17%	2%

9.2.3 Higher Level Controls

The most relevant controls to be included at the 'higher level' are structural area characteristics, specifically the index of multiple deprivation deciles. These are included for all respondents in the survey in England and split the interval level English index of multiple deprivation (IMD) into decile groups (all Welsh respondents are set to missing). There are four different relatively methodologically consistent versions (2000, 2004, 2007 and 2010 with plans for a 2015 update[30]) of the IMD, and the one present here appears to be either the 2000 or 2004 version (it is not entirely clear from the documentation which and this is a point that will be returned to shortly). The IMDs are deprivation indexes based at a small area level (which level depends on the year) and were created by researchers from the University of Oxford for the Department of the Environment, Transport and the Regions, later known as the Office of the Deputy Prime Minister (ODPM), a department that has gone through numerous name changes

[30] There is also a 1998 version called the Index of Local Deprivation and although similarities in method are still apparent to the 2000 version some rather significant changes were made.

and is now called the Department for Communities and Local Government (DCLG). Both the 2000 and 2004 methods follow a similar methodology which can be found in the relevant documentation (Noble et al., 2004). The 2004 overall measure is created from a set of subdomain measures that represent small area indices in themselves for income; employment; health; education, skills and training; barriers to housing and services; the living environment deprivation; and a crime domain. These are formed by taking administrative data on measures such as the unemployment claimant count (employment domain) or the number of emergency admissions to hospital (health), the percentage of adults and children in Income Support households (income) or the number of certain types of crime in the output area. These measures are well known, well used and well tested and details can be found in the Department of the Environment, Transport and the Regions (2000) or the Office of the Deputy Prime Minster (Noble et al., 2004) technical reports.

Two subdomains that deserve specific attention are the Income Deprivation Affecting Children Index (IDACI) and the crime measure. The first, the IDACI, is a further subset of the income subdomain and is defined as the percentage of children in the area under 16 who were living in families in receipt of IS (income support) and JSA-IB (income-based Job Seekers Allowance) or in families in receipt of WFTC/DPTC (Working Families Tax Credit and Disabled Person's Tax Credit) whose equivalised income is below 60% of the median before housing costs (in both the 2000 and 2004 data). This was not given the name IDACI in the 2000 measure but was called the "child poverty index". In the OCJS, it is *labelled* IDACI (as in the 2004 iteration) though given the *code* "dchild", in keeping with the 2000 data. The second, the crime subdomain, was compiled for the 2000 data but was found to be lacking data quality and was not included in the overall measure, yet a version of it is present in the OCJS data. It is not clear whether this is the omitted 2000 crime scale or the robustly measured 2004 version, although one may hazard a guess that it is the latter. Fundamentally confusing and due to poor documentation, this could be problematic. A key difference between the 2000 and 2004 indices is that the 2000 set were based at ward level, an administrative electoral division that does not meet the exact standard of the lower super output area level (the smallest Office for National Statistics area classification and

holding between 1,000 and 3,000 people or 400 to 1,200 households) base for the later version. Wards are still of interest as they represent small areas roughly ten to 20 minutes' walk across. They are, however, larger than LSOA, and vary significantly in size, averaging around 5,500 individuals (larger in urban areas) and this variability was the very reason for the development of the super output area classification scheme (see discussion in section 6.4.1). The earliest OCJS is based on 2003 surveys, but was compiled during 2004 so it is possible that the 2004 IMD could have been included, but it seems tight in terms of the machinery of government and release dates. Personal communication with the NatCen team who worked on the survey found that LSOA is present on the unrestricted micro-data (though un-releasable), but whether the IMD measure is the 2000 or 2004 version is unknown, though there is a slight inclination to believe that it is the 2004 due to the inclusion of the crime domain and the LSOA markers. Despite this confusion as to the exact measure present, the possible importance of the IMD and subdomains means that they should not be rejected, rather the possible confusion noted and the measures carried forward. Alongside the overall index the subdomains of crime, income and the 'IDACI' are of interest, though these cannot all be included together due to collinearity, though can be included individually in different specifications. The IMD and its related subdomains are all coded such that a higher number (on the scale one to ten) represents greater deprivation.

As well as the official measure of area structure each interviewer, when arriving at the respondent's household, completed a set of questions about the conditions that they observed in the local area, namely '*How common is litter or rubbish in the immediate area?*', '*How common is vandalism, graffiti or damage to property?*' and '*How common are homes in poor condition?*'. Each of these was answered on a 4 point Likert scale (1 = not at all common, 2 = not very common, 3 = fairly common, 4 = very common) and when analysed in a measurement model fits a one factor solution extremely satisfactorily with CFI = TLI = 1, RMSEA = 0.000 and WRMR = 0.001 and all standardised factor loading over 0.9. As a latent variable any missing values on the manifest variables does not lead to missing values on the latent part as the distributional assumptions mean Mplus estimation does not need complete data. This can be included in the model and referred to as neighbourhood condition, a slight difference in nomenclature from Brunton-Smith

and Sturgis's (2011) analogous British Crime Survey measure that they refer to as neighbourhood disorder. It was felt that neighbourhood condition better represented this concept (and this was discussed in section 6.4.4). As this is assessed by the interviewer, rather than the interviewee, this is an exogenous variable.

One final area level measure that is to be included is a rural/urban split and is present in the data splitting whether the respondent is living in an urban area with more than 10,000 people (80%), town or fringe (10%) or village, hamlet and isolated dwelling (also 10%).

Chapter 6 included school and peer influences as possible risk factors of interest and these would fit here as higher level controls. There are two reasons why these shall not be pursued. The first is that they are not the focus of the current work, which focuses on area level mechanisms at the higher level. The second is that most are simply not available in exogenous, rather than mediating, form. For school influences, for instance, there are a range of questions on opinions about school[31], but these (in the same way that the area and family level variables were included from Chapter 7) would act as endogenous mediating variables. Including these would require the model specification outlined in Chapter 7 to be extended and, without strong theoretical reasons for their inclusion and previous measurement models that can be aped, this would cloud the understanding of what is currently present. Variables such as expulsion or suspension could be included exogenously but have relatively low report rates (148, 6%, had been suspended and 54 had been expelled) and, again, this is not the focus so these are left out. This could be the subject of future work.

One set of measures that stands as something as an exception is peer influences, or to use common language, gang membership. This was the subject of a specific OCJS Ministry of Justice paper (Sharp et al., 2006) which found that around 6% of respondents were members of what they defined to be delinquent youth groups, or DYGs. The definition applied is a version of the 'Eurogang Network' standard

[31] such as "My school has clear rules about behaviour" and "It is easy to play truant or skip lessons at my school"

(Weerman et al., 2009) which includes as part of its definition that "the group has engaged in delinquent or criminal behaviour together in the last 12 months" and therefore there is a strong circularity with any offending dependent: a member of a DYG is only defined as such *if* they offended. More broadly, the definition of a 'gang' has received considerable attention (to say the least, see for instance Cloward and Ohli, 1993) and the Eurogang definition has been subject to considerable debate (Aldridge et al., 2011). This circularity in the possible OCJS definition and the varying ways in which a gang can be conceptualised leaves it as too broad a subject to include here but again could form part of future work.

9.2.4 Review of controls for inclusion and a note on missing values

Table 9.2 summarises the controls that are to be included in the model and whilst there are some obvious missing covariates (ethnicity, a decent measure of income, anything about schools or gangs) these represent a good number of the possible risk factors, and when coupled with those already in the model as mediators they match well and seem relatively comprehensive. Chapter 6 demonstrated that the individual attitude represents much of what has previously been cast as individual risk factors. At the family level the omission of family size is a weakness but cannot be overcome due to data issues, neither can instances of child abuse or time spent in care, as already mentioned in Chapter 6. Parental conflict and family disruption are proxied through separation from the father, and family criminality (to a limited extent in that it relies on the child's recall) is present.

Table 9.2 Control variables for inclusion

Control	Variable	Missing values[1]	Measurement type and notes
Individual			
Gender	Sex	None	Binary, reference male
Age	Age	None	Continuous, range 10 to 16
Family			
Parental divorce	Whether lives with father and if not how often they see	106 (4.8%)	Ordinal: live together, separate but see, and separate but don't see and binary together/separated. Reference together.
Parental criminality	Combination of reported of parental trouble with police/time in prison	188 (8.6%)	Ordinal, never, trouble, prison. Reference never.
Socioeconomic status	NSSEC of the Household reference person	195 (8.9%)	7 category ordinal variable with close relationship to income. Reference category chosen to be intermediate.
Higher			
Index of multiple deprivation and subdomains	Variously IMD decile, crime decile, income decile and IDACI	159 (7.4%) – including all 151 Welsh respondents	These cannot all be included at once and can be considered either as continuous or, more strictly, ordinal. For simplicity these will be approached as continuous variables. Low score is less deprivation.
Neighbourhood Condition	Formed interviewer impressions of area	None	Latent standardised continuous variable
Rural/Urban	Urban/fringe/rural	9 (<1%)	Ordinal with reference category the urban environment

[1] number and percentage of missing values in the longitudinal dataset

In the measurement and structural chapters, little attention was paid to missing values, yet throughout the discussion in this section the number of missing values has been highlighted repeatedly, despite being relatively low at under 9% for all variables. The reason for this is that the variables discussed in this chapter are exogenous covariates and are added to the model without the distributional assumptions of components within the model. This means that when a control variable response for a certain respondent is missing, the person is deleted (case-wise) from the dataset, reducing the sample size.

Including all three family level controls would result in 477 cases being deleted from the dataset, with 440 cases missing on just one value. Adding the IMD variable

increases the missingness to 595 cases, or 27.2% of the sample (adding in the rural/urban split does not increase the number missing). With such a reduction in sample sizes, it is good practice therefore that these components be added as controls sequentially and checks made. Adding each control individually, however, would lead to an unmanageable number of coefficients to interpret and therefore a level of parsimony is needed. As a compromise then, this work shall take the form of adding individual controls, then all of the family and then all of the higher order components together.

9.3 Control Variable Results

Table 9.3 presents something of a smorgasbord of results from the models described above. The results from four different models are presented sequentially, listed as the base model (from Chapter 7), with age and gender as controls, then with the family level covariates added and finally with the area level components. Whether the results are presented in standardised form or not depends on whether the original variable has a meaningful scale. In the case of the neighbourhood condition, assessed by the interviewer and formed as a latent variable, or individual attitude the results are standardised, whereas for categories such as gender, socioeconomic class or guardian's trouble with the police these are in terms of comparisons to the (stated) reference category. For continuous variables such as age and the decile of the index of multiple deprivation they are in terms of a one unit change in the original variable. Results are presented if they are significant at the 90% level (no star) through to 99.9% (three star). Model fit for each model is adequate, despite the lower need for this in structural models, with an RMSEA around 0.035 for each, and a CFI and TLI around the 0.91 to 0.92 level. The dependent variable is the dichotomous reported offending/did not offend in the previous 12 months as used in Chapter 7.

The first four columns of numbers, the effects on the mediators, are regressions of the mediating variables onto the control variables and, as these mediators are latent continuous variables, represent the effect of a change in standard deviations of the mediator either by one standard deviation (in the case of latent controls), one unit (for the continuous), or in reference to the omitted category (for the categorical) of the listed

Table 9.3 Model results

	Effects on mediators				Direct effect on offend	Indirect effects through mediators				Total effect on offend	Proportion of mediation			
	ND (total)	CE (total)	FAR	FAD		ND (total)	CE (total)	FAR	FAD		ND (total)	CE (total)	FAR	FAD
Pathway in figure 9.1	a+b*e	b	c	d	g	(a+b*e) *h	b(i+e*h)	c*j	d*k	Sum all paths				
Base Model														
Individual Attitude	0.33***	0.28***	0.57***	0.7***	0.72***	0.12***	n/s	n/s	-0.11	0.61***	19%***	n/s ind	n/s ind	-18%
Model 1														
Individual Attitude	0.24***	0.27***	0.53***	0.59***	0.68***	0.09***	n/s	n/s	n/s	0.58***	16%***	n/s ind	n/s ind	n/s ind
Female (ref male)	n/s	0.14**	n/s	-0.08	-0.31***	n/s	n/s	n/s	n/s	-0.3***	n/s ind	n/s ind	n/s ind	n/s ind
Age	0.06***	0.02	0.08***	0.15***	0.13***	0.05***	n/s	n/s	n/s	0.14***	38%***	n/s ind	n/s ind	n/s ind
Model 2														
Individual Attitude	0.2***	0.26***	0.52***	0.58***	0.7***	0.08**	n/s	n/s	n/s	0.58***	13%**	n/s ind	n/s ind	n/s ind
Female (ref male)	n/s	0.12*	n/s	-0.1	-0.29***	n/s	n/s	n/s	n/s	-0.31***	n/s ind	n/s ind	n/s ind	n/s ind
Age	0.06***	n/s	0.07***	0.15***	0.11***	0.05**	n/s	n/s	n/s	0.13***	41%**	n/s ind	n/s ind	n/s ind
Live together (reference)														
Separate but see	n/s	0.14*	n/s	n/s	n/s	n/s	0***	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Separate don't see	0.19***	0.3***	0.24	n/s	n/s	0.17*	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Guardians - No trouble (reference)														
Guardians - trouble with police	0.18***	0.17	0.41**	0.31**	0.35*	0.16*	n/s	n/s	n/s	0.36**	45%*	n/s ind	n/s ind	n/s ind
Guardians - Prison	0.31**	n/s	0.45*	0.33	0.54*	0.28*	n/s	n/s	n/s	0.67**	42%*	n/s ind	n/s ind	n/s ind
Employers or high professionals	n/s	n/s	n/s	-0.32**	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Lower managerial and professional	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Intermediate (reference)														
Small Organisations and own account worker	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Lower supervisory	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Semi-routine	0.16*	n/s	n/s	n/s	n/s	0.14*	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Routine	0.13	0.36**	n/s	n/s	n/s	0.12	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Model 3														

	Effects on mediators				Direct effect on offend	Indirect effects through mediators				Total effect on offend	Proportion of mediation			
	ND (total)	**CE (total)**	**FAR**	**FAD**		**ND (total)**	**CE (total)**	**FAR**	**FAD**		**ND (total)**	**CE (total)**	**FAR**	**FAD**
Pathway in figure 9.1	a+b*e	b	c	d	g	(a+b*e) *h	b(i+e*h)	c*j	d*k	Sum all paths				
Individual Attitude	0.2***	0.26***	0.52***	0.56***	0.69***	0.07**	n/s	n/s	n/s	0.59***	12%**	n/s ind	n/s ind	n/s ind
Female (ref male)	n/s	0.1	n/s	-0.12*	-0.31***	n/s	n/s	n/s	n/s	-0.32***	n/s ind	n/s ind	n/s ind	n/s ind
Age	0.06***	n/s	0.08***	0.16***	0.11***	0.05**	n/s	n/s	n/s	0.13***	39%**	n/s ind	n/s ind	n/s ind
Live together (reference)														
Separate but see	n/s	0.13	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Separate don't see	0.16**	0.25**	0.29*	n/s	n/s	0.14*	n/s	n/s	n/s	0.25	54%*	n/s ind	n/s ind	n/s ind
Guardians - No trouble (reference)														
Guardians - trouble with police	0.16**	n/s	0.44**	0.29**	0.34*	0.14*	n/s	n/s	n/s	0.36**	38%*	n/s ind	n/s ind	n/s ind
Guardians - Prison	0.29**	n/s	n/s	0.34	0.46	0.25*	n/s	n/s	n/s	0.6**	43%*	n/s ind	n/s ind	n/s ind
Employers or high professionals	n/s	n/s	n/s	-0.36**	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Lower managerial and professional	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Intermediate (reference)														
Small Organisations and own account worker	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Lower supervisory	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Semi-routine	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Routine	n/s	0.29*	n/s	n/s	0.32	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
IMD decile	0.03***	0.07***	n/s	n/s	-0.03	0.03**	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Neighbourhood condition	0.25***	0.16***	n/s	0.08*	n/s	0.09**	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Urban (reference)														
Town or urban fringe	n/s	-0.19*	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Rural	-0.2**	-0.49***	n/s	-0.18	-0.33*	-0.17*	n/s	n/s	n/s	-0.34**	50%*	n/s ind	n/s ind	n/s ind

*** = 99.9% significant, ** = 99% significant, * = 95% significant, no star = significant at 90% level. Lower significance otherwise

variable. It is worth recalling from Chapter 7 the direction in which these moderators are coded; higher scores are 'more criminogenic', i.e. worse disorder, lower collective efficacy, lower reinforcement or less harsh discipline.

Individual attitude has a relatively stable effect across the model specification, with more criminogenic people reporting more disorder, lower collective efficacy, worse familial reinforcement and lower discipline, with the largest effects being on the familial discipline. Age and gender have similarly stable interpretation across the models with females generally reporting worse collective efficacy and slightly better discipline. For age, for every extra year, reports of disorder seem to increase at about the same rate as lower reinforcement but low discipline is reported as increasing at a greater rate. Focusing on the second model (which is consistent with the third, with area added) whether the guardian has been in trouble with the police and how seriously seems to worsen the perceptions of disorder, reinforcement and discipline. The father being separated from the family home, and the respondent not seeing him, seems to also increase the reported levels of disorder, efficacy and reinforcement. There does not seem to be any real consistent gradient to effects by socioeconomic class in comparison to the intermediate group, with only a few significant results (the highest households appear to have less-strict levels of discipline, the semi-routine and the routine seem to report worse disorder and collective efficacy) that for the main disappear once area structure is also controlled for, indicating that either the effect was erroneous, or linked to the neighbourhood of the respondent.

Finally, there are some interesting effects on the mediators by the structural area measures. Whilst none seem to have consistent and strong effects on the familial moderators, a worse area (as measured by the IMD) increases perception of disorder and low collective efficacy slightly but significantly, and as a ten-point scale means that respondents in the most deprived areas have rather different views to those in the least. Neighbourhood condition increases perceptions of disorder, though the effect is lower than would be expected given the broad similarity in the two measures and points to the difference between the individual respondent's view of their area and the view recorded by the interviewer. Neighbourhood condition also lowers collective efficacy

significantly, but as a standardised variable the effect size of this is broadly comparable to the effect of IMD. Living in a small town or an urban fringe increases perceived collective efficacy, and living in a rural area lowers perceptions of disorder and raises perceptions of collective efficacy (both when compared to living in an urban area).

The next column, titled *direct effects on offending*, is useful in that it gives a handle on the effect that the control variables have on offending without the mediation but in this modelling framework, as there are no instances of full mediation, it is reiterated in a more complete manner in the *total effect on offending* column. As the ultimate regressions onto the dependent variable are probit regressions, as discussed extensively in Chapter 4, interpretation of these values is imprecise and depends on the values of the other components within the model. They are, however, indicative of effect size. In short, and focusing on model 3:

- a higher criminogenic propensity (measured via individual attitude) increases offending (as in the original model, though with a slightly muted effect);
- females report offending less than males, equivalent to about half a standard deviation change in individual attitude;
- older respondents report more offending;
- being separated from the father and not seeing them increases offending over whether the parents live together, but separation still with contact seems to make no difference;
- the guardians being in trouble with the police seems to have rather large effects on reported offending, larger than the gender effect, and if they have been in prison increases it still further; and,
- there are no total effects on reported offending by the structural characteristics of area that are included in the model, bar living in a rural environment (which lowers reported offending).

As for indirect effects through the mediators, and the coupled proportions of mediation, there are few effects to be noted. This is either as the mediator has no effect, or that the total effect of the control variable on offending has no relevance. There are, of course, exceptions and all effects are found to operate through neighbourhood

disorder. Increases in individual attitude are amplified when the respondent also reports lower neighbourhood disorder (and this is partly caused by the routing through lower perceived collective efficacy as previously highlighted in Chapter 7). The age effect, too, is amplified by reporting worse disorder, as is separation from parents, guardian's trouble with the police and living in rural setting. In summary, *perceiving* the area to be worse, to have more disorder, amplifies the effect of these risk factors, and part of that amplification can operate through the perception of the collective efficacy of the area.

What has not been presented here is the modelling using the alternative classification of the index of multiple deprivation, i.e. the subdomains. These results can be found in Appendix A and hold very few differences to the overall IMD measure. This is not overly surprising, though was worth checking, given that the raw LSOA based figures for the 2004 index correlate at over 0.9 for income, IDACI and the overall measures, with the crime correlating slightly lower at 0.6 (and in the OCJS data there is a similarly strong linear relationship).

9.4 Control Discussion

The results presented in this chapter provide an absolutely central component of the substantive findings in this thesis. There are three particularly important findings. The first is that structural characteristics of the family (divorce and criminality) seem to increase the likelihood of offending, shown by the large and significant total effects of parental separation with no contact with the father and by the direct effect of family criminality, yet socioeconomic status seems to have no effect. The second is that the structural characteristics of the area (the deprivation level as measured by the index of multiple deprivation or the subdomains and the interviewer assessed neighbourhood condition) does not seem to impact on reported offending, nor does it change the importance of the mediation by perception of neighbourhood. The third finding is the manner in which the mediating perceptions of the area amplify the individual and family level risk factors, and that this amplification is again operating though neighbourhood disorder.

It should be stressed that individual attitude again plays a vital role in determining the rate of self-report offending, as was previously found in Chapter 7, though with a slightly lower effect. With such a range of other coefficients and findings it would be possible to overlook this most central of components. The results from the previous modelling seem to be robust to the inclusion of these control variables and this chapter does not supersede those results found in Chapter 7 but adds to the knowledge that can be taken from this thesis as a whole. So on top of the effect of the criminogenic propensity and its mediation via the disorder (with efficacy) found in Chapter 7, this current chapter also shows that the other risk factors mentioned in the previous chapter are also mediated in the same manner.

The relationship of self-report offending with age and with gender is not new, though the presence of substantively large effects for both of these variables gives credence to the model, as it would be of concern were these not apparent. Some of the corollaries from this approach in terms of the apparent relationships between the control variables and the perceived mediators are also interesting, that age increases perceptions of all of the mediators, barring collective efficacy, and that females report higher collective efficacy but weaker discipline is noteworthy, if not something that leads to clear policy implications (though could play a role in targeting of any top-down policy implementation).

This is illuminating analysis but there are of course a number of limitations. Only a certain set of controls have been added to the model and it is possible that some of relevance have been missed, such as family size or residential instability. These were unable to be measured using the OCJS and their omissions are unfortunate, though the most important components were included.

Although more than satisfactory in terms of substantive findings, it would also be possible to refine the modelling technique still further. No interaction effects have been included and there could very well be a divergence between, say, the effects of age dependent on the gender of the respondent, or that respondents with criminogenic families live disproportionately in poorer, highly disordered neighbourhoods. This latter point about criminogenic families, however, is not reflected in the data: though

there is a slight gradient towards the percentage of respondents living in the poorest neighbourhoods (if treating the IMD as continuous, the average IMD score is 5.8 for those with a parent who has been in trouble with the police or in prison, whereas for the other group it is 6.5 and therefore less deprived) this difference is not statistically significant (one tailed t-test $p = 0.105$). One interaction that would be deserving of more attention were the results more clearly significant is between the socioeconomic class of the household and the area deprivation level; 25% of the employers or higher professionals live in the least deprived areas, with only 4% in the most deprived and conversely for the routine worker households 2% are in the wealthiest areas compared to 15% in the worst. Given that neither of these covariates seem to affect offending though, there is little reason to suspect that their combination would.

It is also possible that females and males have different behavioural models altogether and that, say, collective efficacy takes on a different role as a mediator dependent on the gender of the individual. This analysis is achievable but would greatly elongate the chapter and there does not seem to be justification in the literature strong enough to imply that gender is such a defining characteristic of young adolescent's perceptions of their environment. The analysis would of interest, though would most likely reduce to that presented here in terms of findings. Parsimony is the name of the game and this analysis, though experimented with away from the write-up and lacking any results, won't be explored further.

Chapter 10

Towards a Longitudinal Model

10.1 Scope of the Chapter

The majority of the analysis up to this point has taken place on a dataset that was specifically created to include only those who were interviewed in all of the 2004, 2005 and 2006 waves of the OCJS study. This has been to give consistency across the piece, and with such low attrition between the waves (losing just 567 respondents) it has no impact upon the results, but use of the longitudinal components of the data has been conspicuous by its absence. This final analytical chapter will attempt to utilise, as far as practicably possible, this extra information and will complete the analytical picture.

It would not have been possible to attempt this longitudinal analysis without the work that has gone before: without the measurement chapter there would have been no components upon which to base the analysis; without the first structural definitions there would be no framework within which to link these concepts; without attempting to model different dependents there would be a tendency to over-stretch that framework; and without the careful procession through the various different control variables to be added there could be erroneous components added to the model that do not shed any significant light upon the matter at hand. To quote Little (2013), "The starting point for a panel model will depend on your theoretical expectations and the guidance from past work. The more theory you have about which effects will be direct and which ones will be adequately captured as indirect relations, the better" (p189).

Longitudinal mediation analysis, particularly with a SEM methodology, is not a simple procedure and many checks need to be made before proceeding. Unfortunately the data at hand only pass some of these (as will be demonstrated) and the complexity of the model renders the most thorough analysis unachievable. As throughout the work, a pragmatic approach will be taken. The central question that this chapter will address is

how the position that a respondent reports they find themselves in at the beginning of the survey affects their reported behaviour towards the end of the period. Before moving on to taking this pragmatic approach, it is first necessary to address why the more 'complete' longitudinal model cannot be constructed.

10.2 An Overview of Longitudinal Modelling with SEM

10.2.1 Measurement Invariance

Of primary importance when conducting longitudinal analysis is that the measures being used are the same at each cross-sectional time point. When using manifest variables as the basis for analysis this assessment is formed from considerations of question reliability and replicability but within SEM this assumption is testable. This is called factorial invariance and is conducted in the same way in which multi-group analysis invariance is tested. Little (2013) provides a short table outlining the different types of invariance that should be considered along with the mathematical formulas that match to each level. The substantive points are repeated in Table 10.1 leaving the mathematics aside and adding an explanation of what each level means.

Table 10.1 Level of factorial invariance

Level	Invariance type	Alternative name	Meaning
0	Configural	Pattern invariance	The model specifications (number of factors and loading pattern) are the same at both time points
1	Weak factorial	Metric invariance Loading invariance	The regression coefficients of each manifest item onto the factors are statistically the same across time
2	Strong factorial	Scalar invariance Intercept invariance	The mean value of latent variables are the same across time
3	Strict factorial	Error variance invariance Residual invariance	The residuals, or error terms, for each manifest item are statistically the same

The first level, level 0 or configural invariance, is more of a qualitative assessment. Are the patterns of associations the same at both time points, and do both models fit (as assessed in the way as described in Chapter 5)? This is straightforward to judge and holds true for all the measurement models that have been used in this work. In testing the 'higher' levels of invariance this model is used as the baseline to which to compare.

The process proceeds by fitting the configural model for each time point (or group, if conducting multi-group analysis), then placing constraints upon the model and assessing whether the resultant implied covariance matrix is significantly different from the baseline models. To assess weak factorial invariance, the regression coefficients are constrained to be equal with no other constraints; for strong factorial invariance the means and the intercepts of the regression coefficients are also set to be equal; and for strict invariance the residuals or error terms of the manifest items are also constrained. Assessment of equality of the model is typically conducted via chi-squared tests and applying a 99% significance level of confidence that both are the same (where the number of degrees of freedom is a function of the number of regression coefficients/means/intercepts/residuals that are forced together that would not have already been fixed for identifiability). Some Monte Carlo simulation studies (Cheung and Rensworth, 2002 and Meade et al., 2006) have suggested that a change in overall CFI of less than 0.01 would be better suited as the chi-squared test is highly influenced by trivial fluctuations. As with a good many tests used in this work and in SEM modelling in general, these rules are not hard and fast but merely guidelines (Little, 2013) and a judgement based on multiple approaches is best.

As there is debate as to the testing criteria for invariance, so too is there some debate as to the level of invariance that is needed before longitudinal modelling can proceed. Little (2013) recommends that strict factorial invariance is not used, in almost all situations. The reason being that this not only constrains the indicator specific information, but also the manifest variables' random error to be the same across time; a rare and overly prescriptive situation. As to whether weak or strong invariance is needed, Little actually puts forward a middle ground, 'partial strong invariance', and claims this to be the bare minimum. Partial strong invariance means that the means/intercepts of the latent variables and the loading of the manifest items are the same, with the exception of one or two (loadings or intercepts).

Mplus can obviously be used to assess all types of invariance though in some cases is a little more involved. Mplus7 implemented a set of commands to shorten checks on single order latent measurement models which do not have a mix of binary and ordinal

variables. To assess more complex models a little more attention needs to be paid, especially with categorical data as the threshold values take the role of the intercepts. Checks for familial reinforcement, family discipline, collective efficacy and neighbourhood disorder separately or combined with their respective partner (area or family) are straightforward and the results are in Table 10.2. The only failure when the most strenuous criteria is applied is found with the chi-squared test when testing invariance of both discipline and reinforcement together between 2004 and 2006 but, as these chi-squared tests are thought to be less reliable than the change in the CFI, all constructs pass and have strong factorial invariance.

Table 10.2 Testing strong factorial invariance for family and area latent variables

Domain(s)	**2004 - 2005**			**2005 - 2006**			**2004 - 2006**		
	$\Delta\chi^2$	**p**	**ΔCFI**	$\Delta\chi^2$	**p**	**ΔCFI**	$\Delta\chi^2$	**p**	**ΔCFI**
Area									
Neighbourhood disorder	4.43	0.11	0	1.79	0.41	0	1.45	0.48	0
Collective efficacy	14.79	0.32	0	12.55	0.48	0.001	18.45	0.14	0.002
Both together	20.71	0.24	0	14.57	0.63	0.003	14.13	0.65	0.002
Family									
Familial reinforcement	1.55	0.46	0.002	0.61	0.73	0	1.52	0.46	0.001
Familial Discipline	4.64	0.59	0.003	4.81	0.57	0.02	9.18	0.16	0.002
Both together	9.01	0.53	0.001	15.03	0.13	0.01	30.52	<0.001	0.002

To test the invariance of the individual attitude construct is where the modelling becomes a little more involved. In that this is a second order construct a few extra steps are needed. Chen et al. (2005) outline the procedure needed to test second order constructs by sandwiching checks at the different orders into those presented in Table 10.1. After fitting the configuration model, the first order factor loadings should be tested for equivalence, then the second order, before moving on to testing the first and second order intercepts and means. They also test for strict invariance at each level, though in keeping with Little's guidance this is deemed over-prescriptive for the current work. Table 10.3 presents the results and these demonstrate that overall, the measures are equivalent across the years. It appears that the 2004 measure is slightly

less consistent, in that when it is compared to either of the other two years it fails the chi-squared test, but judged by the CFI criteria it can still be claimed to be equivalent.

Table 10.3 Testing strong factorial invariance for individual attitude

Domain(s)	2004 - 2005			2005 - 2006			2004 - 2006		
	$\Delta\chi^2$	p	ΔCFI	$\Delta\chi^2$	p	ΔCFI	$\Delta\chi^2$	p	ΔCFI
First order strong invariance	51.30	0.06	0.001	38.28	0.36	0.001	78.75	<0.001	0
First order strong, second weak invariance	52.08	0.05	0.001	39.36	0.36	0	82.45	<0.001	0.001
First and second order strong invariance	83.62	<0.001	0.001	48.91	0.18	0.002	170.53	<0.001	0.002

10.2.2 Sample Size and Power

As noted in Chapter 4, the OCJS questionnaire varies for respondents of different ages. Of particular relevance to the current work is that the questions that make up individual attitude are only asked of those aged 16 years or under. A respondent who was 16 in the 2004 wave was asked one set of questions and at the next interview occasion (assuming that they had had a birthday in-between) they would be asked another set. When trying to link these question across time, therefore, an immediate restriction that needs to be put on the analysis is that all respondents are aged 16 or under in 2006. This reduces the sample size from a possible 2,190 (as has been used when modelling based on the panel data cross-sectionally) to 1,558 and only covers those aged 14 or under in 2004 (though there are three respondents who were 15 in 2006 and presumably must had been interviewed just after a birthday in 2004 and just before in 2006). This was the sample size used in the preceding section.

A sample size of just over 1,500 respondents seems relatively large, certainly larger than the types of sample size used in standard political opinion polls or the like. When assessing fit of longitudinally modelled latent constructs the number of degrees of freedom increases, and hence the power needed to assess fit also increases. For instance, if a latent variable were to have three indicators (manifest variables), this would mean

that when modelling cross-sectionally it would have 24 degrees of freedom. When modelled longitudinally, the degrees of freedom doesn't quite increase exponentially (as for each measurement occasion residuals amongst repeated items are set to be correlated), but it does indeed go up significantly. Little (2013, p131) provides an example and demonstrates that the degree of freedom over three years for a three item latent construct is 261. For more detail, consultation of Little's book is encouraged but the take-home message is that the degrees of freedom increase and consequently the ability to detect misfit decreases. In short, to make reliability judgements about measurement models in a longitudinal setting requires larger sample sizes.

Detecting whether a model fits, however, is only half of the argument; the other being detecting whether a given coefficient is significantly different from zero (or any other specified value). Little (2013) explains that the sample size needed for a SEM model depends on a number of factors, effect size, degrees of freedom and model complexity and rejects traditional rules of thumb for determining sample size such as ten cases per parameter (Bentler and Chou, 1987) but is clear in stating that "complex models, such as continuous time models, longitudinal multi-trait/multi-method models, models with higher order factor structures and multilevel longitudinal models… typically make greater demands on the data and, as such, would require larger sample sizes in order to gain greater precision"(Little, 2013, p127). A full discussion about the intricacies of the calculations for adequate sample size is somewhat beyond the scope of the current work but the guidance is clear. The models in the current work are complex and hold higher order factor structures, so whilst the 1,500 size is not too small, caution needs to be exercised and parsimonious models are the aim.

10.2.3 Longitudinal Structure

Before moving on to start to model the longitudinal structure of the data the first consideration is whether the spacing of the measures is meaningful and adequately covers the development period of interest, and it is of importance that there is some level of change between time points. If one were to be developing a study from scratch, a careful review of the theoretical literature on the developmental time frames would be the initial starting point. For valid reasons of survey methodology and the need for

datasets to address numerous research questions this consideration is often not possible and data collection defaults to the norms of annual or biennial. With the OCJS, data was collected each year and it not clear that this time frame is indeed the most suitable for measuring change. Theoretically, it has been argued in Chapter 7 that criminogenic propensity develops before the age of 10 (the minimum aged respondent in the current work) and is stable thereafter and this can now be assessed. The evidence presented for changes in familial and area perceptions is less well developed and can also be judged here, but whether these cover the developmental epoch needed is unknown.

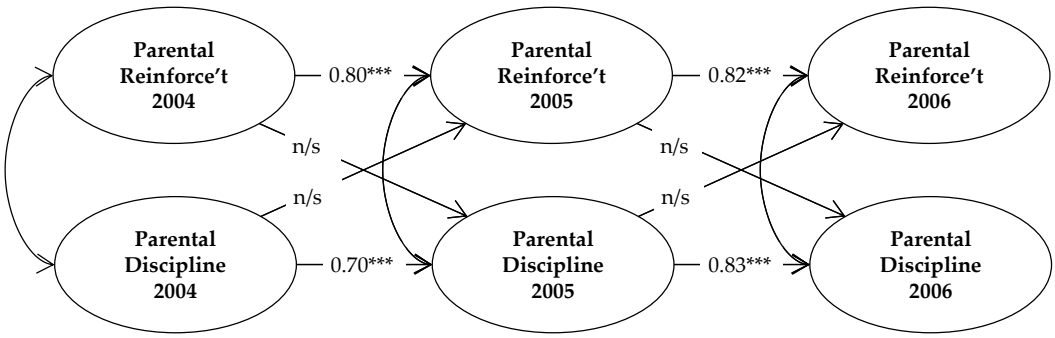
More pertinent is a question about the structure to impose: "with multiple waves of data, a large number of possible relationships can be estimated over time… choosing an appropriate [longitudinal] model to fit data can be challenging" (Little, 2013, p183) and the start point for a panel model will depend on theoretical expectations and guidance from past work. Unfortunately "it is difficult to have a strong theory about all the relationships that will emerge in a longitudinal model for the simple reason that many of the relations will be new and novel in the context of a given study and the context of the particular set of constructs being included in the analysis" (ibid., p189). As with the approach in Chapter 7 to the structural model, it is therefore appropriate to start small and build up models from that low base. By working through the relationships over time between the latent constructs in each domain independently a light can be shone upon the interactions of these elements in a more reliable manner, and some previous statements can be tested empirically.

The first step is to create relatively straightforward *simplex* change models. These test the effect of a construct at time point 1 on the later time points, testing whether change is steady, how much change there is, and whether the time-to-time relationship is affected by other contextual influences. When a construct is regressed onto its previous incarnation this is called an *autoregressive* model, when they are regressed onto a related variable at an earlier time point it is called *cross-lagging*. Single construct autoregressive models are the simplest and test how one item changes over time, cross-lagged models test how one construct affects another. A typical simplex model will include a small number of constructs and have both autoregressive and cross-lagged pathways. As the

latent constructs that have been the main focus of the work naturally split into family, area and individual this distinction will be kept here, starting with the family level relationships. The structure of these models is best represented as path diagrams and the manifest items will again be omitted for the sake of presentability. The residual errors of each item are allowed to correlate across time to account for time point to time point biases.

The first sets of models to be approached are the constructs family reinforcement and familial discipline. Although it is possible to construct models that look at each individually, the results of these are rather trivial and therefore it is okay to jump straight into creating an autoregressive and cross-lagged model; this can be seen, with estimated standardised parameters, in Figure 10.1. These results point to three important findings. Firstly, the respondents reporting of their perception of their familial reinforcement and discipline are very steady over time, with year on year values of over 0.7 for all regressions from time point 2004 to 2005 or 2005 to 2006. This represents an extremely strong relationship; correlations in the social sciences rarely reach such high level and show there is very little change across the study time in respondents' perception of these constructs.

Figure 10.1 Autoregressive and cross-lagged simplex model of family level constructs

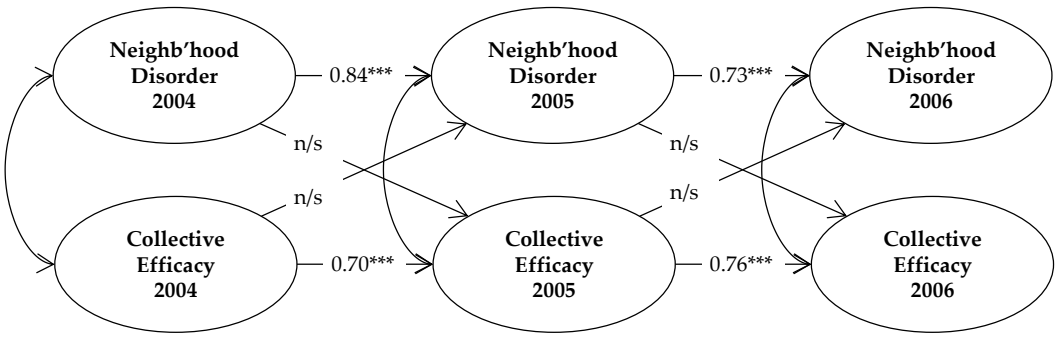


The second observation is that found by a slight modification of the model to allow the 2004 level to regress onto the 2006 version (not pictured), as the coefficient of the pathway from 2004 to 2006 is lower than the value found by multiplying the 2004 to

2005 by the 2005 to 2006 value[32] for both reinforcement and discipline this demonstrates what Little (2013) refers to as "first-order discrete-time auto-regressive lag" (p188) rather than a higher order structure. What this means with a less obtuse sounding interpretation is that the relationship between 2004 and 2006 is not affected by changes in contextual factors, i.e. it is not affected by an external influence and that the model is appropriate and hence so is the identification of stability. The third reflection is that the cross-lagged loadings are not significant. This simply means that the discipline reported at an earlier time point does not affect the reinforcement later (and vice versa) and therefore the covariances between the two constructs at one time (the curved arrows) are synchronic rather than diachronic.

Applying the same model to the area level latent constructs is shown in Figure 10.2 and gives the results that neighbourhood disorder and collective efficacy do not affect each other across time, as shown by the insignificant cross-lagged paths, but this representation does not account for the clear theoretical dependency of neighbourhood disorder on collective efficacy at any one given time as discussed in Chapter 7.

Figure 10.2 Autoregressive and cross-lagged simplex model of area level constructs



A slight variation in this cross-lagged/autoregressive simplex model is needed to account for the known dependency of neighbourhood disorder on perceptions of collective efficacy. This is shown in Figure 10.3 and by adding a cross-sectional regression of neighbourhood disorder onto collective efficacy, indirect effects of

[32] The value of familial reinforcement in 2004 on 2006 is 0.27 compared to 0.43 for the pathway 2004 to 2005 multiplied by 2005 to 2006. For discipline it is 0.12 compared to 0.49, respectively.

collective efficacy at the earlier time point to neighbourhood disorder at the second are introduced. These results are displayed in Table 10.4.

Figure 10.3 Modified autoregressive and cross-lagged simplex model of area level constructs

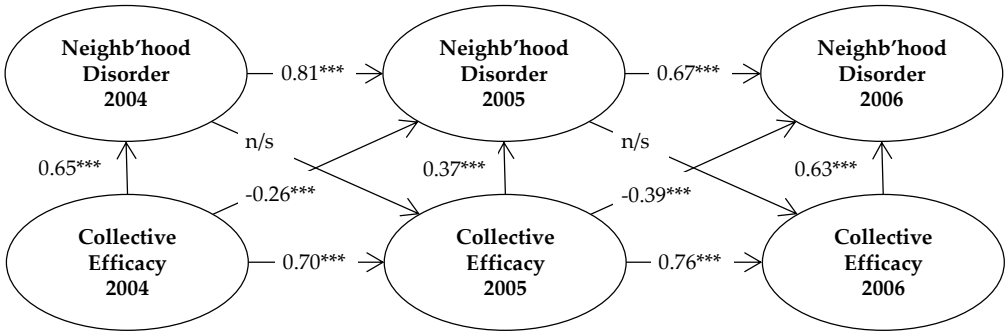


Table 10.4 Total and indirect effects of time point 1 collective efficacy on time point 2 neighbourhood disorder

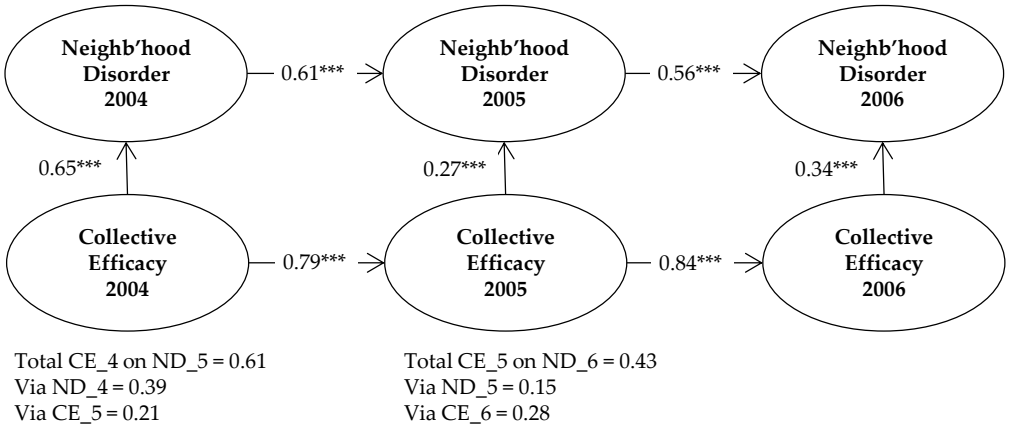
Pathway	2004 to 2005	2005 to 2006
Direct effect	-0.26***	-0.39***
Indirect via time 2 CE	0.26***	0.48***
Indirect via time 1 ND	0.53***	0.25***
Indirect via time 1 ND and time 2 CE	n/s	n/s
Total effect	0.54***	0.36***

The most important finding, again, is the extremely high stability of the measures across time, represented by the horizontal arrows from each latent variable into its later version. With the levels of change this low, it is questionable whether there is enough variation across time to consider working these into a more complete longitudinal model and this will be returned to at the end of this section. Other important findings are mainly to be found in Table 10.4 in conjunction with the figure. Specifically, it appears at first glance that 'worse' perception of collective efficacy at the first time point imply 'better' perceptions of neighbourhood disorder a year later, but once the indirect effects are taken into account the reverse is true, and (as expected) worse collective efficacy leads to a more negative view of neighbourhood disorder. The relationship appears stronger from 2005 to 2006 than for 2004 to 2005.

In light of the total effect being positive and theoretical expectation, as well as the simple cross-lagged structure in Figure 10.2 showing no time to time relationship, the pathway from time 1 collective efficacy to time 2 neighbourhood disorder (for both 2004

to 2005 and 2005 to 2006) has no place in the model. It seems appropriate to reformulate the model once more and remove the cross-lag. This is displayed in Figure 10.4.

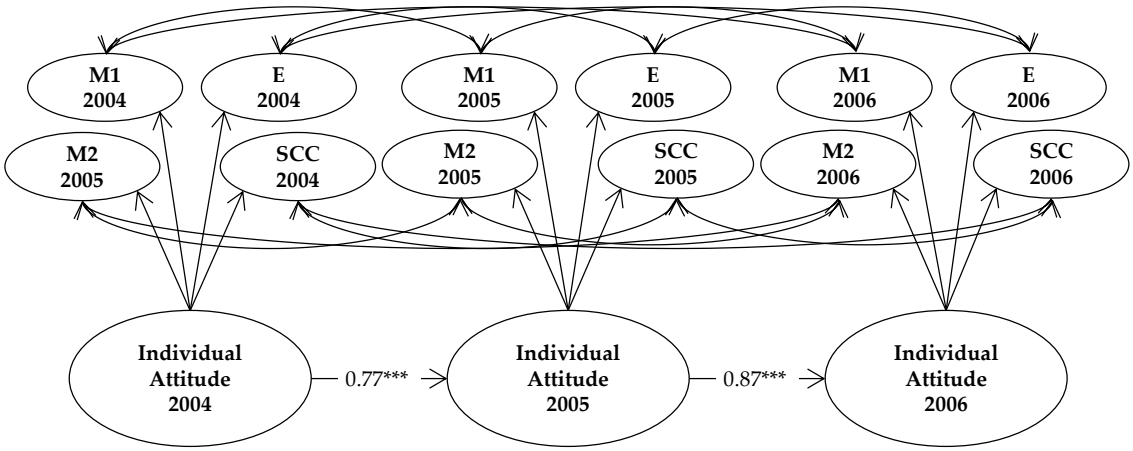
Figure 10.4 Modified autoregressive simplex model of area level constructs with no cross-lag



The regression coefficients from collective efficacy in 2005 to neighbourhood disorder in 2005, and similarly for 2006, appear to be much lower than the 2004 counterpart. However, this is a symptom of the high stability of the constructs. Over 60% of the variance in the latter time point collective efficacy is explained by its previous measurement and in the latter two time periods the cross-sectional relationship is absorbed by the relationship the previous year. This model seems most sensible, and in the context of the two alternatives and their results, it is fundamentally telling a story of stability across time in each of the domains and no cross-time relationships that are not a symptom of the temporal relations.

The final component to be tested longitudinally is criminogenic propensity, alternatively referred to as individual attitude. The theory is clear, this is expected to be constant over time and any departure from this would call into question the theoretical debate that allowed it to be placed as an exogenous component. As a second order latent variable the presentation in Figure 10.5 is slightly different from what has gone before in that the first order latent components are pictured with covariances added between them (meaning that their errors are correlated over time). These are added as covariances rather than regressions to force the relationship to be at the higher level.

Figure 10.5 Autoregressive simplex model of individual attitude



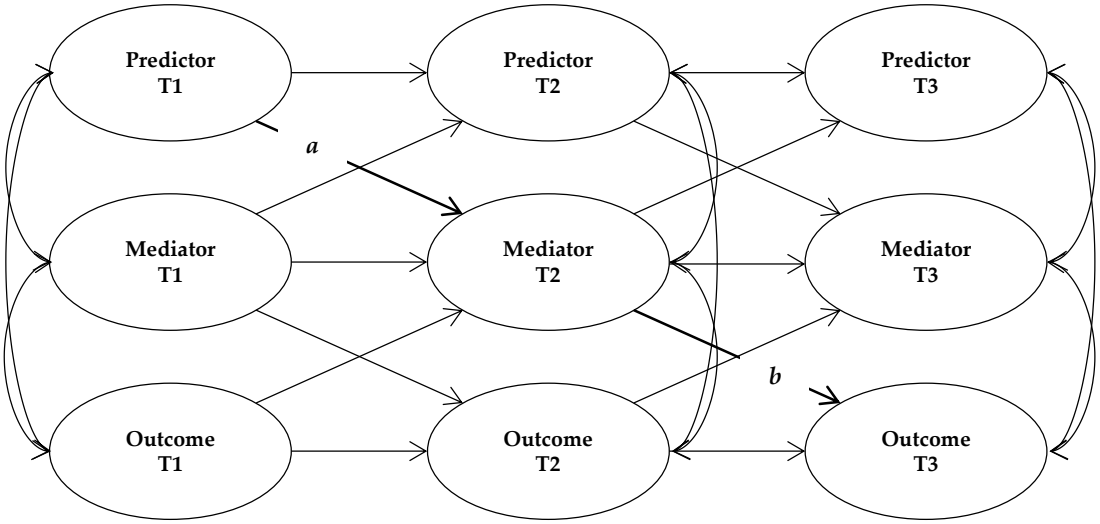
M1 = non-age specific morals; M2 = age specific morals; E = empathy; SCC = self-control/contentment

These results empirically demonstrate the stability of this measure of individual attitude over time. Furthermore, a model that is run allowing the 2006 measure to regress onto the 2004 leads to the regression coefficient being insignificant, meaning that it is entirely mediated by the 2005 time point and there are no contextual factors affecting the relationship (particularly this excludes an age-graded change). The theoretical conclusion in Chapter 7 holds true empirically: individual attitude does not change over the data collection period for the respondents in the survey.

10.2.4 Longitudinal Mediation Analysis

Cole and Maxwell (2003) provide a collection of tips for structural equation modelling of mediation process and suggest a series of five steps. Appropriately constructed longitudinal mediation analysis, in their reckoning, is the primary way to establishing causality: "In that mediation is a causal chain involving at least two causal relations… and a fundamental requirement for one variable to cause another is that the cause must precede the outcome *in time*" (p559, emphasis added). They outline a system of modelling, a slightly simplified version of which is reproduced in Figure 10.6, which captures the mediation relationships in their most general form. A mediation can be said to exist if the pathway a*b from the predictor variable via the mediator to the outcome is significant.

Figure 10.6 Schematic for longitudinal mediation



This is not a model that can simply be picked off the shelf and applied however, as the five tests that come hand in hand with the model framework sequentially test the constructs' interactions *before* moving on to the mediation identification. A full explanation of the steps would be rather longwinded but the takeaway message is one of testing (strong) factorial invariance, followed by creating a saturated model (that is, with all model components allowed to correlate with all others) to identify possible confounding variables or paths and then imposing restrictions by setting some correlations to regressions and constraining others to zero. The piece (Cole and Maxwell, 2003) has been influential and forms the cornerstone of the guidance given by Little (2013) about mediation in his comprehensive treatment of longitudinal SEM. The important consideration for the current work, however, comes in the warnings and requirements that one's data must live up to.

The importance of the timing of the data collection for mediation analysis is even more critical than for longitudinal modelling in general. How much time does it take for the predictor to have an effect on the mediator? And how long does that mediator take to have an effect on the outcome? These questions are non-trivial and Little recommends that a data collection designer run pilot studies and extensively review previous findings. As mentioned at the beginning of this previous section this is not possible and was not done for the OCJS, so it is left to consider theoretically how these components might affect each other and there are some immediate pitfalls that present themselves.

The first is the very nature of the response variable, the question "*In the last 12 months, that is since [month and year 12 months ago] have you [committed offence type X]?"* This situates the offending in the past, whilst questions about family and area simply ask about the respondents' opinion *at the time* of the interview. In the cross-sectional analysis that has gone before, these differential lags were not considered noteworthy due to the tacit assumption that 'that is just how surveys work' and results are drawn regardless. The warning presented here is not intended to call into question the entirety of survey design and implementation, but when dealing with the specific subject of mediated longitudinal analysis the applied statistical guidance is clear - these considerations can no longer be neatly swept under the rug and ignored. It is beyond the secondary data analyst to be able to unpick exactly what cognitive processes were at work when the respondent answered the questions but it seems safe to assume that the responses to the immediate questions that constitute the individual, family and area latent variables have some effect on behaviour either side of the interview's arbitrary cut in the timeline of offending behaviour that the respondent is (or is not) engaged within. Without this precision a solid mediation analysis of the type that Cole and Maxwell build is not reliably possible.

The second pitfall specifically concerns the mediators. As this work uses a range of measures of possible mediation, to model all simultaneously would be assuming that the development of these is equal. That is to say that their causal link from the predictor (individual attitude) at the first time point to second (be that the 2004 to 2005 or 2005 to 2006) is the same for each mediator. Do perceptions of family discipline and collective efficacy have the same time related dependency on levels of individual attitude? This question set cannot be analysed from this dataset as it would require experimental manipulation of the time lags to investigate, as indeed Little recommends as the start point for any longitudinal mediation analysis. It would be possible to model each separately but that would be a huge injustice to the complexity of the situational responses that are present in the dataset and theorised to be of importance. The interpretative difficulties of working with such a complex set of pathways should also not be underestimated.

The final barrier to be mentioned is the within-time dependencies that exist between the individual attitude and the situational mediators as discussed in Chapter 7 and that have been the cornerstone of the work thus far. Little (2013) recommends allowing predictor and mediator to simply co-vary at any one time point as "the focus is solely on whether there is evidence for mediation in the absence of other parameter estimates" (p301). Given the clear expectation these situational factors interact with the individual attitude of the respondent at every moment, to define the model in this way would be unsuitable. The context that the respondent experiences can be expected to interact with their decisions at a later time point, but it appears from the literature to more pertinently be at the time of decision making as well. The within-time dependencies therefore must be included.

10.3 Approach to Be Taken: A Lagged Model

Given the stability (in the sense of Kenny's, 1979, definition that stability "refers to unchanging levels of a variable over time", p231–232) of the constructs present in this work, the reduction in the sample size due to the changes in the questionnaire and the problems with running a 'full mediation' analysis as described by Cole and Maxwell that were highlighted in the previous section, there are some significant barriers to conducting longitudinal analysis in the most complete manner. An achievable alternative is needed and the case where the dependent offending variable is allowed to move forward in time from the cross-sectional set of mediation pathways present in this work since Chapter 7 fits the bill. This certainly gives a parsimonious model but to label it longitudinal would probably be overstepping the mark. For this reason it is best described as a lagged model; lagged in the sense that the first time point (2004) set of mediations are allowed to predict offending at later time points.

At first reading this may appear an *overly* simplistic venture but this approach has a lot going for it. Firstly, the within-time dependencies that have been widely discussed and used throughout this work are easily incorporated. They have good rationale and have produced interesting, enlightening and sensible results. The results will also provide a continuous set of comparable effect estimates. As the models will be the same on all but one variable, comparability of the effects will be automatically possible, and any

movement from significance to insignificance, or in terms of the point estimate size, will be of substantive importance.

More importantly, however, this approach will neatly dodge all of the bullets that could scupper a more complicated longitudinal analysis. As every respondent was asked the same offending questions in each wave regardless of age, the sample size will not decrease. The confusion about the time frame to which the question refers will also not be an obstacle as the analysis will shed light on how situational factors (as perceived) at a given moment will affect reports at that time *or* at some point in the future. The cross-sectional dependencies between the mediators will be included in the analysis as has previous been theorised and the complexity of trying to unpick and interpret indirect cross-time effects will be removed. The assumption that each mediator develop at the same rate will also be removed. And this system is sensible from an analytical perspective. Section 10.2.3 demonstrated by looking at simplex change models for the latent variables that there is a great deal of consistency over time in these measures, with standardised regression coefficients between the years of over 0.6 in each case. Of course there is some stability in offending too (as shown in Table 4.3) but the level is lower than for the other constructs present.

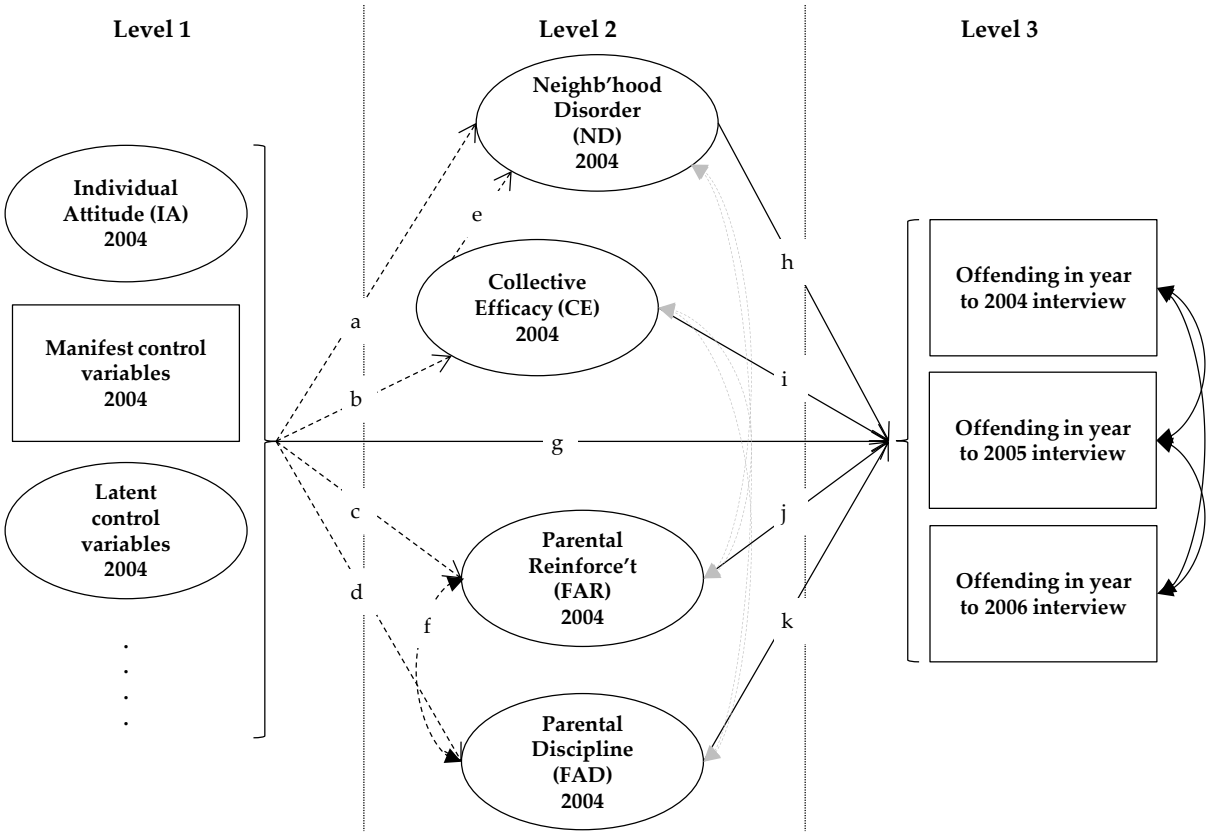
What this analysis will answer, therefore, is how the position that a person reports experiencing at time 1 affects their behaviour at the time, one year and two years later. This will highlight the time variant importance of these factors. Two sets of models will be run: the first will be akin to the analysis in Chapter 7 and will not have controls and the second will be along the lines of Chapter 9 and with the same controls added.

10.4 Lagged Model Results

Figure 10.7 presents the most general form of the path diagram for this model in the same manner as used throughout this work, and as such the control variables are included in the diagram even though for the first set of results they are not included in the analysis. Although this representation holds three dependent offending variables on the right hand side of the figure, these models were also run with each year's dependent variable separately for clarity of the results. As with chapter 8, the

coefficients of regression of the level 1 variables onto the mediators and those between mediators (that is, pathways a through to f and the faint covariances between the mediators) do not vary across the model specifications and are the same as in Chapter 7 (for the uncontrolled model) or Chapter 9 (for the version with controls added). These estimates will be left out of the presentation of results.

Figure 10.7 General lagged model path diagram



10.4.2 Individual Attitude-only Results

The results from the model with only individual attitude added are presented in Table 10.5 and immediately tell a fascinating story. The r-squared of the dependent self-report offending variable (representing the amount of variance that is explained by the model) decreases year on year from 45% in the 2004 model, to 24% in the 2005 case and down to 13% for 2006, meaning that the model does a less good job of predicting whether a person reports offending the further forward the dependent moves. This is reflected in the lowering of the coefficient in the sixth column of the table that reports the total

effect of individual attitude on offending for the later time periods. Despite this, individual attitude still plays the primary role in the model.

Table 10.5 Lagged model results with individual attitude only

Direct effect on offend	Indirect effects through mediators				Total effect on offend	Proportion of mediation			
	ND (total)	CE (total)	FAR	FAD		ND (total)	CE (total)	FAR	FAD
2004 offending with individual attitude only									
0.72***	0.12***	n/s	n/s	-0.11	0.61***	19%***	n/s ind	n/s ind	-18%
2005 offending with individual attitude only									
0.56***	0.06*	n/s	n/s	-0.11*	0.46***	12%*	n/s ind	n/s ind	-23%*
2006 offending with individual attitude only									
0.45***	n/s	n/s	n/s	-0.13*	0.32***	n/s ind	n/s ind	n/s ind	-41%*

*** = 99.9% significant, ** = 99% significant, * = 95% significant, no star = significant at 90% level.

The big story, however, is how the role of the mediators changes. In the cross-sectional work it has consistently been the neighbourhood disorder mediation effect (with around 50% operating though collective efficacy) that has shown the highest level of significance. The proportion of mediation has been roughly equal to that of familial discipline but discipline has hovered on the edge of significance. These results are clear: as the offending measure moves further away from the situational factors present at the beginning of the survey the relevance of neighbourhood disorder greatly diminishes to the point of insignificance whilst the importance and significance level of familial discipline substantially increases. In plain English this means that neighbourhood only mediates at any one time point, whilst family situation mediates against future behaviour (which the fall in the mediation of the former, coupled with the rise in the mediation of the latter demonstrates).

Further, taken in conjunction with the results presented in Figure 10.1, the cross-time stability of reported familial discipline, these results also shed some light onto why the coefficient of the estimate is negative. Remember that a higher value on familial discipline latent variable represents perceptions of a less strict disciplinary regime, shown by the respondents' answers to questions about the trouble they would get into for fighting, skipping school, spraying paint on a wall or smoking cannabis. A negative

coefficient on the mediating effect therefore means that those whose parents are perceived to be *more* strict seem to report offending at a higher rate. This goes against expectations. However, the stability of the construct over the three years in this study means that it is probably fair to assume that this is a continuation of the stability of the home environment's disciplinary regime from younger childhood. According to developmental work reviewed earlier, this has been shown to be the key time for the development of individual attitude and often (at least partly) attributed to poor disciplinary regimes. These results taken together with theory, implies a reversal of the causality of family discipline and moves its effect to before the period under study as has been mooted on numerous occasions throughout this work (though of course this is 'informed speculation').

10.4.3 Model Results with Controls Added

Table 10.6 presents a set of result similar to those in Table 9.3 but with a reduced set of results for the model incorporating the controls used in the previous chapter. For reasons of space and relevance the estimates of the indirect effects have been left out, leaving only the total effects and the proportion of mediation via each mediator. Any controls that were added in Chapter 9 and found to be insignificant across all three dependent offending measures are omitted from the presentation, though all the controls were added and assessed for significance.

Table 10.6 Lagged model results with controls

		Proportion of mediation			
	Total effect on offending	**Neighb'h'd disorder (total)**	**Collective efficacy (total)**	**Familial reinforcement**	**Familial discipline**
2004 control model					
Individual attitude	0.59***	12%**	n/s ind	n/s ind	n/s ind
Female (ref male)	-0.32***	n/s ind	n/s ind	n/s ind	n/s ind
Age	0.13***	39%**	n/s ind	n/s ind	n/s ind
Live together (reference)					
Separate but see	n/s	n/s tot	n/s tot	n/s tot	n/s tot

		Proportion of mediation			
	Total effect on offending	**Neighb'h'd disorder (total)**	**Collective efficacy (total)**	**Familial reinforcement**	**Familial discipline**
Separate don't see	0.25	54%*	n/s ind	n/s ind	n/s ind
Guardians - No trouble (reference)					
Guardians - trouble with police	0.36**	38%*	n/s ind	n/s ind	n/s ind
Guardians - Prison	0.6**	43%*	n/s ind	n/s ind	n/s ind
IMD	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Urban (reference)					
Town fringe	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Rural	-0.34**	50%*	n/s ind	n/s ind	n/s ind
2005 control model					
IA	0.45***	8%	n/s ind	n/s ind	n/s ind
Female (ref male)	-0.22***	n/s ind	n/s ind	n/s ind	n/s ind
Age	0.03	82%	n/s ind	n/s ind	n/s ind
Parents live together (reference)					
Parents separate but see	0.22*	n/s ind	n/s ind	n/s ind	n/s ind
Parents separate & don't see	0.22	n/s ind	n/s ind	n/s ind	n/s ind
Guardians - No trouble (reference)					
Guardians - trouble with police	0.31*	n/s ind	n/s ind	n/s ind	n/s ind
Guardians - Prison	0.64**	n/s ind	n/s ind	n/s ind	n/s ind
IMD decile	-0.03*	-56%	n/s ind	n/s ind	n/s ind
Urban (reference)					
Town fringe	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Rural	-0.2	n/s ind	n/s ind	n/s ind	n/s ind
2006 control model					
IA	0.36***	n/s ind	n/s ind	n/s ind	n/s ind
Female (ref male)	-0.37***	n/s ind	n/s ind	n/s ind	n/s ind
Age	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Parents live together (reference)					
Parents separate but see	0.25**	n/s ind	n/s ind	n/s ind	n/s ind
Parents separate & don't see	n/s	n/s tot	n/s tot	n/s tot	n/s tot

		Proportion of mediation			
	Total effect on offending	**Neighb'h' d disorder (total)**	**Collective efficacy (total)**	**Familial reinforcement**	**Familial discipline**
Guardians - No trouble (reference)					
Guardians - trouble with police	0.31*	n/s ind	n/s ind	n/s ind	n/s ind
Guardians - Prison	0.44	n/s ind	n/s ind	n/s ind	n/s ind
IMD decile	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Urban (reference)					
Town fringe	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Rural	n/s	n/s tot	n/s tot	n/s tot	n/s tot

The results present a much more complicated picture than those in the model without the controls. There are some consistent results, some differences and some that should be deemed spurious, and care needs to be taken with interpretation. Looking first at the total effects on offending for individual attitude, it can be seen that the size of the coefficients again decreases as the dependent moves further forward. Of all the continuous controls present it is the only one that is consistently significant but its effect size reduces by just less than 40% from the 2004 model to the version with offending in 2006. This is coupled with a decrease in the r-squared of the dependent (from 47% to 27% to 21% respective to the years).

It seems natural to compare the other coefficients to the value present for individual attitude given its theoretical centrality and it can be seen that although the coefficient for females, compared to males, in the 2004 and 2005 model is around a half of the size (meaning that an average female would be expected to report offending about as much as a male half standard deviation below the mean value on individual attitude) by the 2006 estimates the relationship is closer to one standard deviation. Although the respondents here are all below the age of 17 in 2004 perhaps this is the starting sign of the divergence in the development of offending rates between the sexes.

Whether the guardian of the respondents had been in trouble with the criminal justice system seems a consistently important factor. Although the significance and point

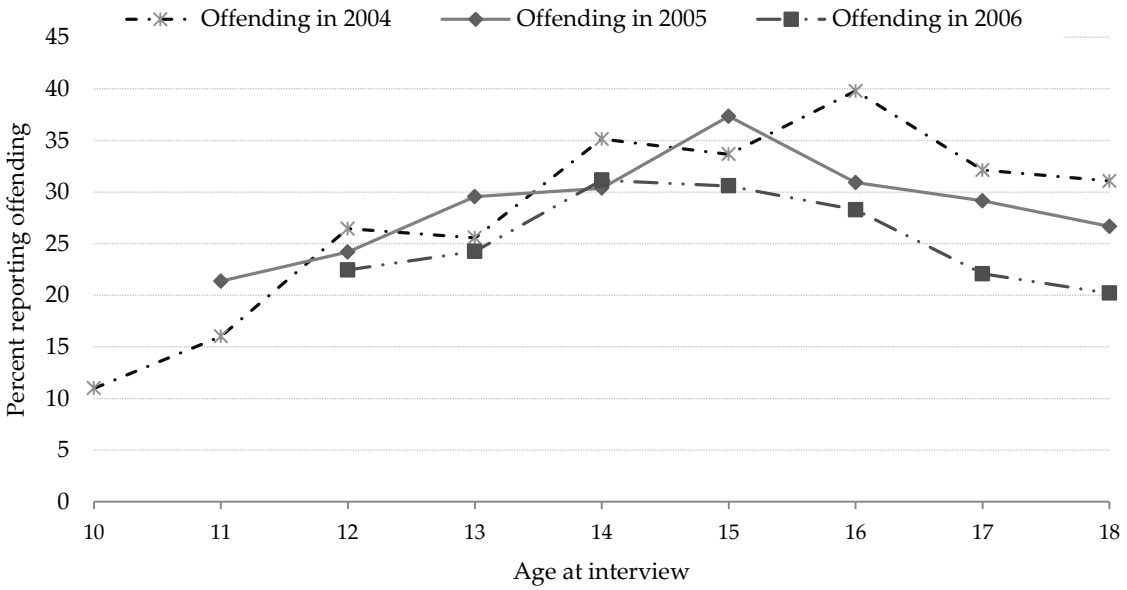
estimate reduces over time, it is the consistency over the models for this relatively small group (only 317 respondents answering that their parents had had any contact with the CJS) that is the important finding. The effect is large, somewhere near a one standard deviation increase in individual attitude for those whose guardian had been to prison and perhaps around half that for any police contact. Living in a rural environment was significant in Chapter 9 but in the 2006 model it is not, despite the point estimate being relatively even from 2005 to 2006, in the latter year it falls out of significance. The IMD decile of the respondent's dwelling has what must be deemed a spurious significant result in the 2005 model because it is far from significant in the other two years and with such a small estimate (just 0.03) is substantively insignificant regardless.

A peculiar finding is the disappearance of age as a significant control variable in the 2006 model. This follows a significant reduction from the 2004 level in the 2005 model. This is surprising given the central role age plays in criminological theory. Admittedly, this is self-report offending and the age-related gradients are less pronounced than the arrest data that the age-crime curve is traditionally associated with but one would expect that increasing age would be a consistent predictor of higher offending. Does this call into question the relevance of age, the reliability of the model or the survey?

The answer seems to be somewhere between the latter two. To call into question the entire age-crime curve would be an outrageous claim, but the age effect in this particular dataset seems unclear. Figure 10.8 displays the pattern of self-report offending in each year by the respondent's age at interview, so the 2006 line only starts at age 12. The 2004 line includes some people who were 17 or 18 in the first wave and are therefore not included in the model but their reports are presented here for comparison. Whilst this chart clearly shows some differences between the waves in terms of the pattern of reported offending prevalence, especially for those aged 16 at the 2004 interview compare to those age 16 in the 2006 wave, the results in the age range 12 to 15 at interview are not *too* dissimilar to call the measure completely into question. There could be an element of respondent fatigue by the time the respondents gets to the third wave and they are consequently under-reporting their offending but this cannot be deduced from the data. Focusing on the 2006 line for a moment, it can be

seen that the differences by age flatten out past 16 in this wave and this is also seen, to a lesser extent, in the 2004 and 2005 waves (though they are at quite a higher level). So the flattening of variation by age in offending once the respondents are in the third wave and between 12 and 18 years of age could cause the lack of results. Regardless of this discussion, the observation weakens the confidence that can be put in the model and this strategy does not trump all others in the thesis: rather, these results need to be taken along with the others and form a *part* of the evidence rather than the *culmination* of the work.

Figure 10.8 Self-report offending by age at interview for each wave



The indirect effects though the mediator variables present a simple picture: there is no apparent and consistent mediation by any of the constructs. The most obvious explanation for this movement from the results found in the previous subsection is that the mediation is controlled away by parental divorce and whether the parents had been in trouble with the CJS. This reinforces the statement made in the previous section about the importance of the family environment become increasingly apparent over a longer time span than can be seen with cross-sectional work. The fact that family discipline mediates the relationship in the uncontrolled model, and then once controls are added that highly influence this meditating variable (as shown in Table 9.3 from the previous chapter), this effect disappears implies that the relationship previously

highlighted is well observed: perceptions of area matter at any one time, but over a longer time frame it is the family environment that matters most.

10.5 Discussion

The aim of this chapter was to incorporate the longitudinal elements of the available data into the analysis. To manage this, a pragmatic lagged model was constructed, one that used the established cross-sectional relationship defined earlier in the thesis and allowing the dependent variable to 'move forward in time'. This was necessary as the most complete mediated longitudinal model as outlined by Little (2013) proved to be beyond the abilities of the data. In assessing the possibilities of the data in terms of mediated longitudinal modelling, two interesting findings in and of their own right were found: 1) each construct was found to be well defined and equivalent in each year of the survey, adding weight to the measurement in Chapter 6; and 2) there was high stability year-on-year in the reports of each of the concepts, particularly the individual attitude component. Problems of sample size, questionnaire design changes, the time frame of the questions and the assumptions that were placed on the data, as well as the complication of assessing multiple mediations in a longitudinal structure all culminated to mean that an alternative to the most thorough longitudinal treatment was needed and the lagged approach was defined. A major strength of this approach, alongside neatly sidestepping the issues that were present with the data, was that it gave comparable results to those that have been found in Chapters 7 and 9. As such, the models presented in this chapter should be seen as supporting and comparable to the work that has previously been defined, though in some senses each of the results from previous chapters are a subset of those found here. It is in the comparability of the estimated effects on each year's dependent variables that the major strength of the analysis lies.

Ultimately, two sets of model results were presented, alongside a set of checks of construct stability, and all of these hold value on which to draw conclusions. Perhaps the most exciting finding is the model without the controls showing that the familial discipline that the respondent reports at the beginning of the study works as a mediator against offending 'two years down the line' whilst the temporally important

neighbourhood characteristics fade into insignificance. The stability of the discipline construct over the study also gives a hint as to why the apparent effect seems to be that a more lax regime actually seems to discourage offending, rather than the other way round. The stability of the measure of discipline over the study can be extrapolated to mean that when individual attitude was developing approximately the same regime could have been experienced, affecting the development of individual attitude in the first place. This extrapolation is clearly rather speculative, but this reversal of the effect would make sense theoretically and has tentative evidence to support the position.

This finding would suggest that although the neighbourhood in which a respondent lives is pertinent to their current behaviour (the Chapter 7 result) it is more important to improve the home environment, or the family setting, to reduce offending behaviour over the longer term. The policy implication would quite clearly be that to do something immediate about a child's delinquent behaviour it would be sensible to try to improve their view of the neighbourhood, their perception of the local disorder (the presence of drunks, noisy neighbours, etc.), and this could be achieved either by directly addressing those factors, or by increasing their perception of the collective efficacy, the 'good neighbourliness', of the area. However, to tackle the root cause of that behaviour and lower its occurrence over the whole of adolescence, it is the individual and the family environment that needs to be supported, and most likely supported from a young age.

The result from the first model, without controls and showing mediation by family perception when looking at offending two years later, was not entirely consistent through the work adding controls to the model (the second substantive model presented in this chapter). The mediation appears to be drowned out by the other measures of family problems, parental separation and parental involvement with the criminal justice system, meaning that there are no apparent mediating characteristics when looking at offending in 2006. This contrasts with the 2004 results where neighbourhood disorder was still a mediator when controls were added, making 'a bad situation worse' as found in the previous chapter. Although there are limitations to the latter analysis and there are some factors that do not conform to expectation which call

the model into slight question (particularly the lack of an age effect), the result seems to imply that external factors are just, if not more, relevant on long-term offending development than the internal perception that the respondent reports. Gender, parental separation and whether the respondent's guardian has been in trouble with the CJS seem to play strong defining roles in the self-report offending of adolescents between the ages of 10 and 16.

The main component that shines though as the most pertinent factor for determining people's offending is. however, their criminogenic propensity, operationalised here as the individual attitude. An important corollary of this work is that it has been empirically demonstrated that this is indeed stable over the course of the OCJS data collection phase, which matches with the theoretically-based claims made in Chapter 7. Measuring the exact context of development of individual attitude is outside the scope of the data present. As all work in the DLC field points to the fact that its development is complete by the age at which this study picks up participants assessing this development would certainly need a more detailed and specific survey design, but once defined by the steps in Chapter 6 the centrality of the concept is not called into question.

Chapter 11

Discussion and Conclusion

11.1 Introduction

This thesis has had two overarching aims, to investigate the role that area plays as a risk factor associated with offending behaviour for those in adolescence (10- to 16-year-olds) and to answer this question using a pre-collected generic dataset, the Offending, Crime and Justice Survey. By employing structural equation modelling (SEM) coupled with detailed theoretically informed measurement models and structural relationships, and focusing on a nationwide analysis, this work has demonstrated that with appropriate consideration and discussion of limitations, it is possible to build enlightening models of the risk factors associated with offending and to uncover new insight into perceptions of neighbourhood, their interrelationship and mechanisms.

11.2 Recap of Background and Research Specification

The initial motivation for the work was to attempt to integrate the well-established development and life theories of offending (as summarised in Farrington, 2005) and their influential risk factor findings (Farrington and Welsh, 2007; Rutter, Giller and Hagell, 1998; Jolliffe and Farrington, 2010; Blackburn, 1993) with the fields of social disorganisation/collective efficacy theory (Sampson, 2012), and the effects of area more generally. Put another way, the aim is to investigate how the area in which someone lives acts as a risk factor for their offending behaviours. Much is known about the various risk factors that can contribute to the likelihood of individuals becoming involved in offending behaviour with there being a well-established evidence base for individual risk factors such as low self-control, impulsivity, low empathy or a lack of morals. There is also a great deal of work that has captured the family type factors associated with offending such as poor parental supervision, childhood physical punishment, low parental reinforcement, erratic discipline and inconsistent rewards,

low parental warmth, child abuse and neglect, and parental conflict. Less is known about risk factors that operate at a higher level, such as school, peer or (the main focus of this work) area level factors. That is not to say that this question hasn't been addressed, and there have been some notable exceptions that have focused in exactly on how the area affects delinquent development (Wikström and Loeber, 2000; Odgers et al., 2009; McVie and Norris, 2006a) and have generally found neighbourhood characteristics to have small but noticeable effects on offending behaviours, usually only for those in the most deprived areas. These reports have often found it hard to separate out these from more general selection effects of 'certain types of people being found in certain types of areas'. Causation has been a major issue that has been tough to unpick.

This lack of specific knowledge of an area's effect on offending behaviour *of the individual* sounds surprising due to the historic importance that area research has had in criminology. From the early days of the Chicago school and the hugely influential work of Shaw and McKay (1942) around social disorganisation, the neighbourhood has played a central part in criminological theory. However, due to a number of criticisms of the work on social disorganisation theory, and to some extent the focus on area more generally, such work fell out of favour during the 1960s and 1970s only to be picked up again through the impact of Wilson and Kelling's (1982) 'broken windows' thesis and subsequently reformulated within Sampson's (Sampson et al., 1997, Sampson et al., 2002) collective efficacy work. The focus of these works, however, was often not specifically on the interplay between the individual, their area and their offending behaviour but addressed wider questions of why certain areas have higher crime rates, higher levels of victimisation or greater fear of crime. Some work has touched on the question of whether the area in which an individual lives leads to their greater involvement in criminal behaviours, notably the Odgers et al. (2009) research that examined 5- to 10-year-olds and calls specifically for an investigation into young adolescents to further knowledge, but these works are unexpectedly thin on the ground.

In some respects, perhaps the lack of attention that this precise type of inquiry has received should be expected, and anticipated due to the (perceived) data requirements needed for area level analysis (Lupton, 2003). The majority of the risk factor type findings have been derived from specific cohorts or relatively geographically constrained locations, at best a city (such as South-east London or Boston – details are available in Chapter 2) and the attention of those with data detailed enough has often led to different focuses. The Project on Human Development in Chicago Neighbourhoods, from which Sampson based a good deal of his work (Sampson, 2012), collected a quite staggering array of data on all aspects of Chicago, and out of this he and colleagues started to formulate collective efficacy theory and 'ecometrics'. That wealth of data available to the researchers has meant, however, that he and his colleagues were able to tackle different, perhaps more policy-relevant questions such as whether neighbourhoods cause crime per se, rather than the narrower question of whether living in a certain type of area increases the adolescent individual's offending behaviour. The focus of work based on the Peterborough Adolescent and Young Adult Development Study, a project run by Wikström based in the University of Cambridge, is about 'bad people in bad places' and has not (until now at least) delved into the question of 'what makes the person bad', whilst the research goals of the Edinburgh Study of Youth Transitions and Crime has tended primarily to focus on the role that institutions play in altering people's lives.

Despite this lack of attention to the exact overlap between the risk factor paradigm and the long tradition of neighbourhood analysis, both research schools have been extremely influential and, either by a subconscious seep into the research community psyche or by individuals on the design committees of large scale data collection exercises making sure that their area of interest is represented in survey questionnaires, sets of questions that relate to the measurement of the risk factors and to the measurement of neighbourhood level process regularly appear, to one extent or another, on publically available datasets designed to facilitate secondary data analysis amongst the academic community. This is true of the Citizenship Survey (Laurence and Health, 2005), the Crime Survey for England and Wales (formerly known as the British Crime Survey) and of the third of the triumvirate of surveys, the Offending, Crime and

Justice Survey (OCJS). The latter panel survey was commissioned by the Home Office and ran from 2003 to 2006, collecting a range of questions relating to the established risk factors and area characteristics for those aged 10 to 16[33], the exact age group for which Odgers et al. (2009) called for more analysis and the subjects of this work.

This led to a set of research questions that were split into two realms, the methodological and the substantive. The methodological were focused on creating appropriate measurements, creating suitably reliable structures that take account of both the measurement limitation and theoretical interrelations. The substantive, on the other hand, focused on making sense of these from a criminological perspective, matching the measurement models to existing work and then interpreting the output from the structural modelling procedure. These two goals therefore are indelibly interwoven and this dichotomy was highlighted throughout as being somewhat false, but was helpful to delineate the goals of the work analytically. The specific research questions were derived from an analytical framework that, it was claimed, could be applied to many different research areas. These amounted to:

1. Can, and how can, criminogenic risk factors be measured?
2. How should they be linked together?
3. What does this tell us about offending behaviour for those aged 10 to 16?

These three questions are the basis of the investigation and give a sequential ordering to the work. The first means, once appropriate data upon which to base the modelling have been found, can a range of criminogenic risk factors be defined from it? This is mainly a methodological step as it requires attention to the detail of measurement structures that have previously been used and to mirror them in the current work. Once these have been established in this first measurement step, how can they be linked together to form a robust model of offending risk factors? Again, this falls more onto the methodological side and necessitates the establishment of causal chains that link the factors together. At both of these steps limitations of the data became apparent,

[33] Full details of the survey design can be found in Chapters 3 and 6, but for the purposes of this section it is adequate to highlight only that the 10- to 16-year-old group were captured.

necessitating constraining the resulting narrative and analysis. The final question falls more onto the substantive side and is where interpretations of the model are given, and checks as to applicability and consistency with previous findings give rise to confidence about reliability of the modelling strategy and any new twists to any findings become apparent. This is where the majority of the substantive findings are held, but is also where the limitations of the modelling strategy come into sharp focus. This combination of both methodological and substantive aims to the thesis means that this work makes contributions on both sides. The methodological in that it demonstrates that with appropriate adherence to theory and previous work better use can be made of secondary data and the substantive in that it can shed light on a realm of investigation that has previously gone under-analysed due to perceived data constraints.

11.3 Data and Method

Before moving on to the modelling part of the work both a data source and a methodology were needed and were discussed in Chapter 4 (data) and Chapter 5 (method). In the review of the datasets to conduct this work on (Chapter 4) it was seen that the OCJS is unique in that it is the only nationally representative (representative of England and Wales and unique to the world) longitudinal survey with detailed self-report offending questions. Covering people aged 10 to 25 and designed specifically to measure prevalence of offending and drug use in the general population of England and Wales, the dataset also holds a range of lifestyle, health, socioeconomic, neighbourhood and individual attitudinal questions. These were deemed to be suitable as to provide a basis of the measurement of risk factors. Crucially the self-report offending module was seen to have followed best practice and due to this should be as reliable as most social science measures. However, for reasons of practicalities and some slightly more opaque design decisions with the data collected there were some issues with questioning structure. Most pertinent were the changes in the questionnaire between the age groups 10 to 16 and 17 to 25. The impact of this was at first not troublesome as it merely restricted the analysis to the population of interest, but it did have an impact on longitudinal modelling possibilities as will be discussed later in this conclusion. The questions relating to neighbourhood changed based on those aged 10 to

15 and 16 to 25 meaning some issues with this question set needed to be addressed further (and were, in Chapter 6). Finding that the 2003 wave questionnaire changed significantly from those conducted over the following three years, and with the highest level of attrition between any two waves, this first wave was removed from the analysis and work proceeded, in the main, on a 2004 – 2006 longitudinal dataset for consistency over the whole piece and with a total sample size of 2,190 respondents aged 10 to 16 years (out of a possible 2,757 in the 2004 cross-sectional data, 79% of the total). By removing this 2003 wave and forming a dataset based on the 2004 to 2006 data there were no weights available to apply but it was shown that with complex modelling the analysis would not suffer from this, and would have been a sensible decision to make regardless.

Structural equation modelling (SEM) was seen to be the most appropriate technique. One of the main attractions of this was the ability to form and test measurements models, and then link these together to account for the theoretical interrelations of concepts included in the model. Graphical, rather than mathematical, displays were highlighted as being particularly useful to SEM approaches as they allow clear and visually understandable representations to be used. A further attraction of the technique was the importance that is placed upon theory in any analysis: this branch of criminology has a long tradition of theoretically informed debate and as such is an ideal test bed for SEM approaches. The ontology that SEM enforces was discussed, and it was argued that a realist position was needed and this implied that the latent variables that were defined via the measurement models need to be considered to cause the responses, rather than simply taking them to be a numerical trick to simplify discussions. This had an impact on the discussion of results. Causality was found to rest on a number of assumptions of the relationships between the latent variables and by applying the technique to survey data such as that available in the OCJS the central role that theory takes in allowing interpretation was again highlighted as vital. Implementing this analytical task within a SEM framework, discussing each measurement model and subsequent linkage between the concepts from both an analytical and theoretical position, allowed robust and rigorous models to be created,

avoiding the possible pitfalls of data mining and results in substantive results that rested on clearly articulated steps. This is a key advantage of the methodology.

11.4 Summary of Results

The results here are discussed as a whole with the substantive and the methodological woven together. As there is a certain sequential ordering to the steps to be made, and these in part reflect the relative importance of the method or the interpretation, some parts of the following discussion fall more on one side or the other and this division will be highlighted. The presentation here, and throughout the thesis, follows the intuitive analytical framework of: measurement; structure; interpretation.

11.4.1 Measurement of Risk Factors

In Chapter 6 the focus was on the measurement of risk factors and as such was a predominantly methodological exercise demonstrating that with appropriate attention to theory and previous work accurate measurement is possible with non-specific datasets. The factors were split into three levels in the same manner as Farrington and Welsh (2007): individual, family and area. On the individual attitude side, a term which is used synonymously with criminogenic propensity throughout the thesis, a second order latent variable structure was defined that captured elements of the first order components of opinions about breaking age-specific rules, a more general moral dimension relating to breaking of laws (predominantly stealing and where the victim was perceived to be able to afford the loss), a measure of empathy and self-control/self-contentment. Crucially, it was argued theoretically that in the literature these elements have often been measured in relatively interchangeable manners and "a story of a closely related and often overlapping melee of individual psychological components that have been seen to be related and predict offending or deviancy" (section 6.2.5) was told. The solution fit the data extremely well with strong face validity and can be seen in Figure 6.3. Some individual domains were missing from the measurement framework, notably any measure of the 'big five' personality traits, though it was argued that these are likely to be strongly correlated with that which the strategy was able to measure. This was a pragmatic approach to the measurement of criminogenic

individual propensity and may not be as rigorous as, say, Grasmick et al.'s (1993) self-control scale nor Jolliffe and Farrington's (2005) basic empathy scale but with the items available and the theoretical discussion to back it up should be considered a parsimonious alternative.

To measure family level risk factors a detailed review of the available questions revealed a rather complicated set of manifest items with numerous routing and missing values issues. A selection was used to measure attitudinal (or psychological) opinions about the familial setting with others used to create factual (or external) factors; a separation that will be returned to in the structural discussion in section 11.4.2. The attitudinal variables that were measurable were perceptions of parental discipline and parental reinforcement, both of which were measured by four items and were relatively strongly correlated with each other, and fit the data excellently (this can be seen in Figure 6.4). Parental conflict, a measure of disruption within families (by divorce/separation combined with visiting frequency) and an indication of any familial criminality could also be measured. Together these were deemed to be the most important risk factors.

For the area level measurement a review of the dataset separated the administrative (such as index of multiple deprivation and rural/urban) from the attitudinal and it was the latter that was the focus. The aim was to create measures of neighbourhood disorder, local fear of crime and neighbourhood collective efficacy as perceived by the respondents when asked *"about 15-20 minutes' walk or 5-10 minutes' drive from your current home"*. This self-defined psycho-social measure of the home area meant that the respondents could define their own sphere of influence and avoid the problems of using administrative data and divisions. Neighbourhood disorder and fear of local crime were relatively simple to measure once the question set was considered in full, although some items were removed from the analysis as they were felt to be looking at slightly different domains. For collective efficacy, a vital component for this thesis due to the theoretical importance of this construct in criminology and broader social sciences, the measurement was slightly more difficult. After review of previous measures that have been operationalised in a variety of different ways, eight items from

the survey were chosen. Due to the data collection process where changes in the administered questionnaire occurred from those aged 10 to 15 and 16 and over, with the younger group being asked a subset of the eight questions, a further reduced set of five items was taken forward. Analysis of those aged 16 showed that the two measures, the full and subset of items, correlated excellently with no statistical difference between the two. This was used to justify the measurement of collective efficacy with the reduced set throughout the thesis. The final measurement of the area level components is Figure 6.7.

The measurement process required detailed review of previous measurement structures that have been applied and in a great many cases (such as the individual attitude first-order components, if not the second order structure) have been used elsewhere. Theoretical considerations were used to fill gaps where the reduced set of items in the OCJS survey meant that full and complete replication of previous works was not possible. The process was a success, with a range of well-fitting and robust measures created that were seen to closely match against the risk factors that should ideally be included for work to progress. This demonstrated that with appropriate attention to detail, measurement of complex concepts is possible with this parsimonious, pragmatic approach, even when using datasets that were not specifically designed to tap complex latent psychological constructs.

11.4.2 Structuring Relationships between Psychological Risk Factors

The structural work in Chapter 7 defined a theoretically sound analytical structure and hence the early part of the chapter fell more onto the methodological side, but a methodology rooted in theoretical discussion. A central debate of this chapter was the theoretical argument that criminogenic propensity (operationalised as individual attitude) develops in early to mid-childhood (before the age of 10 and due to reasons such as childhood neglect or parental problems) and then stays relatively fixed, perhaps with a slow age-graded decrease over the life time. This was seen as being a key construct in all developmental criminological theories from situational action theory (Wikström et al., 2012), Sampson and Laub's age-graded theory of offending (Sampson and Laub, 1993), Moffitt's (1993) dual taxonomy, Gottfredson and Hirschi's (1990)

General Theory of Crime, Wright, Tibbets and Daigle's (2008) neuro-psychological approach to the development of offending behaviour and in Farrington's (2005) Integrated Cognitive Antisocial Potential (though this was reviewed in the background in Chapter 2 and not revisited in Chapter 7). Beyond the identification of a stable and (relatively) fixed-from-an-early-age criminogenic propensity that matched that measured in Chapter 6, each theory specifically acknowledged that this attitude would interact with environmental factors such as family or physical space and that this interaction would vary based on the characteristics of the setting.

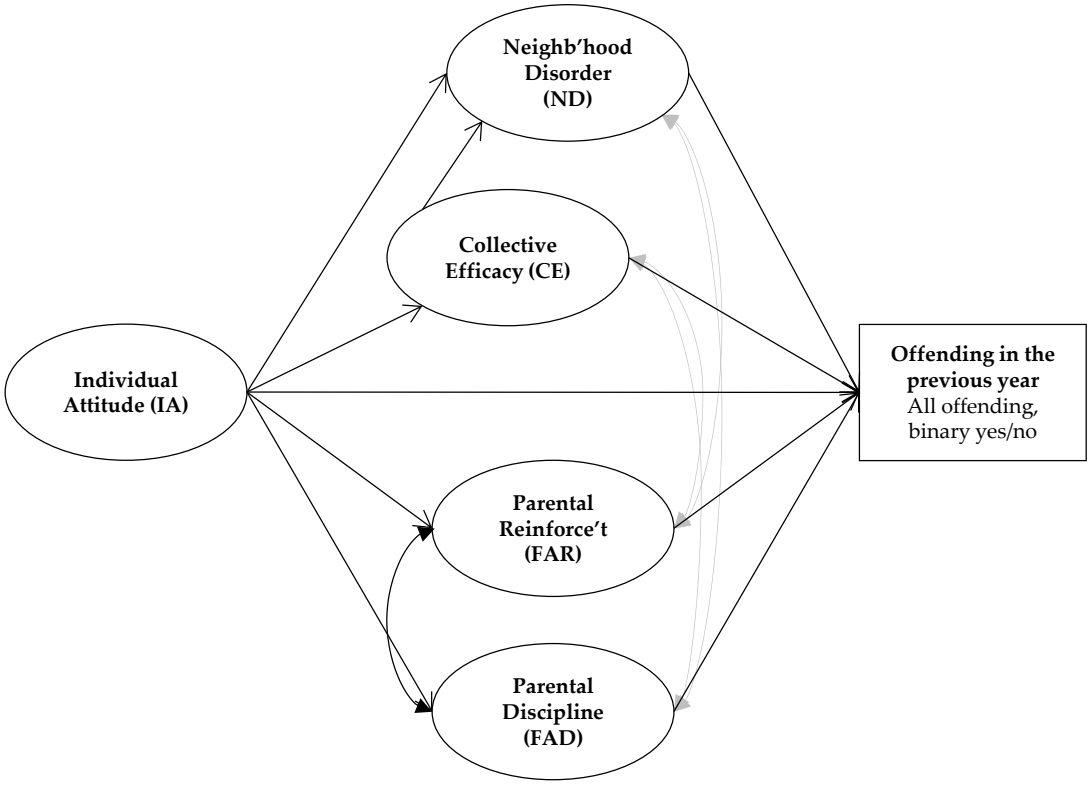
The manner of this interaction required further discussion and, after acknowledging that the latent measures of observations of family and area were based on the individual's psychological perceptions (expressed in a professionally conducted survey) the only way to adequately capture these was via mediation analysis. This was arrived at through two separate strands of argument. The first was by elimination of the alternatives of covariance and moderation. Covariance was seen not to do justice to the theoretical argument that an individual's psychological state affects their perception of the higher level factors, as is well documented in developmental theories. The rejection of moderation rested on the enforced realist ontology discussed in section 5.4.1; for moderation to be appropriate it would mean that there would be *another* psychological construct that was responsible for the combination of the answers to the items measuring the independent variable and items measuring the moderator. Not only did this seem improbable but (footnoted) analysis found this not to fit. Beyond merely accepting mediation by default, psychological models such as that being dealt with imply mediation analysis as the additional variable (family or area level) in the system is a property of the respondent that transforms the relationship between the individual's criminogenic propensity and their self-report offending, and has been a central observation of mediation analysis as far back as Baron and Kenny's (1986) seminal work.

Once the theoretical schematic was established, each domain was tested for its relationship with self-reported offending. Individual attitude showed a strong relationship with the family level constructs of reinforcement, while discipline had a

slightly weaker relationship. Both analyses were straightforward. The area components took a little reformulation and fear of local crime was removed from the analysis as it was felt there was not enough theoretical evidence to link the component into this investigation. Once removed, a suppression effect (Tu et al., 2008) of collective efficacy became apparent and once theory was revisited and closely scrutinised, a dependency of neighbourhood disorder to be caused by variation in collective efficacy was incorporated into the analysis, a move supported by other works.

The final model that was defined is reproduced in Figure 11.1 (without estimated coefficients and leaving aside all manifest variables/the first order components of the individual attitude latent construct). Each relationship examined in the earlier discussions was incorporated into this model. Specifically, the self-report offending dependent (dichotomous yes/no self-reported offending in the previous 12 months) is present on the right hand side of the diagram whilst the placement of the individual attitude on the very left-hand side indicates its stability and mediated interaction with the psychologically perceived higher level latent constructs of family discipline/reinforcement and collective efficacy/neighbourhood disorder. Similarly, the dependency of disorder on levels of perceived collective efficacy is also present in the regression (single-headed arrow) between those two measured concepts. The covariances between the mediating variables are presented by light grey double-headed arrows but was found to be insignificant in all models and eventually omitted from the diagrams. This specification thus becomes the 'psychological map of interrelations' between the attitudinal latent constructs and became a central part of the thesis. Its strength lies in the manner in which it was derived, via sequential steps that leave little possibility for a reformulation without calling into question either the stability of individual attitude (present in all developmental theories), the dependency of perceptions of neighbourhood disorder on perceptions of collective efficacy (Sampson and Raudenbush, 1999; Reisig and Cancino, 2004) or the clear reasoning behind defining psychological interactions as mediation analysis (Baron and Kenny, 1986; Mackinnon, 2008; Little, 2013).

Figure 11.1 Empirical specification of the relationship between perceived latent constructs



Although the main aim of the chapter was to demonstrate that it is possible, given good measurement, to link criminological concepts to provide a basis for understanding the risk factors associated with offending, it only stands if the substantive conclusions that derive from it make sense. To that end, an interpretation in terms of a fictional respondent gave the following outline of an individual with a higher criminogenic propensity:

- they live in a household where they receive a lower level of parental reinforcement, or at least perceive it as such;
- and they perceive living in a home environment with lower levels of discipline (a more permissive home environment);
- but these differential levels of perception of the family environment have no significant effect on whether the respondent self-reports offending (although the level of discipline verged on significance).

The more criminogenic individual is also likely to:

- perceive the area in which they live as having lower levels of collective efficacy;
- but the perception of the collective efficacy appears to have no significant direct effect on the rate of offending;
- but the more criminogenic individual is likely to rate the area as having a higher level of disorder and anti-social behaviour;
- and this effect is amplified by the perceptions of lower collective efficacy, meaning that the perceptions of disorder are made worse by perceptions of low collective efficacy;
- the perception of the disorder in the neighbourhood seems to significantly and substantially increase the rate at which the person would self-report offending.

These findings raise a number of questions and some substantively significant findings. Firstly, a large number of studies have found evidence for a family level effect on offending behaviour yet this strategy does not, or does so with only emergent significance. The emergent significance of the mediation through discipline was in the opposite direction to that which would be expected, which led to the hypothesis that, as the family environment has been deemed a factor in the development of the individual attitude in early childhood (before the age of 10) and there will undoubtedly be some level of consistency over the childhood to adolescence transition, the reversal could be a symptom of family environment from the past. This became a key discussion throughout the remaining analysis in later chapters.

For the area level factors the implication is clearer yet less well-observed in previous works. Whereas the collective efficacy of an area (perceived or measured from other sources) has been shown, rather inconsistently, to affect offending (and often only amongst those in deprived areas) this work points to a tangled interrelation of the perceptions of area characteristics. Perception of neighbourhood disorder is higher, and collective efficacy is perceived to be lower, among individuals with more criminogenic attitudes. A negative perception of neighbourhood disorder plays an important role in shaping whether an individual offends beyond just their criminogenic propensity, with those living in areas that they perceive to have one standard deviation higher disorder offending at around the same rate as an individual with a fifth of a standard deviation

higher score on their criminogenic propensity measure. Put another, more precise way: the perceived disorder mediates around 20% of the relationship. Further lower collective efficacy amplifies this effect (with just over half of the mediation operating through this latent construct) even though when analysed individually, collective efficacy appears to have no effect.

Overall, Chapter 7 demonstrated that it is indeed possible to create a viable model of the psychological constructs that lie behind a young individual's answer to the set of questions present in the OCJS, and the role that theoretical positions played in this formulation was extremely important. These relationships gave substantively important findings to be further examined in later analysis: the level of perceived disorder in the neighbourhood seemed to increase self-reported offending, with the levels of family discipline bordering significance meaning that a definitive statement could not, as yet, be made.

11.4.3 Modelling Different Types of Offending

Before moving on to start to add controls to the cross-sectional model this thesis first considered different types of offending. Motivated by papers that have studied specific types of crime (such as Sampson, Raudenbush and Earls, 1997; Browning, 2002; Morenoff et al., 2001; McVie and Norris, 2006b) and by a wish to test the different definitions of offending (such as the unrestricted self-report offending frequency which was treated as ordinal and then linear) that are present in the OCJS, Chapter 8 started with a review of the measures available. Finding only a limited number of different types of offences being reported the work moved on to consider three subtypes of offending that were labelled property offences (predominantly theft from school or a shop), violent offences (any type of assault, with or without injury) and serious offences (mainly assault with injury). Alongside this a grouping of those who reported the most frequent offending (over six or more offences in any one year) was created to echo the importance of prolific and priority offenders (i.e. PPOs, see Hopkins and Wickson, 2012). The relatively small numbers of respondents reporting each offence type was highlighted early on as problematic for finding significant results, and a strategy to run

the analysis on larger datasets (such as the cross-sectional 2004 data or using a pooled version of the 2004 to 2006 combined dataset) was put forward.

Individual attitude was still the dominant factor in the each of the models and reading the mediation results across the multiple specifications of the dependents/datasets led to tentative findings that neighbourhood disorder was a consistently strong mediator for all and violent offending. Property offending, which arguably requires more forethought, and serious offending, including more extreme violence, had lower mediation and hence the individual propensity plays a more important role as the deciding factor. For frequent offending this mediation was analytically even less apparent, indicating that the respondent's behaviour is harder to keep in check and moving them closer to the chronic offender type. Neither of the collective efficacy total effect or familial reinforcement mediated in any model specification nor for any offending type. Familial discipline presented a mixed picture in keeping with the Chapter 7 findings, often bordering a significant effect, and with a point estimate of approximately the same magnitude as the neighbourhood disorder total effect across most offence types and most specifications. However, for this mediator the point estimates across different datasets (the longitudinal, 2004 and all years combined) rendered the results a little too unreliable to allow firm conclusions and family discipline again proved hard to unpick. Overall, these results were deemed less reliable due to the separation between the crime types and the (approaching) low response rate. Whilst still valid, especially the disorder reduction across different crime types, it was decided not to extend this subgroup analysis in the later part of the thesis due to possible unreliability, and instead to focus on overall offending/non-offending.

11.4.4 Adding Non-psychological Constructs

The central concern of the next part of the work was to add risk factors that had previously been omitted to the modelling strategy. Chapter 7 structured and analysed only the psychological (or self-perceived) risk factors whilst Chapter 9 added the non-psychological items that were present in the dataset and able to be measured. The results from this piece of analysis therefore fall more onto the substantive side, though as a cross-sectional piece of modelling the methodology that is used should be widely

applicable to other cross-sectional data sources, albeit a methodology that rests on the discussions in Chapters 6 and 7. The key contribution of this section is that the analysis allows the effect of the perception risk factors to be assessed alongside the external measures.

The first step was to consider what could be controlled in the model using the OCJS data. This took the form of reviewing the dataset and the available items, some of which had previously been discussed and left for inclusion from the measurement section. At the individual level these were age and gender, for clear theoretical and well-established criminological reasons, but sample size/lack of appropriate items meant that other controls couldn't be assessed (notably ethnicity). For the family/household level (terms used synonymously) parental divorce was proxied through whether the respondent reported living separately from the father combined with contact frequency, whilst parental criminal involvement was whether that respondent reported the parent being in trouble with the police or subsequently jailed. The derivation of both of these was made in the measurement chapter and seen to be reliable. Socioeconomic classification of the household was chosen over family income as the latter was deemed unreliable. Higher level controls included neighbourhood condition, assessed by the interviewer and via a latent exogenous variable out of three items (presence of litter, vandalism and poor condition housing), a rural/urban split and the decile group of the Index of Multiple Deprivation (Office for the Deputy Prime Minister, 2004) and its subdomains of income, crime and the Income Deprivation Affecting Children Index (IDACI). School and peer (and gang) possible risk factors were left from the analysis as outside of the focus of this work.

Analysis proceeded sequentially, adding first individual, then family and finally area level control variables, though this summary shall focus on the full model (named model 3 in Chapter 9). The first substantive results came from looking at the effect of the control variables on the mediators. Individual attitude had a similar relationship with the mediators as explained in the structural results from Chapter 7. Females reported worse collective efficacy and slightly better discipline than males. For respondents a year older, reports of disorder increased at about the same rate as

reinforcement decreased and with perceptions of discipline reported as decreasing at about twice the rate (a change of 0.16 standard deviations per year older). The father being separated from the family home and the respondent not seeing him increased the reported levels of disorder, poor efficacy and reinforcement (with coefficients of 0.16, 0.25 and 0.26 of a standard deviation in the dependent respectively). Socioeconomic class appeared not to affect any mediating perceptions. A worse area as measured by the IMD, or any of the subdomains that were seen to be extremely similar to the overall measure, increased perception of disorder and low collective efficacy slightly but statistically significantly. Living in a small town or an urban fringe increased perceived collective efficacy by a fifth of a standard deviation, and living in a rural area lowered perceptions of disorder and raised perceptions of collective efficacy by a fifth and half of a standard deviation respectively (when compared to living in an urban area).

The effects of the controls when analysed together can be summarised as:

- a higher criminogenic propensity (measured via individual attitude) increases offending (as in the original model, though with a slightly muted effect);
- females report offending less than males, equivalent to about half a standard deviation change in individual attitude;
- older respondents report more offending;
- being separated from the father and not seeing them increases offending over whether the parents live together, but separation still with contact seems to make no difference;
- the guardians being in trouble with the police seems to have a rather large effect on reported offending, larger than the gender effect, and if they have been in prison increases it still further; and,
- there are no total effects on reported offending by the structural characteristics of area that were included in the model, bar living in a rural environment, which lowers reported offending.

Increases in individual attitude are amplified when the respondent also reports lower neighbourhood disorder and this is partly caused by the routing through lower perceived collective efficacy. This result is consistent, though slightly small (with a

proportion of mediation of 12% compared to 19%) than in the uncontrolled structural work, meaning that some of the relationship was absorbed by the control variables. The age effect, too, is amplified by reporting worse disorder, as is separation from parents, guardian's trouble with the police and living in a rural setting. In summary, perceiving the area to have worse disorder makes these bad situations worse. An important piece of information is that the addition of the structural area characteristics did not remove the effect of the perceptions of the area as mediating effects and this shows that these perceptions play a larger role in governing offending behaviour than the structural characteristics (as were available to be measured from this data source).

11.4.5 Incorporating Longitudinal Data

After successful creation of good measurement models with the limited variables available from the OCJS, and finding suitable structure into which to fit both the perceived and the non-psychological variables, the final piece of analysis in Chapter 10 attempted to incorporate the longitudinal possibilities of the data. After some checks on the invariance of the measurement models over the course of the study found them to be consistent across years, in a similar manner to Chapter 7 the work started examining the relationship within each of the 'levels' of risk factors individually. This helped clarify relationships to be incorporated into later analysis and provided as a corollary some interesting findings. Specifically, it was shown that the reporting of within-family level of discipline and reinforcement were extremely stable over time, and the same could be said for reports of neighbourhood disorder and collective efficacy once a suitable model was pinned down. Perhaps the most important finding of the preliminary analysis was that the individual attitude construct was extremely stable over the course of the study as would be predicted from the theoretical literature, which led to increased confidence in the measurement (although it could be viewed that if the measurement validity were already established, as claimed, it would lead to increased confidence in the theory!).

Creating the most robust longitudinal analysis in the sense that Little (2013) explains proved beyond the data and perhaps would not have been warranted given the stability of the constructs over the period of the OCJS data collection. A suitable

alternative specification using a lagged structure and looking at how the essentially cross-sectional model used as the basis from Chapter 7 onwards predicted self-report offending sometime in the future (either 2005 or 2006) was used. This specification had the added bonus of neatly dodging a number of problems with a full longitudinal structure that the data limitations enforced. A model without controls, and then with, were constructed and some fascinating results came to light.

The uncontrolled model gave insight into how the mediators vary over time. Although neighbourhood disorder was the only consistent cross-sectional mediating factor in the model, as the dependent variable moved further forward in time the significance of this diminished to insignificance. This was accompanied by an increase in the significance (both statistical and substantive) of the mediation by family discipline. This means that although the neighbourhood in which someone lives affects their behaviour at any one time, with worse disorder (partly caused by worse perceptions of collective efficacy) increasing self-reported offending by around a fifth of a standard deviation compared to what would be expected based merely on their individual attitude, two years later this effect has disappeared and it is the family discipline that matters, and matters greatly at two-fifths of a standard deviation of individual attitude. Although the negative coefficient demonstrated that stricter disciplinary regimes increase offending (against expectations), the stability of that concept over the three years of the study and extrapolating backwards to before the study began (where the respondent was in young childhood) went some way to explaining why. It could be that this negative sign is a symptom of the continuation of a harsh disciplinarian regime that was (part of) the cause of problem during the development of the individual attitude the first place.

The lagged model with controls went some way to reinforcing this statement. When the controls were added, the inclusion of separation of the parents and family criminal justice system involvement meant that the mediation effect over the longer term by perceptions of familial discipline disappeared. Both parental separation and parental police contact in 2004, however, increased the likelihood of self-reported offending two year later, and with a large effect (approximately equivalent to just under a standard deviation of individual attitude for both). Given that the these two control variables

(separation and parental police contact) were seen to be strongly related to perceptions of family discipline in the cross-sectional control work (reported in Table 9.3 of Chapter 9), the removal of the mediation becomes apparent and the importance of family level influences *in general* on self-report offending over the longer time period is clear. So family circumstance matter, and matter greatly (both structurally and as perceived) but their relevance is seen over the longer term. On the other hand, the neighbourhood level factors seem to be a) operating through perceptions and b) only relevant at any one time point.

Although a central finding there are some issues with the reliability of the results that means these conclusions need to be made in conjunction with other findings from the thesis rather than as a replacement. The insignificance of age is one worry, though the flattening off of changes in self-reported offending after the age 14 or so (shown in Figure 10.8) goes some way to assuaging doubt. The second worry is the precision of the model: as the dependent goes further into the future from the cross-sectional interaction of the mediating variables and the controls, the results become less reliable and the myriad of other factors not included in the analysis could play an unknown role.

11.4.6 Overview of Results

It is hoped that the results is this work form something of a coherent whole. The measurement section laid the basis for further work by providing a platform upon which the rest could be built, whilst the structural discussions and resulting model specification led to a reliable structure that went on to become the central part of the thesis. This gave the first insight into the importance of perceptions of collective efficacy and neighbourhood disorder, and highlighted how the role of collective efficacy only operates through neighbourhood disorder. After a slight detour into attempting to stretch the analysis to cover subtypes of offending which gave results a little too tentative to conclude much from (with confidence at least), the addition of controls to the modelling strategy showed that from a cross-sectional viewpoint, it seems that perceptions of neighbourhood disorder and collective efficacy are more important in shaping offending behaviour than the deprivation level of the area, the socioeconomic

status of the family or perception of the characteristics of the family environment. Structural family problems, as measured by parental separation and parental trouble with the criminal justice system, also seem to have a direct effect on the behaviour of the adolescent, associated with increases in their self-reported offending behaviour. The final piece of the jigsaw came from the lagged model and told a story of the reducing role of temporal perceptions of area on future behaviour and that it is actually the family environment that is most important in determining offending behaviour over the long term.

These results have all been achieved from analysis conducted on a publically accessible dataset without the need for area level identifiers or special access to sensitive variables (or expensive specialised data collection), but some of the substantive findings are novel. The interaction of neighbourhood disorder with collective efficacy and their impact on self-report offending does not appear to have been uncovered in the literature before, and the subsidence of the neighbourhood to be taken over by the family in terms of longer-term offending patterns similarly is a new take on work that has gone before. To be able to make these conclusions has taken a good deal of space, with theoretically based measurement and structuring along with parsimonious steps taken to combat vague, incomplete data and/or definitions. This process is certainly longer than a standard journal publication and the sequential nature of the work means that to have treated the thesis as a coherent whole was the correct decision.

11.5 Limitations of the Data Available

There are two types of limitations in this work; those within the current approach that have been an integral part of the analyses and those that flow from the nature of the data. Given that the stated goal of this work was to attempt to investigate this realm with existing data it would be amiss to allow too much focus to fall on this latter set and these wider limitations will be revisited in the next section and discussed in terms of implications for data collections and survey design. Most of the existing limitations to the analysis have been an integral part of the narrative already and have very much shaped the results and analytical direction. This current section therefore focuses on the existing limitations that have not already been made explicit.

The first of the existing limitations to discuss here were those imposed by the software. Although Mplus was deemed the most versatile for this investigation, especially with its advanced and well-established range of estimator algorithms (in particular the WLSMV for categorical data) and its computational power, some features that could be wished for are not present. Notably here is the link function available for the modelling of a dichotomous outcome variable with categorical data. This was limited to the probit link and therefore the interpretation of coefficients instead of being in terms of "holding all else constant…" were in terms of "holding all else constant *at their mean value*". To some extent this was not too troublesome as standardisation was used throughout the work and holding a value at its mean of zero eliminated it from most calculations. The predicted probabilities used mainly in Chapter 7 and the proportion of mediations approximation (which superseded the predicted probabilities and became the main interpretive tool) did suffer an imprecision, and results had to be read in comparison to each other rather than as statements of absolute value (for instance the proportion of mediation for neighbourhood disorder is derived as 19% but the discussion has phrased this as "around a fifth" and made rough comparisons to other values). This is more of an inconvenience than a major limitation, but is worth noting.

A more serious limitation was the questionnaire design and the changes to question sets that were asked of respondents that varied depending on their age. Although the majority of the changes occurred between 10 to 16 years old and those aged 17 plus, for the area level questions the changes happened for those 15 and below and the older group. Although the 10- to 16-year-old group was seen as being of greatest analytical interest, the change in the questionnaire impacted the longitudinal analysis and had an alternative approach not been taken could have greatly reduced sample size. This also means that these results cannot be checked for equivalence for the older group.

Another apparent limitation with the questionnaire design was the loss of the 2003 wave. This was mainly due to attrition between that wave and the subsequent ones, but also some additional questions were added that became crucial to the analysis. Although the loss of the 2003 wave was unfortunate, in many ways the team in charge of the administration of the OCJS should be commended in recognising the problems

that were occurring, both in terms of attrition and question shortages, and addressing these in time for the second wave.

Missing data in general has not been too much of an issue due to the way in which latent variables and the WLSMV estimator computes latent variable scores. This uses full information and has been shown to provide consistent estimates when "the amount of missing data is not substantial" (Asparouhov and Muthén, 2010, p8) as here. Missing questions did mean that the collective efficacy latent variable had to be run on a reduced set of questions but this was unproblematic. When adding covariates however, such as the IMD (not valid for those in Wales) and various family level questions, the missing data resulted in a few (around 10%) of the respondents falling out of the analysis. Better documentation and coding in the production of the data would have helped here, but the impact was overall rather manageable.

It would be possible to lament the lack of variation across the offending subtypes, the reliability of the income questions, the unavailability of area level identifiers, interviewer characteristics (which have been shown to be important in the field of fear of crime and could very well have an impact here; see Brunton-Smith, 2006, for more details) or other missing questions but these worries shall be saved for this wish list for future data. Overall, despite the data issues, the thesis has demonstrated that this analytical approach, a methodical, rigorous and theoretically informed procession through each step can provide fresh insight into problems that have traditionally been considered too complicated to answer.

11.6 Implications and Wider Limitations

After reviewing the findings of this thesis, both methodological and substantive, the following discussion will consider the implications of the results from a policy perspective to reduce offending behaviour, for approaches to modelling complex topics using existing data and in terms of the design of future surveys.

11.6.1 Policy Implications

The substantive results presented in this thesis lead to some rather clear policy implications. Firstly, although the focus of the work has been elsewhere and has tried to focus on the environment that an individual experiences and lives within, the centrality of the individual's criminogenic propensity has been shown to be a central issue. In every model that has included this component it has played a pivotal role in the relationship with offending, and analysis has confirmed what would be expected from the developmental theory, that this individual attitude is relatively stable (though when and how it develops is outside the scope of the data, previous research indicates it is before the age of 10). This should not be ignored, and policies to address this enduring type of criminogenic propensity would clearly be the most appropriate in terms of reducing offending behaviour over the long term. This point shall be picked up again shortly.

This work has shown that at a national level both neighbourhood and family characteristics matter, but that they matter in slightly different respects. Perceptions of neighbourhood disorder were shown to have a strong influence over the offending behaviour of the adolescent only at one particular time point, not one that lasts. This is different to what Wikström (2012) might call a criminogenic setting in that the respondent is not reflecting on their offending behaviour and the setting found at that moment, but on their impression of their local area as they see it in general. Given a population of adolescents therefore, and a need to do something to address offending behaviour immediately then the response should be to implement a policy that addresses the neighbourhood character. The manner in which the neighbourhood disorder and collective efficacy interact with offending is a new finding and means from a policy perspective that whilst it may appear at first glance that measures to reduce perceptions of disorder might be the primary method of achieving this goal, it would also be possible to implement a process to also try to boost perceptions of collective efficacy, which should in turn help reduce perceptions of neighbourhood disorder.

The analysis, however, demonstrated that this relationship is transitory, and therefore any longer-term policy implications must reflect this difference. Family character, both perceived and structural, seems to have an effect on offending over the longer term. This leads to a mix of potential policy implications depending on which of these it is felt need to be addressed and the over-time causality that is being examined. The final two sets of models indicated how over the long term the perceptions of family discipline mediated against offending but when other measures of family level problems (separation and trouble with the criminal justice system) are added, this effect evaporates. Would a policy implication therefore be that it is best to keep families together and not arrest any parents? Of course not, these external factors most probably represent themselves in the young person's life in other ways and the fact that the addition of these covariates means the disappearance of the *perceptions* of family characteristics effect clearly points to this. It means there is a general family disruption problem at play. A more enactable policy implication would therefore be that given a negative family context, which may be indicated by problematic divorce or parental involvement with nefarious behaviour, then the child within the household must be supported *over the long term*, with help put in place to make a suitably normative framework to address their behaviour and, importantly, their *perceptions* of their family's behaviour. Perhaps addressing their neighbourhood would help at the time, but without processes in place to help the family situation over a longer period they are likely to fail.

Although the direction of the mediation effects of the family measures was initially counter-intuitive, meaning that harsher discipline increased offending, whilst an argument could be made about negative effects of an overly prescriptive atmosphere it must be remembered we are dealing with national data here and therefore this type of household is relatively unlikely to be dominating the coefficient of this effect. The work understood this reversal by pointing to the stability of the family environment as demonstrated in Chapter 10, extrapolating this to be a remnant of an earlier age and therefore reversing the causality with the individual's attitude. This obviously puts the cause back onto the individual criminogenic propensity and the central part that this has played in every aspect of the modelling presented here and the theoretical

expectations that it is built upon. The policy implication over a longer time period, to reduce offending behaviour from the whole life course, therefore, is along the familiar lines that highlights the importance of a nurturing and positive family environment from an early age.

One final policy implication to note was touched on very briefly in the concluding remarks to Chapter 9, and that is by looking at the effect of the control variables on the mediators, where a suitable policy initiative designed and implemented to address the perception of neighbourhood, then knowledge of the effects of the structural characteristics (such as age, gender or divorce) on those perceptions could help with the targeting of those interventions. This has not been the focus of the current work but a revised analysis placing more importance on unpicking these effects could lead to useful knowledge.

11.6.2 Methodological Implications for Existing Data

This section shall focus on methodological implications that should be worked into any analytical approach once the data collection process is complete. In the entirety of this thesis, work has progressed only on an existing dataset that can be downloaded simply by registering an online account, and there are many hundreds more such datasets available to researchers. The OCJS by the standards of these national surveys is not particularly large nor particularly complex, but by approaching each concept's measurements and then the relationships between these concepts from the bottom up and building forward one step at a time, some results have been achieved that have thus far eluded researchers. Perhaps criminology lends itself to this type of approach more so than other realms of the social sciences as it has a long history of detailed theoretical debate, often approached with numerous different methodologies in many different academic traditions, both qualitative and quantitative and from psychology through to econometrics. The sequential ordering of the analysis, however, with open confrontation of the limits of generalisability, and the simple analytical framework consisting of asking what can be measured, can these measures be structured and what are the results *with equal attention paid to each part* is surely widely applicable.

Structural Equation Modelling, SEM, very much helps in being able to achieve these goals and it is unsurprising that the methodology has enjoyed a huge uptake over the last 30 years. Armed with increasingly more sophisticated computer technology it is likely that the method will carry on being picked up by more and more researchers, especially given the increased functionality available in the recently released STATA14 with its excellent graphical user interface (notably the asymptotic distribution free estimator, bootstrapping and the implementation of Satorra–Bentler adjustments). The ability to create non-mathematical intuitive presentations of the data structure and results via path diagrams is a huge boon, which allows the method to be understood by a wide range of people with limited statistical knowledge. Its versatility once one is adept in the processes of model specification mean that it can replicate a huge variety of techniques that are familiar elsewhere: multi-level modelling, econometric models of simultaneous equations, latent growth curve modelling and beyond.

This does not mean that structural equation modelling is a panacea for all problems but that social science, psychology and economics could do well to pay more attention to measurement frameworks in their original format rather than rushing forward with subsets that may not accurately capture the original concept. The very notion of collective efficacy is a case in point: in the original definition Sampson et al. (1999) perceive collective efficacy to have two distinct dimensions, social trust and informal social control, but it is common to find work that used vastly reduced numbers of items to 'measure' this concept and the weaknesses of these measures are not adequately addressed. For example, Laurence and Heath (2005) use just the informal social control elements of the definition to formulate their 'collective efficacy' measure, and after a quick Cronbach alpha to give reliability, use it to predict changes in community cohesions and find an extremely strong effect without acknowledging that community cohesion is exactly one half of the original definition. To be clear, it is not the intention to say that any works necessarily become invalid if they work with reduced measurement, indeed this very thesis makes exactly some of these assumptions, but the awareness of the nuances of the measurement instrument and their relationships to other concepts can be quickly lost and this needs to be considered, whatever statistical technique is being applied.

Causality within SEM approaches is also often misconstrued for the simple fact that it is attractive to interpret an arrow coming out of one circle and going into another as a causal relationship. In reality this is exactly the same as a regression and considering the implication of a regression coefficient is an absolute must for all courses that attempt to teach that technique. SEM is no different. This does not mean that any particular results are dubious, far from it, but what it does mean is that as with any statistical modelling the use of language is extremely important. In this thesis there has been a deliberate attempt to try to use as neutral a causal language as possible, difficult as it may be. Where causal language and statements are made these need to be in a tightly controlled manner, which full consideration of the model necessitates. As in regression, it is possible to use causal language, but only when it is certain that there is a cause and there is an effect.

11.6.3 Wider Limitations and Implications for Data Collection and Survey Design

Given that this work has largely managed to create measures of complex risk factors using only secondary data it is tempting to conclude that there is little improvement that could be made to any data collection process, and that the methodology can iron out any issues post-hoc. This simply is not the case and a good deal of this thesis has been dedicated to trying to rectify problems that need not have occurred. Given that there are a number of well-established scales used to measures parts of what has been termed here individual attitude (Grasmick et al., 1995; Whiteside and Lynam, 2001; Jolliffe and Farrington, 2006), in an ideal world if a survey were to be conducted then to acknowledge that these are the basis for a question set, label them as such and point to the literature from which they are drawn in the documentation would no doubt ease the researcher's job. This would require someone with specific knowledge to sit on the design committee and input into the questionnaire design actively, but this shouldn't be too overbearing a responsibility. Trying to match the questions present on the OCJS to theoretical and previous definitions has been a relatively tough job and it does not appear this has been done from this data source before, this despite the survey being close to ten years old and available to researchers for quite some time.

The inclusion of a more comprehensive set of measures would also be beneficial. Jolliffe and Farrington (2010) spend a good deal of their chapter in the SAGE Handbook of Criminological Theory discussing the impact of intelligence on offending, and conclude that on balance low intelligence and low attainment seem to have a relationship with greater offending. Similarly, measures of personality types, though seen probably to be closely related to those measured here, would have been helpful. This is mainstream knowledge that does not seem to have been acknowledged and incorporated into the design of the survey and would be relatively straightforward to fix.

One abiding limitation of this work has concerned the issue of causation between family and individual risk factors, and working out which way round they should go. It has only been possible to speculate, based on results, a way to unpick this. Given the stability of the individual attitude and well researched expectation that this develops in early childhood, it would be a relatively straightforward step to design a short survey to be administered to the parent or guardian of the primary respondent asking them about their opinion of the child, perhaps specifically asking about how they were in younger childhood. This type of additional survey could also ask about parental judgement of their discipline, school achievement and would also allow better capture of household level questions such as income (awfully recorded in the OCJS) and socioeconomic status.

To try to improve the area level measures, though, would probably be to ask too much. The great lengths that the PHDCN and the PADS+ have gone through to get a measure of the characteristics of an area, using detailed observation or extremely large scale specific surveys in relatively constrained geographical spaces (at most a city) would clearly be prohibitively expensive. One possibility could be geographical co-location of surveys such as running the (larger) Citizenship Survey in the same locations as those of a future attempt at data collection related to self-report. If the sampling strategy was aligned closely enough it would be possible to get another view, perhaps even another four or five views, of the collective efficacy of an area (for example) but would this be an improvement on an individual's self-perception of an area? Wikström et al. (2012) retrieved 13 responses per output area (the smallest unit of Office for National Statistics

Geographical classifications, smaller than Lower Super Output Area and an average of around 120 households) to be able to measure their characteristics of a neighbourhood, so would four or five actually help the measurement here? Is it not the opinion of the individual that matters most? Perhaps a combination of the two, but to replace the individual's own assessment with that of a handful of completely unrelated people would not be adequate. And this operates under the assumption that that level of joined-up survey administration is even possible! Perhaps in relation to area measurement of non-administrative factors there *are* only two possibilities, what people perceive in and of themselves *or* taking a brute force (and expensive) more laborious approach.

11.7 Future Work

This thesis gives many departure points for further work, ranging from the immediate to the distantly possible. On the immediate side it should be possible to address some of the aspects which have thus far been defined outside the scope of this investigation. School and peer effects have not been included in this work at all, yet their role in the development of delinquent behaviour has been highlighted by numerous studies (see Farrington and Welsh, 2007, or Blackburn, 1993, p173). Peer influences in particular are a key part of Moffitt's developmental taxonomy, and descriptive work on the OCJS highlighted increased offending activity amongst those in 'delinquent youth gangs'. The aim of this work was not to look into these aspects, and with good reason as investigating area has taken a good amount of space, but this is certainly something that could be looked at in future analysis and using exactly this approach.

Perhaps an even simpler extension would be to change the dependent completely. In this current work the focus has been on self-report offending and only self-report offending. This has been for reason of focus but it would be possible to change this completely to a measure of other delinquent behaviour. Broader anti-social behaviour can be measured from the OCJS dataset with a specific set of questions asked and the survey's stated aim is to look at drug usage, and hence how these risk factors relate to these could be included in an analysis. McVie and Norris (2006a) found that the

neighbourhood had different effects on drug use compared to self-reported delinquent acts, and this analysis could easily be achieved using this framework.

Another interesting avenue for further exploration is the link between fear of crime and offending behaviour. This was touched upon very briefly in Chapter 7 and then moved away from as it was felt there was too little theory to be able to include this in the analysis, given that SEM is a hypothesis-testing method rather than a theory-generating mechanism. Questions on self-report offending are typically not included on datasets that aim to investigate fear of crime or victimisation, as the questions needed to get a reliable measure are time consuming to design and administer. As these are included in the OCJS, with the right theorising at the outset this question could be addressed from this source.

An attractive further piece of work is to try to replicate these results from an alternative dataset. Now that these relationships have been established and the methodology to tap measurement of the psychological risk factors demonstrated, it would be possible to conduct this analysis on other datasets with an appropriate range of questions. One source that immediately springs to mind is the International Self-Report Delinquency Study (ISRD). Launched in 1992 by the Ministry of Justice in the Netherlands, the aims of this study are to measure the prevalence of different types of delinquent behaviour in industrialised countries, examine sources of cross-national variation and the development of the self-report methodology (Junger-Tas and Terlouw, 1994). The first wave involving 12- to 15-year-olds in 13 countries was conducted between 1992 and 1993; the second expanded to 31 countries, with the same aged participants, was conducted between 2005 and 2007 and was made publically available in January 2015. The third wave of the study is currently in the field, covering 35 countries, and focuses on the empirical integration of Situational Action Theory, Institutional Anomie Theory and Procedural Justice Theory. Further areas of interest are theories of social control, criminal opportunity and social disorganisation. Together these could provide an extremely rich dataset to test the findings of this work and more (see the project website for more details). Particularly, SEM is capable of analytically delving into whether

empirical specifications are equivalent across countries and the ISRD3 would be an ideal test bed for such work.

Appendix A

Table 11.1 Model results for subdomains of the index of multiple deprivation, from section 9.3

	Effects on mediators				Direct effect on offend	Indirect effects through mediators				Total effect on offend	Proportion of mediation			
	ND (total)	CE (total)	FAR	FAD		ND (total)	CE (total)	FAR	FAD		ND (total)	CE (total)	FAR	FAD
Pathway in Figure 9.1	a+b*e	b	c	d	g	(a+b*e) *h	b(i+e*h)	c*j	d*k	Sum all paths				
Model 3 with Crime IMD														
Individual Attitude	0.19***	0.25***	0.52***	0.56***	0.7***	0.07**	n/s	n/s	n/s	0.59***	11%**	n/s ind	n/s ind	n/s ind
Female (ref male)	n/s	0.09	n/s	-0.12*	-0.31***	n/s	n/s	n/s	n/s	-0.32***	n/s ind	n/s ind	n/s ind	n/s ind
Age	0.06***	n/s	0.08***	0.16***	0.12***	0.05**	n/s	n/s	n/s	0.13***	38%**	n/s ind	n/s ind	n/s ind
Live together (reference)														
Separate but see	n/s	0.13	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Separate don't see	0.18***	0.3**	0.28*	n/s	n/s	0.15*	n/s	n/s	n/s	0.23	68%*	n/s ind	n/s ind	n/s ind
Guardians - No trouble (reference)														
Guardians - trouble with police	0.17**	n/s	0.44**	0.29**	0.33*	0.14*	n/s	n/s	n/s	0.35**	40%*	n/s ind	n/s ind	n/s ind
Guardians - Prison	0.34***	n/s	n/s	0.35	0.44	0.28*	n/s	n/s	n/s	0.58*	48%*	n/s ind	n/s ind	n/s ind
Employers or high professionals	n/s	n/s	n/s	-0.37**	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Lower managerial and professional	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Intermediate (reference)														
Small Organisations and own account worker	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Lower supervisory	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Semi-routine	0.14	n/s	n/s	n/s	n/s	0.11	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Routine	n/s	0.35**	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
IDACI decile	0.03***	0.06***	n/s	n/s	n/s	0.02**	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot

	Effects on mediators				Direct effect on offend	Indirect effects through mediators				Total effect on offend	Proportion of mediation			
	ND (total)	CE (total)	FAR	FAD		ND (total)	CE (total)	FAR	FAD		ND (total)	CE (total)	FAR	FAD
Pathway in Figure 9.1	a+b*e	b	c	d	g	(a+b*e)*h	b(i+e*h)	c*j	d*k	Sum all paths				
Neighbourhood condition	0.29***	0.21***	n/s	0.11**	n/s	0.1**	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Urban (reference)														
Town or urban fringe	n/s	-0.18*	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Rural	-0.19**	-0.43***	n/s	-0.17	-0.27	-0.16*	n/s	n/s	n/s	-0.28*	58%*	n/s ind	n/s ind	n/s ind
Model 3 with Income IMD														
Individual Attitude	0.2***	0.26***	0.52***	0.56***	0.69***	0.07**	n/s	n/s	n/s	0.59***	12%**	n/s ind	n/s ind	n/s ind
Female (ref male)	n/s	0.11*	n/s	-0.11*	-0.31***	n/s	n/s	n/s	n/s	-0.32***	n/s ind	n/s ind	n/s ind	n/s ind
Age	0.06***	0.03	0.08***	0.16***	0.11***	0.05**	n/s	n/s	n/s	0.13***	40%**	n/s ind	n/s ind	n/s ind
Live together (reference)														
Separate but see	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Separate don't see	0.15**	0.24*	0.28*	n/s	n/s	0.13*	n/s	n/s	n/s	0.25	52%*	n/s ind	n/s ind	n/s ind
Guardians - No trouble (reference)														
Guardians - trouble with police	0.16**	n/s	0.44**	0.29**	0.34*	0.14*	n/s	n/s	n/s	0.35**	38%*	n/s ind	n/s ind	n/s ind
Guardians - Prison	0.29**	n/s	n/s	0.34	0.46	0.25*	n/s	n/s	n/s	0.6**	42%*	n/s ind	n/s ind	n/s ind
Employers or high professionals	n/s	n/s	n/s	-0.35**	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Lower managerial and professional	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Intermediate (reference)														
Small Organisations and own account worker	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Lower supervisory	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Semi-routine	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot

	Effects on mediators				Direct effect on offend	Indirect effects through mediators				Total effect on offend	Proportion of mediation			
	ND (total)	CE (total)	FAR	FAD		ND (total)	CE (total)	FAR	FAD		ND (total)	CE (total)	FAR	FAD
Pathway in Figure 9.1	a+b*e	b	c	d	g	(a+b*e)*h	b(i+e*h)	c*j	d*k	Sum all paths				
Routine	n/s	0.28*	n/s	n/s	0.31	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Crime decile	0.04***	0.08***	n/s	n/s	n/s	0.03**	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Neighbourhood condition	0.23***	0.15***	n/s	0.08*	n/s	0.08**	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Urban (reference)														
Town or urban fringe	n/s	-0.22*	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Rural	-0.17**	-0.44***	n/s	-0.16	-0.34*	-0.15*	n/s	n/s	n/s	-0.34**	43%*	n/s ind	n/s ind	n/s ind
Model 3 with IDACI IMD														
Individual Attitude	0.2***	0.26***	0.52***	0.56***	0.7***	0.07**	n/s	n/s	n/s	0.59***	12%**	n/s ind	n/s ind	n/s ind
Female (ref male)	n/s	0.12*	n/s	-0.11*	-0.31***	n/s	n/s	n/s	n/s	-0.32***	n/s ind	n/s ind	n/s ind	n/s ind
Age	0.06***	0.03	0.08***	0.16***	0.12***	0.05**	n/s	n/s	n/s	0.13***	39%**	n/s ind	n/s ind	n/s ind
Live together (reference)														
Separate but see	n/s	0.12	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Separate don't see	0.14**	0.23*	0.28*	n/s	n/s	0.12*	n/s	n/s	n/s	0.24	51%*	n/s ind	n/s ind	n/s ind
Guardians - No trouble (reference)														
Guardians - trouble with police	0.15**	n/s	0.44***	0.28**	0.34*	0.13*	n/s	n/s	n/s	0.35**	37%*	n/s ind	n/s ind	n/s ind
Guardians - Prison	0.27**	n/s	n/s	0.33	0.47	0.24*	n/s	n/s	n/s	0.59**	40%*	n/s ind	n/s ind	n/s ind
Employers or high professionals	n/s	n/s	n/s	-0.35**	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Lower managerial and professional	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Intermediate (reference)														
Small Organisations and own account worker	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot

	Effects on mediators				Direct effect on offend	Indirect effects through mediators				Total effect on offend	Proportion of mediation			
	ND (total)	**CE (total)**	**FAR**	**FAD**		**ND (total)**	**CE (total)**	**FAR**	**FAD**		**ND (total)**	**CE (total)**	**FAR**	**FAD**
Pathway in Figure 9.1	a+b*e	b	c	d	g	(a+b*e) *h	b(i+e*h)	c*j	d*k	Sum all paths				
Lower supervisory	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Semi-routine	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Routine	n/s	0.28*	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Income decile	0.04***	0.09***	n/s	0.02*	n/s	0.03**	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Neighbourhood condition	0.22***	0.13***	n/s	0.07	n/s	0.08**	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Urban (reference)														
Town or urban fringe	n/s	-0.22*	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s	n/s tot	n/s tot	n/s tot	n/s tot
Rural	-0.16*	-0.42***	n/s	n/s	-0.33*	-0.14*	n/s	n/s	n/s	-0.33*	42%*	n/s ind	n/s ind	n/s ind

*** = 99.9% significant, ** = 99% significant, * = 95% significant, no star = significant at 90% level. Lower significance otherwise

Bibliography

Akers, R. L. (1991). Self-control as a General Theory of Crime. *Journal of Quantitative Criminology*, *7*(2), 201–211.

Akers, R. L. (1999). *Criminological Theories: Introduction and Evaluation*. Chicago: Routledge.

Aldridge, J., Medina-Ariz, J., and Ralphs, R. (2012). Counting Gangs: Conceptual and Validity Problems with the Eurogang Definition. In F.-A. Esbensen and C. L. Maxson (Eds.), *Youth Gangs in International Perspective* (pp. 35–51). New York: Springer.

Alihan, M. (1938). *Social Ecology: A Critical Analysis*. New York: Columbia University Press.

Aluja, A., and Blanch, A. (2007). Comparison of Impulsiveness, Venturesomeness and Empathy Structure in English and Spanish Samples: Analysis of Different Structural Equation Models. *Personality and Individual Differences*, *43*(8), 2294–2305.

Amato, P. R., and Previti, D. (2003). People's Reasons for Divorcing: Gender, Social Class, the Life Course, and Adjustment. *Journal of Family Issues*, *24*(5), 602–626.

Angrist, J. D., and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Arah, O. A. (2008). The Role of Causal Reasoning in Understanding Simpson's Paradox, Lord's Paradox, and the Suppression Effect Covariate Selection in the Analysis of Observational Studies. *Emerging Themes in Epidemiology*, v.*5 p*5.

Arbuckle, J. L. (2006). *Amos (Version 7.0) User Guide*. Chicago: SPSS.

Armstrong, T. A., Katz, C. M., and Schnebly, S. M. (2010). The Relationship between Citizen Perceptions of Collective Efficacy and Neighborhood Violent Crime. *Crime & Delinquency*. 61(1), 121-142

Arneklev, B. J., Cochran, J. K., and Gainey, R. R. (1998). Testing Gottfredson and Hirschi's "Low Self-Control" Stability Hypothesis: An Exploratory Study.

American Journal of Criminal Justice, *23*(1), 107–127.

Arneklev, B. J., Elis, L., and Medlicott, S. (2006). Testing the General Theory of Crime: Comparing the Effects of "Imprudent Behavior" and an Attitudinal Indicator of "Low Self-Control". *Western Criminology Review*, 7(3), 41–55

Asparouhov, T., and Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397–438.

Asparouhov, T., and Muthén, B. (2010). Weighted Least Squares Estimation with Missing Data. *Mplus Technical Appendix*, 1–10.

Aston, E. V. (2015). Are Risk Factors for Drug Use and Offending Similar During The Teenage Years? *International Journal of Drug Policy*, *26*(4), 396–403.

Bader, M. D., and Ailshire, J. A. (2013). Creating Measures of Theoretically Relevant Neighborhood Attributes at Multiple Spatial Scales. *Sociological Methodology*, 44(1), 322–368.

Baldwin, J., and Anthony, B. (1976). *The Urban Criminal: a Study in Sheffield*. London: Tavistock Publications.

Bandura, A. (2000). Exercise of Human Agency through Collective Efficacy. *Current Directions in Psychological Science*, *9*(3), 75–78.

Bandura, A. (2006). Guide for Constructing Self-Efficacy Scales. in Pajares, F. and Urdan, T. (ed) *Self-Efficacy Beliefs of Adolescents*, p307 337. Information Age Publishing Inc, US

Baron, R. M., and Kenny, D. A. (1986). The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173.

Barrett, P. (2007). Structural Equation Modelling: Adjudging Model Fit. *Personality and Individual Differences*, *42*(5), 815–824.

Bartholomew, D. J., Knott, M., and Moustaki, I. (2011). *Latent Variable Models and Factor Analysis: A Unified Approach* (3rd Edition). Chichester, West Sussex: Wiley-Blackwell.

Bartholomew, D. J., Steele, F., Galbraith, J., and Moustaki, I. (2008). *Analysis of Multivariate Social Science Data*. Boca Raton: Chapman and Hall/CRC.

Becker, J., and Roe, S. (2005). Drug Use among Vulnerable Groups of Young People: Findings from the 2003 Crime and Justice Survey. London: Home Office.

Bellair, P. E. (2000). Informal Surveillance and Street Crime: A Complex Relationship. *Criminology*, *38*(1), 137–170.

Bender, R., and Grouven, U. (1998). Using Binary Logistic Regression Models For Ordinal Data with Non-Proportional Odds. *Journal of Clinical Epidemiology*, *51*(10), 809–816.

Bentler, P. M. (1980). Multivariate Analysis with Latent Variables: Causal Modeling. *Annual Review of Psychology*, *31*(1), 419–456.

Bentler, P. M. (1990). Comparative Fit Indexes in Structural Models. *Psychological Bulletin*, *107*(2), 238–246.

Bentler, P. M. (1995). *EQS structural equations program manual*. Multivariate Software.

Bentler, P. M., and Chou, C.-P. (1987). Practical Issues in Structural Modeling. *Sociological Methods & Research*, *16*(1), 78–117.

Berry, J. W., Phinney, J. S., Sam, D. L., and Vedder, P. (2006). Immigrant Youth: Acculturation, Identity, and Adaptation. *Applied Psychology*, *55*(3), 303–332.

Beyers, J. M., Loeber, R., Wikström, P. O., and Stouthamer-Loeber, M. (2001). What Predicts Adolescent Violence in Better-Off Neighborhoods? *Journal of Abnormal Child Psychology*, *29*(5), 369–381.

Birks, D., Townsley, M., and Stewart, A. (2012). Generative Explanations of Crime: Using Simulation to Test Criminological Theory. *Criminology*, *50*(1), 221–254.

Bjerk, D. (2010). Thieves, Thugs, and Neighborhood Poverty. *Journal of Urban Economics*, *68*(3), 231–246.

Björkqvist, K., Österman, K., and Kaukiainen, A. (2000). Social Intelligence – Empathy = Aggression? *Aggression and Violent Behavior*, *5*(2), 191–200.

Blackburn, R. (1993). *The Psychology of Criminal Conduct: Theory, Research And Practice*.

Chichester, England; New York: John Wiley.

Blanc, M. L. (2006). Self-Control and Social Control of Deviant Behaviour in Context: Development and Interactions Along the Life Course. In P.-O. H. Wikström and R. J. Sampson (Eds.), *The Explanation of Crime, Context, Mechanisms and Development* (pp. 195–242). Cambridge: Cambridge University Press.

Blau, J. R., and Blau, P. M. (1982). The Cost of Inequality: Metropolitan Structure and Violent Crime. *American Sociological Review*, *47*(1), 114–129.

Blumstein, A., and National Research Council (U. S.) Panel on Research on Criminal Careers. (1986). *Criminal Careers and "Career Criminals"*. Washington, DC: National Academy Press.

Youth Justice Board (2014). *Youth Justice Statistics 2012/13*. Youth Justice Board / Ministry of Justice Statistics bulletin available online https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/278549/youth-justice-stats-2013.pdf

Bock, R. D., and Bargmann, R. E. (1966). Analysis of Covariance Structures. *Psychometrika*, *31*(4), 507–534.

Bollen, K. A. (1987). Total, Direct, and Indirect Effects in Structural Equation Models. *Sociological Methodology*, *17*(1), 37–69.

Bonin, E.-M., Stevens, M., Beecham, J., Byford, S., and Parsonage, M. (2011). Costs and Longer-Term Savings of Parenting Programmes for the Prevention of Persistent Conduct Disorder: A Modelling Study. *BMC Public Health*, *11*(1), 803.

Borsboom, D. (2005). *Measuring the Mind: Conceptual Issues in Contemporary Psychometrics*. Cambridge: Cambridge University Press.

Bottoms, A. E. (2007). Place, Space, Crime and Disorder. In R. Morgan, M. Maguire, and R. Reiner (Eds.), *The Oxford Handbook of Criminology* (4th ed., pp. 528 – 574). Oxford: Oxford University Press.

Bottoms, A. E., and Wiles, P. (1986). Housing Tenure and Residential Community Crime Careers in Britain. In Reiss, A. J. and Tonry, M. (Eds) *Crime and Justice*, *Vol*

8: Communities and Crime p101–162. Chicago, University of Chicago Press

Bottoms, A. E., and Wiles, P. (2002). Environmental Criminology. In M. Maguire, R. Morgan, and R. Reiner (Eds.), *The Oxford Handbook of Criminology* (3rd ed., pp. 620–656). Oxford: Oxford University Press.

Bradford, B., and Myhill, A. (2014). Triggers of Change to Public Confidence in the Police and Criminal Justice System: Findings from the Crime Survey for England and Wales Panel Experiment. *Criminology and Criminal Justice 15*(1)

Braithwaite, J. (1997). Charles Tittle's Control Balance and Criminological Theory. *Theoretical Criminology*, *1*(1), 77–97.

Brantingham, P. L., and Brantingham, P. J. (1993a). Environment, Routine, and Situation: Toward a Pattern Theory of Crime. In R. V. G. Clarke and M. Felson (Eds.), *Routine Activity and Rational Choice* (p. 259–294). Piscataway, NJ: Transaction Publishers.

Brantingham, P. L., and Brantingham, P. J. (1993b). Nodes, Paths and Edges: Considerations on the Complexity of Crime and the Physical Environment. *Journal of Environmental Psychology*, *13*(1), 3–28.

Breathnach, S. (2002). *Emile Durkheim on Crime and Punishment: An Exegesis*. Boca Raton: Universal Publishers.

Brezina, T. (1998). Adolescent Maltreatment and Delinquency: The Question of Intervening Processes. *Journal of Research in Crime and Delinquency*, *35*(1), 71–99.

Browning, C. R. (2002). The Span of Collective Efficacy: Extending Social Disorganization Theory to Partner Violence. *Journal of Marriage and Family*, *64*(4), 833–850.

Browning, C. R., Dietz, R. D., and Feinberg, S. L. (2004). The Paradox of Social Organization: Networks, Collective Efficacy, and Violent Crime In Urban Neighborhoods. *Social Forces*, *83*(2), 503–534.

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: Guilford Press.

Brunton-Smith, I. (2008). *Local Areas and Fear of Criminal Victimisation: Applying Multilevel Models to the British Crime Survey*. University of Surrey Thesis.

Brunton-Smith, I., and Jackson, J. (2012). *Urban Fear and Its Roots in Place*. in Ceccato, V. (ed.) *Urban Fabric of Crime and Fear*. London: Springer, pp. 55-82.

Brunton-Smith, I., Jackson, J., and Sutherland, A. (2014). Bridging Structure and Perception on the Neighbourhood Ecology of Beliefs and Worries about Violent Crime. *British Journal of Criminology*, 54, 503–526.

Brunton-Smith, I., and Sturgis, P. (2011). Do Neighbourhoods Generate Fear of Crime? An Empirical Test using the British Crime Survey. *Criminology*, *49*(2), 331–369.

Brunton-Smith, I., Sutherland, A., and Jackson, J. (2013). The Role of Neighbourhoods in Shaping Crime and Perceptions of Crime. In D. Manley, M. van Ham, N. Bailey, L. Simpson, and D. Maclennan (Eds.), *Neighbourhood Effects or Neighbourhood Based Problems?* (pp. 67–87). Springer Netherlands

Bryman, A. (2004). *Social Research Methods*. Oxford: Oxford University Press.

Buckley, W. (1967). *Sociology and Modern Systems Theory*. Oxford, England: Prentice-Hall.

Budd, T., Sharp, C., Mayhew, P., and Britain, G. (2005). *Offending in England and Wales: First Results from the 2003 Crime and Justice Survey*. London: Home Office.

Buonanno, P., Montolio, D., and Vanin, P. (2009). Does Social Capital Reduce Crime? *Journal of Law and Economics*, *52*(1), 145–170.

Burgess, E. W. (1925). The Growth of the City. In R. E. Park, E. W. Burgess, and R. D. McKenzie (Eds.), *The City*. Chicago: University of Chicago Press.

Bursik, R. J. (1988). Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects. *Criminology*, *26*(4), 519–552.

Bursik, R. J., and Grasmick, H. G. (1995). Neighborhood-Based Networks and the Control of Crime and Delinquency. In H. D. Barlow (Ed.), *Crime And Public Policy: Putting Theory To Work* (pp. 107–130). Boulder, CO: Westview Press.

Bynner, J., Butler, N., Ferri, E., Shepherd, P., and Smith, K. (2000). The Design and

Conduct of the 1999–2000 Surveys of the National Child Development Study and the 1970 British Cohort Study. *Centre for Longitudinal Studies, Cohort Studies, Working Paper*, (1).

Byrne, B. M. (1994). Burnout: Testing for the Validity, Replication, and Invariance of Causal Structure across Elementary, Intermediate, and Secondary Teachers. *American Educational Research Journal*, *31*(3), 645–673.

Byrne, B. M. (2012). *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming*. New York, Routledge.

Byrne, B. M., Stewart, S. M., and Lee, P. W. H. (2004). Validating the Beck Depression Inventory-II for Hong Kong Community Adolescents. *International Journal of Testing*, *4*(3), 199–216.

Canter, R. J. (1982). Sex Differences in Self-Report Delinquency. *Criminology*, *20*(3-4), 373–394.

Capaldi, D. M., and Patterson, G. R. (1991). Relation of Parental Transitions to Boys' Adjustment Problems: I.A. Linear Hypothesis II. Mothers at Risk for Transitions and Unskilled Parenting. *Developmental Psychology*, *27*(3), 489–504.

Capaldi, D. M., and Patterson, G. R. (1996). Can Violent Offenders be Distinguished from Frequent Offenders: Prediction from Childhood to Adolescence. *Journal of Research in Crime and Delinquency*, *33*(2), 206–231.

Carroll, J. M., Rosson, M. B., and Zhou, J. (2005). Collective Efficacy as a Measure of Community. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1–10).

Caspi, A. (2000). The Child is Father of the Man: Personality Continuities from Childhood to Adulthood. *Journal of Personality and Social Psychology*, *78*(1), 158–172.

Lombroso, C. (2006). *Criminal Man*. Durham, NC: Duke University Press.

Chen, F. F., Sousa, K. H., and West, S. G. (2005). Teacher's Corner: Testing Measurement Invariance of Second-Order Factor Models. *Structural Equation Modeling*, *12*(3),

471–492.

Chen, L., Gillenson, M. L., and Sherrell, D. L. (2002). Enticing Online Consumers: an Extended Technology Acceptance Perspective. *Information & Management*, *39*(8), 705–719.

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255.

Cleek, M. G., and Pearson, T. A. (1985). Perceived Causes of Divorce: An Analysis of Interrelationships. *Journal of Marriage and the Family*, 47(1) 179–183.

Cliff, N. (1983). Some Cautions Concerning the Application of Causal Modeling Methods. *Multivariate Behavioral Research*, *18*(1), 115–126.

Clifford R. Shaw, and McKay, H. D. (1942). *Juvenile Delinquency and Urban Areas : A Study of Rates of Delinquency in Relation to Differential Characteristics of Local Communities in American Cities.* (Revised edition).

Cloward, R. A., and Ohlin, L. E. (2013). *Delinquency and Opportunity: A Study of Delinquent Gangs*. Toronto, Canada, Routledge.

Cochran, J. K., Wood, P. B., Sellers, C. S., Wilkerson, W., and Chamlin, M. B. (1998). Academic Dishonesty and Low Self-Control: An Empirical Test of a General Theory of Crime. *Deviant Behavior*, *19*(3), 227–255.

Cohen, A. K. (1955). *Delinquent Boys: the Culture of the Gang*. London: Routledge & Kegan Paul.

Cohen, D. A., Farley, T. A., and Mason, K. (2003). Why is Poverty Unhealthy? Social and Physical Mediators. *Social Science & Medicine*, *57*(9), 1631–1641.

Cohen, D. A., Finch, B. K., Bower, A., and Sastry, N. (2006). Collective Efficacy and Obesity: the Potential Influence of Social Factors On Health. *Social Science & Medicine*, *62*(3), 769–778.

Cohen, D. A., Inagami, S., and Finch, B. (2008). The Built Environment and Collective Efficacy. *Health & Place*, *14*(2), 198–208.

Cohen, L. E., and Felson, M. (1979). Social Change and Crime Rate Trends: A Routine Activity Approach. *American Sociological Review*, 44(4) 588–608.

Cole, D. A., and Maxwell, S. E. (2003). Testing Mediational Models with Longitudinal Data: Questions and Tips in the Use of Structural Equation Modelling. *Journal of Abnormal Psychology*, *112*(4), 558.

Conger, A. J. (1974). A Revised Definition for Suppressor Variables: a Guide to Their Identification and Interpretation. *Educational and Psychological Measurement*, *34*(1), 35–46.

Conway, A. R. A., Macnamara, B. N., and Engel de Abreu, P. M. J. (2013). Working Memory and Intelligence: An Overview. In T. P. Alloway and R. G. Alloway (Eds.), *Working Memory: The Connected Intelligence*. New York; London: Psychology Press.

Cullen, F. T., and Agnew, R. (2006). *Criminological Theory: Past To Present: Essential Readings* (3rd ed.). Los Angeles: Roxbury Park.

Dale, A., Arber, S., and Procter, M. (1988). *Doing Secondary Analysis*. London, Allen & Unwin,.

Dale, A., Arber, S., and Procter, M. (2003). A Sociological Perspective on Secondary Analysis. In C. Seale (Ed.), *Social Research Methods: A Reader*. London; New York: Routledge.

Davidson, R. N. (1981). *Crime and Environment*. London: Croom Helm.

DeLisi, M., and Hochstetler, A. (2002). An Exploratory Assessment of Tittle's Control Balance Theory: Results from the National Youth Survey. *The Justice Professional*, *15*(3), 261–272.

Delisi, M., Hochstetler, A., and Murphy, D. S. (2003). Self-Control Behind Bars: A Validation Study of the Grasmick et al. Scale. *Justice Quarterly*, *20*(2), 241–263.

DeLisi, M., and Vaughn, M. G. (2008). The Gottfredson–Hirschi Critiques Revisited Reconciling Self-Control Theory, Criminal Careers, and Career Criminals. *International Journal of Offender Therapy and Comparative Criminology*, *52*(5), 520–

537.

Department of the Environment, Transport and the Regions. (2000). *Indices of Deprivation 2000*. Government report available online at http://webarchive.nationalarchives.gov.uk/+/http:/www.communities.gov.uk/documents/citiesandregions/pdf/131306.pdf

Dowden, C., and Andrews, D. A. (2004). The Importance of Staff Practice in Delivering Effective Correctional Treatment: A Meta-Analytic Review of Core Correctional Practice. *International Journal of Offender Therapy and Comparative Criminology*, *48*(2), 203–214.

Duckitt, J., Wagner, C., du Plessis, I., and Birum, I. (2002). The Psychological Bases of Ideology And Prejudice: Testing A Dual Process Model. *Journal of Personality and Social Psychology*, *83*(1), 75–93.

Eck, J., Chainey, S., Cameron, J., and Wilson, R. (2005). *Mapping Crime: Understanding Hotspots.* Washington, DC: National Institute of Justice.

Edens, J. F., and Ruiz, M. A. (2006). On the Validity of Validity Scales: The Importance of Defensive Responding in the Prediction of Institutional Misconduct. *Psychological Assessment*, *18*(2), 220.

Elliott, D. S., and Ageton, S. S. (1980). Reconciling Race and Class Differences in Self-Reported and Official Estimates of Delinquency. *American Sociological Review*, *45*(1), 95.

Elliott, D. S., Huizinga, D., and Menard, S. (1989). *Multiple Problem Youth: Delinquency, Substance Use, and Mental Health Problems*. New York, Springer-Verlag.

Ellis, L., Beaver, K. M., and Wright, J. (2009). *Handbook of Crime Correlates*. Calfornia, Elsevier Academic Press.

Enzmann, D., Marshall, I. H., Killias, M., Junger-Tas, J., Steketee, M., and Gruszczynska, B. (2010). Self-Reported Youth Delinquency in Europe and Beyond: First Results of the Second International Self-Report Delinquency Study in the Context of Police and Victimization Data. *European Journal of Criminology*, *7*(2), 159–183.

Evans, T. D., Cullen, F. T., Burton, V. S., Dunaway, R. G., and Benson, M. L. (1997). The Social Consequences of Self-Control: Testing the General Theory of Crime. *Criminology*, *35*(3), 475–504.

Eysenck, S. B., Pearson, P. R., Easting, G., and Allsopp, J. F. (1985). Age Norms for Impulsiveness, Venturesomeness and Empathy in Adults. *Personality and Individual Differences*, *6*(5), 613–619.

Farrall, S., Bannister, J., Ditton, J., and Gilchrist, E. (1997). Questioning the Measurement of the "Fear of Crime": Findings from a Major Methodological Study. *British Journal of Criminology*, *37*(4), 658–679.

Farrington, D. P. (1986). Age and Crime. *Crime and Justice*, vol 7 189–250.

Farrington, D. P. (1992). Criminal Career Research in the United Kingdom. *The British Journal of Criminology*, 32(4) 521–536.

Farrington, D. P. (2003). What has been Learned from Self-Reports about Criminal Careers and the Causes of Offending. London: Home Office.

Farrington, D. P. (Ed.). (2005a). *Integrated Developmental & Life-Course Theories of Offending*. New Brunswick, NJ: Transaction Publishers.

Farrington, D. P. (Ed.). (2005b). The Integrated Cognitive Antisocial Potential (ICAP) Theory. In *Integrated Developmental & Life-Course Theories of Offending* (pp. 73–92). New Brunswick, NJ: Transaction Publishers.

Farrington, D. P. (2007). Childhood Rick Factors and Risk-focused Prevention. In R. Morgan, M. Maguire, and R. Reiner (Eds.), *The Oxford Handbook of Criminology* (4th ed., pp. 602 – 640). Oxford: Oxford University Press.

Farrington, D. P. (2011). *Integrated Developmental and Life-Course Theories of Offending*. Piscataway, NJ: Transaction Publishers.

Farrington, D. P., Coid, J.W., Harnett, L. M., Jolliffe, D., Soteriou, N., Turner, R.E. and West, D.J. (2006). *Criminal Careers up to Age 50 and Life Success up to Age 48: New Findings from the Cambridge Study in Delinquent Development*. Home Office Research, Development and Statistics Directorate.

Farrington, D. P., Snyder, H. N., and Finnegan, T. A. (1988). Specialization in Juvenile Court Careers. *Criminology*, *26*(3), 461–488.

Farrington, D. P., Ttofi, M. M., and Coid, J. W. (2009). Development of Adolescence-Limited, Late-Onset, and Persistent Offenders from age 8 to age 48. *Aggressive Behavior*, *35*(2), 150–163.

Farrington, D. P., and Welsh, B. C. (2007). *Saving Children from a Life of Crime: Early Risk Factors and Effective Interventions*. Oxford; New York: Oxford University Press.

Farrington, D. P., and West, D. (1987). *Cambridge Study in Delinquent Development [Great Britain], 1961-1981*. Inter-university Consortium for Political and Social Research.

Farrington, D. P., and West, D. J. (1990). The Cambridge Study in Delinquent Development: A Long-Term Follow-Up of 411 London Males. In H.-J. Kerner and G. Kaiser (Eds.), *Kriminalität* (pp. 115–138). Berlin; Heidelberg: Springer.

Faulkner, D. (2003). Taking Citizenship Seriously: Social Capital and Criminal Justice in a Changing World. *Criminal Justice*, *3*(3), 287–315.

Felson, M. (2014). Book Review of Breaking Rules: the Social and Situational Dynamics of Young People's Urban Crime. *Journal of Criminal Justice Education*, *25*(2), 254–258.

Fergusson, D. M., Horwood, L. J., and Lynskey, M. T. (1992). Family Change, Parental Discord and Early Offending. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *33*(6), 1059–1075.

Fergusson, D. M., Horwood, L. J., and Lynskey, M. T. (1993a). Ethnicity, Social Background and Young Offending: A 14-Year Longitudinal Study. *Australian & New Zealand Journal of Criminology*, *26*(2), 155–170.

Fergusson, D. M., Horwood, L. J., and Lynskey, M. (1994). The Childhoods of Multiple Problem Adolescents: A 15-year Longitudinal Study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *35*(6), 1123–1140.

Findlay, L. C., Girardi, A., and Coplan, R. J. (2006). Links Between Empathy, Social

Behavior, and Social Understanding in Early Childhood. *Early Childhood Research Quarterly*, *21*(3), 347–359.

Fisher, G., and Ross, S. (2006). Beggarman or Thief: Methodological Issues in Offender Specialisation Research. *Australian & New Zealand Journal of Criminology*, *39*(2), 151–170.

Fitness, J., and Curtis, M. (2005). Emotional Intelligence and the Trait Meta-Mood Scale: Relationships with Empathy, Attributional Complexity, Self-Control, and Responses to Interpersonal Conflict. *E-jounral of Applied Pyschology: Social Section 1*(1), p50 - 62.

Fox, J. (2006). Teacher's Corner: Structural Equation Modelling with the SEM Package in R. *Structural Equation Modeling*, *13*(3), 465–486.

Francis, B., Soothill, K., and Dittrich, R. (2001). A New Approach for Ranking "Serious" Offences: The Use of Paired-Comparisons Methodology. *British Journal of Criminology*, *41*(4), 726–737.

Friedman, L., and Wall, M. (2005). Graphical Views of Suppression and Multicollinearity in Multiple Linear Regression. *The American Statistician*, *59*(2), 127–136.

Gelman, A. (2007). Struggles with Survey Weighting and Regression Modeling. *Statistical Science*, *22*(2), 153–164.

Office of National Statistics Geography. (2011, December 8). Super Output Area (SOA) [Text]. Retrieved June 27, 2015, from http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/super-output-areas--soas-/index.html

Gibbons, D. C. (1979). *The Criminological Enterprise: Theories and Perspectives*. Englewood Cliffs, NJ: Prentice Hall College Div.

Gibbons, S., Green, A., Gregg, P., and Machin, S. (2005). *Is Britain Pulling Apart?: Area Disparities in Employment, Education and Crime*. CMPO Working Paper Series No. 05/120.

Gibson, C. L. (2012). An Investigation of Neighborhood Disadvantage, Low Self-Control, and Violent Victimization among Youth. *Youth Violence and Juvenile Justice*, *10*(1), 41–63.

Glueck, S., and Glueck, E. T. (1951). *Unraveling Juvenile Delinquency.* Cambridge, MA: Harvard Universoty Press.

Gottfredson, M., and Hirschi, T. (1990). *A General Theory of Crime*. Stanford, CA: Stanford University Press.

Grace, J. B., and Bollen, K. A. (2005). Interpreting the Results from Multiple Regression and Structural Equation Models. *Bulletin of the Ecological Society of America*, *86*(4), 283–295.

Grasmick, H. G., Tittle, C. R., Bursik, R. J., and Arneklev, B. J. (1993). Testing the Core Empirical Implications of Gottfredson and Hirschi's General Theory Of Crime. *Journal of Research in Crime and Delinquency*, *30*(1), 5–29.

Grossman, J. R., Keating, A. D., and Reiff, J. L. (Eds.). (2004). *The Encyclopedia of Chicago* (1st edition). Chicago: University Of Chicago Press.

Grover, C. (2010). *Crime and Inequality*. Uffculme, Devon: Willian Publishing

Haas, H., Farrington, D. P., Killias, M., and Sattar, G. (2004). The Impact of Different Family Configurations on Delinquency. *British Journal of Criminology*, *44*(4), 520–532.

Hakanen, J. J., Bakker, A. B., and Schaufeli, W. B. (2006). Burnout and Work Engagement Among Teachers. *Journal of School Psychology*, *43*(6), 495–513.

Hales, J., Nevill, C., Pudney, S., and Tipping, S. (2009). *Longitudinal Analysis of Offending, Crime and Justice Survey 2003-06*. London: Home Office.

Halpern, D. (2005). *Social Capital*. Cam,bridge, UK: Polity.

Hand, D. J. (1994). Deconstructing Statistical Questions. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 157(3) 317–356.

Hand, D. J. (2010). *Measurement Theory and Practice: The World Through Quantification* (1st edition). New York: Wiley.

Harding, D. J., Gennetian, L., Winship, C., Sanbonmatsu, L., and Kling, J. R. (2010). *Unpacking Neighbourhood Influences on Education Outcomes: Setting The Stage For Future Research*. NBER Working Paper No. 16055 National Bureau of Economic Research.

Harmon, M. G. (2013). "Fixed" Sentencing: The Effect on Imprisonment Rates Over Time. *Journal of Quantitative Criminology*, *29*(3), 369–397.

Hartup, W. W., and Rubin, Z. (2013). *Relationships and Development*. Hillsdale, NJ Lawrenace Erlbaum Psychology Press.

Hawkins, D. F., Laub, J. H., Lauritsen, J. L., and Cothern, L. (2000). *Race, Ethnicity, and Serious and Violent Juvenile Offending*. Washington, DC: US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Hayduk, L. A. (1987). *Structural Equation Modeling with LISREL: Essentials and Advances*. Baltimore, MD,: Johns Hopkins University.

Hayner, N. S. (1933). Delinquency Areas in the Puget Sound Region. *American Journal of Sociology*, *39*(3), 314–328.

Hayward, K., and Young, J. (2007). Cultural Criminology. In R. Morgan, M. Maguire, and R. Reiner (Eds.), *The Oxford Handbook of Criminology* (4th ed., p. 102). Oxford: Oxford University Press.

Henry, B., Caspi, A., Moffitt, T. E., and Silva, P. A. (1996). Temperamental and Familial Predictors of Violent and Nonviolent Criminal Convictions: Age 3 to Age 18. *Developmental Psychology*, *32*(4), 614–623.

Hershberger, Scott L. (1994). The Specification of Equivalent Models before the Collection of Data. In A. von Eye and C. C. Clogg (Eds.), *Latent Variables Analysis: Applications for Developmental Research*. Thousand Oaks, CA: SAGE Publications, Inc.

Hills, J. (2010). *An Anatomy of Economic Inequality in the UK: Report of the National Equality Panel*. LSE STICERD Research Paper No. CASE REPORT 60.

Hills, J., Cunliffe, J., Obolenskaya, P., and Karagiannaki, E. (2015). Falling Behind, Getting Ahead: The Changing Structure of Inequality in the UK, 2007-2013. *Social Policy in a Cold Climate, Research Report*, *5*.

Hindelang, M. J., Gottfredson, M. R., and Garofalo, J. (1978). *Victims of Personal Crime: An Empirical Foundation for a Theory of Personal Victimization*. Ballinger Pub. Co.

Hindelang, M. J., Hirschi, T., and Weis, J. G. (1981). *Measuring Delinquency*. Thousand Oaks, CA: SAGE Publications.

Hipp, J. R., and Boessen, A. (2013). Egohoods as Waves Washing Across the City: A New Measure of "Neighbourhoods". *Criminology*, *51*(2), 287–327.

Hirschi, T., and Gottfredson, M. (1983). Age and the Explanation of Crime. *American Journal of Sociology*, 89(3) 552–584.

Hirschi, T., and Gottfredson, M. R. (1993). Commentary: Testing the General Theory of Crime. *Journal of Research in Crime and Delinquency*, *30*(1), 47–54.

Hobcraft, J. (1998). Intergenerational and Life-Course Transmission of Social Exclusion: Influences and Childhood Poverty, Family Disruption and Contact with the Police. *LSE STICERD Research Paper No. CASE015*.

Hogan, R. (1969). Development of an Empathy Scale. *Journal of Consulting and Clinical Psychology*, *33*(3), 307–316.

Hollway, W., and Jefferson, T. (1997). The Risk Society in an Age of Anxiety: Situating Fear of Crime. *British Journal of Sociology*, 48(2) 255–266.

Hooper, D., Coughlan, J., and Mullen, M. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. Electronic Journal of Business Research Methods, 6(1), 53-60.

Hopkins, M., and Wickson, J. (2013). Targeting Prolific and Other Priority Offenders and Promoting Pathways to Desistance: Some Reflections on the PPO Programme Using a Theory of Change Framework. *Criminology and Criminal Justice*, *13*(5), 594–614.

Hu, L., and Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure

Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Ingoldsby, E. M., and Shaw, D. S. (2002). Neighborhood Contextual Factors and Early-Starting Antisocial Pathways. *Clinical Child and Family Psychology Review*, *5*(1), 21–55.

Ishizaki, M., Miura, K., Naruse, Y., and Kido, T. (2004). Construct Validity of the Maslach Burnout Inventory. *Stress and Health*, *20*(5), 255–260.

Jackson, J., Bradford, B., Stanko, B., and Hohl, K. (2012). *Just Authority? Public Trust and Police Legitimacy*. New York: Willan.

Jackson, J., and Gouseti, I. (2014). Fear of Crime: An Entry to the Encyclopedia of Theoretical Criminology. In Mitchell Miller, J. (Ed.), *Encyclopedia of Theoretical Criminology*. Oxford, UK: Wiley-Blackwell.

Jackson, J., and Kuha, J. (2014). Worry about Crime in a Cross-National Context: A Focus on Measurement using European Social Survey Data. *Survey Research Methods, 8 (2). 109-125*

Jackson, J., and Kuha, J. (2015). How Theory Guides Measurement: Examples from the Study of Public Attitudes Toward Crime and Policing. In T. S. Bynum and B. M. Huebner (Eds.), *Handbook on Measurement Issues in Criminology and Criminal Justice*. Hoboken, New Jersey, USA: John Wylie.

Jacobs, J. (1961). *The Death and Life of Great American Cities*. New York: Random House Vintage Books.

John, O. P., and Srivastava, S. (1999). The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. in L. Pervin and O.P. John (Eds.), *Handbook of personality: Theory and research (2nd ed.)* 102–138. New York: Guilford

Johnstone, J. W. C. (1978). Social Class, Social Areas and Delinquency. *Sociology and Social Research*, *63*(1), 49–72.

Jolliffe, D., and Farrington, D. P. (2004). Empathy and Offending: A Systematic Review and Meta-analysis. *Aggression and Violent Behavior*, *9*(5), 441–476.

Jolliffe, D., and Farrington, D. P. (2006). Development and Validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*(4), 589–611.

Jolliffe, D., and Farrington, D. P. (2010). Individual Differences and Offending. In E. McLaughlin and T. Newburn (Eds.), *The SAGE Handbook of Criminological Theory* (pp. 40–56). London: SAGE Publications Ltd.

Jolliffe, D., and Farrington, D. P. (2011). Is Low Empathy Related to Bullying After Controlling for Individual and Social Background Variables? *Journal of Adolescence*, *34*(1), 59–71.

Jöreskog, K. G. (1969). A General Approach to Confirmatory Maximum Likelihood Factor Analysis. *Psychometrika*, *34*(2), 183–202.

Jöreskog, K. G. (1978). Structural Analysis of Covariance and Correlation Matrices. *Psychometrika*, *43*(4), 443–477.

Jöreskog, K. G., and Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago, IL, US: Scientific Software International.

Juby, H., and Farrington, D. P. (2001). Disentangling the Link between Disrupted Families and Delinquency Sociodemography, Ethnicity and Risk Behaviours. *British Journal of Criminology*, *41*(1), 22–40.

Junger-Tas, J., and Marshall, I. H. (1999). The Self-report Methodology in Crime Research. *Crime and Justice*, 25 291–367.

Junger-Tas, J., and Terlouw, G. J. (1994). *Delinquent Behavior Among Young People in the Western World: First Results of the International Self-report Delinquency Study*. Amsterdam: Kugler Publications.

Kagan, J., Reznick, J. S., and Gibbons, J. (1989). Inhibited and Uninhibited Types of Children. *Child Development*, *60*(4), 838–845.

Kaslow, F. W., Lipsitt, P. D., Buka, S. L., and Lipsitt, L. P. (1990). Family Law Issues in Family Therapy Practice: Early Intelligence Scores and Subsequent Delinquency: A Prospective Study. *The American Journal of Family Therapy*, *18*(2), 197–208.

Kearns, A., and Parkinson, M. (2001). The Significance of Neighbourhood. *Urban*

Studies, *38*(12), 2103–2110.

Keeley, J. E. (1991). Seed Germination and Life History Syndromes in the California Chaparral. *The Botanical Review*, *57*(2), 81–116.

Kenny, D. A. (1979). *Correlation and Causation*. New York: Wiley.

Koeske, G. F., and Koeske, R. D. (1989). Construct Validity of the Maslach Burnout Inventory: A Critical Review and Reconceptualization. *The Journal of Applied Behavioral Science, 25*(2), 131–144.

Keynes, J. M. (1936). *General Theory of Employment, Interest and Money*. London, UK: Palgrave Macmillan

Kiecolt, K. J., and Nathan, L. E. (1985). *Secondary Analysis of Survey Data*. Thousand Oaks, CA: SAGE Publications.

Kim, B., Pratt, T. C., and Wallace, D. (2013). Adverse Neighborhood Conditions and Sanction Risk Perceptions: Using SEM to Examine Direct and Indirect Effects. *Journal of Quantitative Criminology* 30 505-526

Kintrea, K., Bannister, J., Pickering, J., Reid, M., and Suzuki, N. (2008). *Young People and Territoriality in British Cities*. York: Joseph Rowntree Foundation.

Kish-Gephart, J. J., Harrison, D. A., and Treviño, L. K. (2010). Bad Apples, Bad Cases, and Bad Barrels: Meta-Analytic Evidence about Sources of Unethical Decisions at Work. *The Journal of Applied Psychology*, *95*(1), 1–31.

Klein, M. W. (1984). Offence Specialisation and Versatility Among Juveniles. *British Journal of Criminology*, *24*(2), 185–194.

Kline, R. B. (2010). *Principles and Practice of Structural Equation Modeling, Third Edition*. New York: The Guilford Press.

Kobrin, S. (1971). The Formal Legal Properties of the Shaw-Mckay Delinquency Theory. In H. L. Voss and D. M. Petersen (Eds.), *Ecology, Crime, and Delinquency*. University of California: Appleton-Century-Crofts.

Kolvin, I., Miller, F. J., Fleeting, M., and Kolvin, P. A. (1988). Social and Parenting Factors Affecting Criminal-Offence Rates. Findings from the Newcastle

Thousand Family Study (1947-1980). *The British Journal of Psychiatry*, *152*(1), 80–90.

Kott, P. S. (2007). Clarifying Some Issues in the Regression Analysis of Survey Data. *Survey Research Methods*, 1, 11–18.

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. (2006a). *Foundations of Measurement Volume I: Additive and Polynomial Representations*. Mineola, NY: Dover Publications.

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. (2006b). *Foundations of Measurement Volume II: Geometrical, Threshold, and Probabilistic Representations*. Mineola, NY: Dover Publications.

Lanier, M. M., and Henry, S. (2009). *Essential Criminology* (3rd edition). Boulder, CO: Westview.

Laub, J. H., and Sampson, R. J. (2003). *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70*. Cambridge, MA: Harvard University Press.

Laurence, J., and Heath, A. (2008). Predictors of Community Cohesion: Multi-level Modelling of the 2005 Citizenship Survey. *London, Communities and Local Government*.

Leeuw, E. D. de, Hox, J. J., and Dillman, D. (Eds.). (2008). *International Handbook of Survey Methodology* (1st edition). New York; London: Routledge.

Leventhal, T., and Brooks-Gunn, J. (2000). The Neighborhoods They Live In: The Effects of Neighborhood Residence on Child and Adolescent Outcomes. *Psychological Bulletin*, *126*(2), 309.

Leventhal, T., and Brooks-Gunn, J. (2003). Children and Youth in Neighborhood Contexts. *Current Directions in Psychological Science*, *12*(1), 27–31.

Leventhal, T., Fauth, R. C., and Brooks-Gunn, J. (2005). Neighborhood Poverty and Public Policy: a 5-Year Follow-Up of Children's Educational Outcomes in the New York City Moving to Opportunity Demonstration. *Developmental Psychology*, *41*(6), 933.

Liao, T. F. (1994). *Interpreting Probability Models: Logit, Probit, and Other Generalized Linear Models*. Thousand Oaks, CA: SAGE Publications.

Lichtenstein, P., Carlström, E., Råstam, M., Gillberg, C., and Anckarsäter, H. (2010). The Genetics of Autism Spectrum Disorders and Related Neuropsychiatric Disorders in Childhood. *The American Journal of Psychiatry*, *167*(11), 1357–1363.

Lilly, J. R., Cullen, F. T., and Ball, R. A. (2014). *Criminological Theory: Context and Consequences* (6th Ed.). Thousand Oaks, CA: SAGE Publications.

Little, T. D. (2013). *Longitudinal Structural Equation Modeling*. New York: Guilford Press.

Little, T. D., Card, N. A., Bovaird, J. A., Preacher, K. J., and Crandall, C. S. (2007). Structural Equation Modeling of Mediation and Moderation with Contextual Factors. in Little, T. D., Bovaird, J. A., and Card, N. A. *Modeling Contextual Effects in Longitudinal Studies*, 207 - 231 NJ: Lawrenace Erlbaum

Little, T. D., Slegers, D. W., and Card, N. A. (2006). A Non-Arbitrary Method of Identifying and Scaling Latent Variables in SEM and MACS Models. *Structural Equation Modeling*, *13*(1), 59–72.

Loeber, R., and Farrington, D. P. (2014). Age-Crime Curve. In G. Bruinsma and D. Weisburd (Eds.), *Encyclopedia of Criminology and Criminal Justice* (pp. 12–18). New York: Springer.

Loeber, R., and Le Blanc, M. (1990). Toward a Developmental Criminology. In Tonry Morris, N. (Eds) *Crime and Justice*, Vol 12 375–473. Chicago: University of Chicago Press

Loeber, R., and Stouthamer-Loeber, M. (1986). Family Factors as Correlates and Predictors of Juvenile Conduct Problems and Delinquency. *Crime & Justice*, In Tonry Morris, N. (Eds) *Crime and Justice*, Vol 7 29–149. Chicago: University of Chicago Press

Lord, F. M. (1967). A Paradox in the Interpretation of Group Comparisons. *Psychological Bulletin*, *68*(5), 304–305.

Lovett, B. J., and Sheffield, R. A. (2007). Affective Empathy Deficits in Aggressive

Children and Adolescents: A Critical Review. *Clinical Psychology Review*, *27*(1), 1–13.

Lupton, R. (2003). Neighbourhood Effects: Can We Measure Them and Does It Matter? *LSE STICERD Research Paper No. CASE073*.

Maassen, G. H., and Bakker, A. B. (2001). Suppressor Variables in Path Models Definitions and Interpretations. *Sociological Methods & Research*, *30*(2), 241–270.

MacKinnon, D. P. (2008). *Introduction to Statistical Mediation Analysis*. New York: Routledge.

MacKinnon, D. P., and Dwyer, J. H. (1993). Estimating Mediated Effects in Prevention Studies. *Evaluation Review*, *17*(2), 144–158.

MacKinnon, D. P., Fairchild, A. J., and Fritz, M. S. (2007). Mediation Analysis. *Annual Review of Psychology*, *58*, 593.

MacKinnon, D. P., Lockwood, C. M., Brown, C. H., Wang, W., and Hoffman, J. M. (2007). The Intermediate Endpoint Effect in Logistic and Probit Regression. *Clinical Trials*, *4*(5), 499–513.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., and Sheets, V. (2002). A Comparison of Methods to Test Mediation and Other Intervening Variable Effects. *Psychological Methods*, *7*(1), 83.

Marsh, H. W., Hau, K.-T., and Grayson, D. (2005). Goodness of Fit in Structural Equation Models. In A. Maydeu-Olivares and J. J. McArdle (Eds.), *Contemporary Psychometrics* (pp. 275–340). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Marsh, H. W., Hau, K.-T., and Wen, Z. (2004). In Search Of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling*, *11*(3), 320–341.

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., S, J., Trautwein, U., and Nagengast, B. (2010). A New Look at the Big Five Factor Structure through

Exploratory Structural Equation Modeling. *Psychological Assessment*, *22*(3), 471–491.

Wolfgang, M. E., Figlio, R. M., and Sellin, T. (1987). Delinquency in a Birth Cohort (New edition). Chicago: University of Chicago Press.

May-Chahal, C., and Cawson, P. (2005). Measuring Child Maltreatment in the United Kingdom: A Study of the Prevalence of Child Abuse and Neglect. *Child Abuse & Neglect*, *29*(9), 969–984.

Mazerolle, L., Wickes, R., and McBroom, J. (2010). Community Variations in Violence: The Role of Social Ties and Collective Efficacy in Comparative Context. *Journal of Research in Crime and Delinquency*, *47*(1), 3–30.

McAra, L. (2003). Negotiated Order: Gender, Youth Transitions and Crime. In *Selected Papers from the British Society of Criminology Conference Bangor* (Vol. 6).

McAra, L., and McVie, S. (2005). The Usual Suspects? Street-life, Young People and the Police. *Criminal Justice*, *5*(1), 5–36.

McAra, L., and McVie, S. (2010). Youth Crime and Justice: Key Messages from the Edinburgh Study of Youth Transitions and Crime. *Criminology and Criminal Justice*, *10*(2), 179–209.

McAra, L., and McVie, S. (2012a). Critical Debates in Developmental and Life-Course Criminology. In R. Morgan, M. Maguire, and R. Reiner (Eds.), *The Oxford Handbook of Criminology* (pp. 531–562). Oxford: Oxford University Press.

McAra, L., and McVie, S. (2012b). Negotiated Order: The Groundwork for a Theory of Offending Pathways. *Criminology and Criminal Justice*, *12*(4), 347–375.

McCrae, R. R., and John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, *60*(2), 175–215.

McGrath, R. E., Mitchell, M., Kim, B. H., and Hough, L. (2010). Evidence for Response Bias as a Source of Error Variance in Applied Assessment. *Psychological Bulletin*, *136*(3), 450.

McVie, S. (2005). Patterns of Deviance Underlying the Age-Crime Curve: The Long

Term Evidence. *British Society of Criminology E-Journal*, *7*, 1–15.

McVie, S., and Norris, P. (2006a). *Neighbourhood Effects on Youth Delinquency and Drug Use*. Edinburgh: University of Edinburgh, Centre for Law and Society.

McVie, S., and Norris, P. (2006b). *The Effect of Neighbourhoods on Adolescent Property Offending*. Edinburgh: University of Edinburgh, Centre for Law and Society.

Meade, A. W., Johnson, E. C., and Braddy, P. W. (2006). The Utility of Alternative Fit Indices in Tests of Measurement Invariance. In *Academy of Management Proceedings* (Vol. 2006, pp. B1–B6). Academy of Management.

Meldrum, R. C., Young, J. T. N., Hay, C., and Flexon, J. L. (2012). Does Self-Control Influence Maternal Attachment? A Reciprocal Effects Analysis from Early Childhood through Middle Adolescence. *Journal of Quantitative Criminology*, *28*(4), 673–699.

Merton, R. K. (1938). Social Structure and Anomie. *American Sociological Review*, *3*(5), 672–682.

Messick, S. (1988). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd edition, pp. 13–103). New York: Macmillan.

Messner, S. F. (1988). Merton's "Social Structure and Anomie": The Road Not Taken. *Deviant Behavior*, *9*(1), 33–53.

Messner, S. F. (1989). Economic Discrimination and Societal Homicide Rates: Further Evidence on the Cost of Inequality. *American Sociological Review*, 597–611.

Messner, S. F., Baumer, E. P., and Rosenfeld, R. (2004). Dimensions of Social Capital and Rates of Criminal Homicide. *American Sociological Review*, *69*(6), 882–903.

Messner, S. F., Raffalovich, L. E., and McMillan, R. (2001). Economic Deprivation and Changes in Homicide Arrest Rates for White and Black Youths, 1967–1998: A National Time-Series Analysis. *Criminology*, *39*(3), 591–614.

Michaelson, J., Pickering, K., Wood, N., and Scholes, S. (2006). *2005 Home Office Citizenship Survey Technical Report*. London: Home Office.

Miller, J., Flory, K., Lynam, D., and Leukefeld, C. (2003). A Test of the Four-Factor

Model of Impulsivity-Related Traits. *Personality and Individual Differences*, *34*(8), 1403–1418.

Mills, J. F., Loza, W., and Kroner, D. G. (2003). Predictive Validity Despite Social Desirability: Evidence for the Robustness of Self-Report Among Offenders. *Criminal Behaviour and Mental Health*, *13*(2), 140–150.

Mitchell, O., and MacKenzie, D. L. (2006). The Stability and Resiliency of Self-Control in a Sample of Incarcerated Offenders. *Crime & Delinquency*, *52*(3), 432–449.

Moffitt, T. E. (1993). Adolescence-Limited and Life-Course-Persistent Antisocial Behavior: A Developmental Taxonomy. *Psychological Review*, *100*(4), 674.

Moffitt, T. E., Caspi, A., Harrington, H., and Milne, B. J. (2002). Males on the Life-Course-Persistent and Adolescence-Limited Antisocial Pathways: Follow-Up At Age 26 Years. *Development and Psychopathology*, *14*(1), 179–207.

Morenoff, J. D., Sampson, R. J., and Raudenbush, S. W. (2001). Neighborhood Inequality, Collective Efficacy, and the Spatial Dynamics of Urban Violence. *Criminology*, *39*(3), 517–558.

Morgan, R., Maguire, M., and Reiner, R. (2012). *The Oxford Handbook of Criminology*. Oxford: Oxford University Press.

Mortel, V. D., and Francisca, T. (2008). Faking it: Social Desirability Response Bias in Self-report Research. *Australian Journal of Advanced Nursing*, *25*(4), 40–48.

Mulaik, S. A. (1986). Toward a Synthesis of Deterministic and Probabilistic Formulations of Causal Relations by the Functional Relation Concept. *Philosophy of Science*, *53*(3), 313–332.

Mulaik, S. A. (2009). *Linear Causal Modeling with Structural Equations*. Boca Raton: Chapman and Hall/CRC.

Muller, D., Judd, C. M., and Yzerbyt, V. Y. (2005). When Moderation is Mediated and Mediation is Moderated. *Journal of Personality and Social Psychology*, *89*(6), 852.

Muthén, B., and Asparouhov, T. (2014). Causal Effects in Mediation Modeling: An Introduction with Applications to Latent Variables. *Forthcoming in Structural*

Equation Modeling.

Muthén, L., and Muthén, B. (1998-2012). *Mplus User's Guide. Seventh Edition.* (Vol. 7). Los Angeles: Muthén & Muthén.

Na, C., and Paternoster, R. (2012). Can Self-Control Change Substantially Over Time? Rethinking the Relationship Between Self- And Social Control. *Criminology*, *50*(2), 427–462.

Nagin, D. S., Farrington, D. P., and Moffitt, T. E. (1995). Life-Course Trajectories of Different Types of Offenders. *Criminology*, *33*(1), 111–139.

Nagin, D. S., and Snodgrass, G. M. (2013). The Effect of Incarceration on Re-Offending: Evidence from a Natural Experiment in Pennsylvania. *Journal of Quantitative Criminology*, *29*(4), 601–642.

Nagler, J. (1994). *Interpreting Probit Analysis*. New York: New York University Press.

Nakhaie, M. R., and Sacco, V. F. (2009). Social Capital, Individual Disorders and Property Offences. *International Journal of Law and Psychiatry*, *32*(6), 392–399.

Neighbourhood Renewal Unit. (2004). *The English Indices of Deprivation 2004: Summary (revised)*. London: Office of the Deputy Prime Minister.

Newman, O. (1973). *Defensible Space; Crime Prevention Through Urban Design.* New York: Macmillan.

Noble, M., Wright, G., Dibben, C., Smith, G., McLennan, D., Anttila, C. Barnes, H., Mokhtar, C., Avenell, D. (2004). Indices of Deprivation 2004. *Report to the Office of the Deputy Prime Minister.* London: Neighbourhood Renewal Unit.

Norris, C. M., Ghali, W. A., Saunders, L. D., Brant, R., Galbraith, D., Faris, P., and Knudtson, M. L. (2006). Ordinal Regression Model and the Linear Regression Model Were Superior to the Logistic Regression Models. *Journal of Clinical Epidemiology*, *59*(5), 448–456.

Odgers, C. L., Moffitt, T. E., Tach, L. M., Sampson, R. J., Taylor, A., Matthews, C. L., and Caspi, A. (2009). The Protective Effects of Neighbourhood Collective Efficacy on British Children Growing Up in Deprivation: A Developmental Analysis.

Developmental Psychology, *45*(4), 942.

Home Office (2004). *Building Communities, Beating Crime: A Better Police Service for the 21st Century* (Vol. 6360). TSO Shop.

Osborn, S. G. (1980). Moving Home, Leaving London and Delinquent Trends. *British Journal of Criminology*, *20*(1), 54–61.

Ouston, J. (1984). Delinquency, Family Background and Educational Attainment. *British. Journal of Criminology*, *24*, 2.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference* (Reprinted with corrections edition). Cambridge; New York: Cambridge University Press.

Pearl, J. (2001). Direct and Indirect Effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence* (pp. 411–420). San Francisco, CA: Morgan Kaufmann Publishers Inc.

Pearson, D. A., McDougall, C., Kanaan, M., Bowles, R. A., and Torgerson, D. J. (2011). Reducing Criminal Recidivism: Evaluation of Citizenship, an Evidence-Based Probation Supervision Process. *Journal of Experimental Criminology*, *7*(1), 73–102.

Piquero, A. R., Farrington, D. P., and Blumstein, A. (2003). The Criminal Career Paradigm. *Crime and Justice*, 30 359–506.

Piquero, A. R., Farrington, D. P., Nagin, D. S., and Moffitt, T. E. (2010). Trajectories of Offending and their Relation to Life Failure in Late Middle Age: Findings from The Cambridge Study in Delinquent Development. *Journal of Research in Crime and Delinquency*, *47*(2), 151–173.

Piquero, A. R., Farrington, D. P., Welsh, B. C., Tremblay, R., and Jennings, W. G. (2009). Effects of Early Family/Parent Training Programs on Antisocial Behavior and Delinquency. *Journal of Experimental Criminology*, *5*(2), 83–120.

Piquero, A. R., and Hickman, M. (1999). An Empirical Test of Tittle's Control Balance Theory. *Criminology*, *37*(2), 319–342.

Piquero, A. R., Jennings, W. G., and Farrington, D. P. (2010). On the Malleability of Self-Control: Theoretical and Policy Implications Regarding a General Theory of

Crime. *Justice Quarterly*, *27*(6), 803–834.

Plomin, R., Kagan, J., Emde, R.N., Reznick, J.S., Braungart, J.M., Robinson, J., Campos, J., Zahn-Waxler, C., Corley, R., and Fulker, D.W. (1993). Genetic Change and Continuity from Fourteen to Twenty Months: The MacArthur Longitudinal Twin Study. *Child Development*, *64*(5), 1354–1376.

Portes, A. (1998). Social Capital: Its Origins and Applications in Modern Sociology. *Annual Review of Sociology*, 24, 1–24.

Portes, A., and Sensenbrenner, J. (1993). Embeddedness and Immigration: Notes on the Social Determinants of Economic Action. *American Journal of Sociology*, *98*(6), 1320–1350.

Pratt, T. C., and Cullen, F. T. (2000). The Empirical Status of Gottfredson and Hirschi's General Theory of Crime: A Meta-Analysis. *Criminology*, *38*(3), 931–964.

Pratt, T. C., and Cullen, F. T. (2005). Assessing Macro-Level Predictors and Theories of Crime: A Meta-Analysis. *Crime and Justice*, 32 373–450.

Pratt, T. C., Cullen, F. T., Blevins, K. R., Daigle, L., and Unnever, J. D. (2002). The Relationship of Attention Deficit Hyperactivity Disorder to Crime and Delinquency: A Meta-Analysis. *International Journal of Police Science & Management*, *4*(4), 344–360.

Preacher, K. J., and Hayes, A. F. (2008). Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models. *Behavior Research Methods*, *40*(3), 879–891.

Preacher, K. J., Rucker, D. D., and Hayes, A. F. (2007). Addressing Moderated Mediation Hypotheses: Theory, Methods, and Prescriptions. *Multivariate Behavioral Research*, *42*(1), 185–227.

Prime, J., White, S., Liriano, S., and Britain, G. (2001). *Criminal Careers of those Born Between 1953 And 1978*. London: Great Britain, Home Office, Research, Development and Statistics Directorate.

Putnam, R. D. (2001). *Bowling Alone: The Collapse and Revival of American Community*.

New York: Simon & Schuster.

Rafter, N. (2004). The Unrepentant Horse-Slasher: Moral Insanity and the Origins of Criminological Thought. *Criminology*, *42*(4), 979–1008.

Reisig, M. D., and Cancino, J. M. (2004). Incivilities in Nonmetropolitan Communities: The Effects of Structural Constraints, Social Conditions, and Crime. *Journal of Criminal Justice*, *32*(1), 15–29.

Riggio, R. E., Widaman, K. F., Tucker, J. S., and Salinas, C. (1991). Beauty is More Than Skin Deep: Components of Attractiveness. *Basic and Applied Social Psychology*, *12*(4), 423–439.

Robbins, B., and Pettinicchio, D. (2012). Social Capital, Economic Development, and Homicide: A Cross-National Investigation. *Social Indicators Research*, *105*(3), 519–540.

Park. R. E. (1920). *Introduction to The Science Of Sociology* (2nd ed.). Chicago: University of Chicago Press

Robins, J. M., and Greenland, S. (1992). Identifiability and Exchangeability for Direct and Indirect Effects. *Epidemiology*, 3 (2) 143–155.

Robinson, J. L., Kagan, J., Steven, J., and Corley, R. (1992). The Heritability of Inhibited And Uninhibited Behavior: A Twin Study. *Developmental Psychology*, *28*(6), 1030–1037.

Robison, S. M. (1936). *Can Delinquency be Measured?* New York: Welfare Council of New York City.

Roe, S., and Ashe, J. (2008). *Young People and Crime: Findings from the 2006 Offending, Crime and Justice survey.* Home Office Statistical Bulletin available online at http://dera.ioe.ac.uk/9140/1/hosb0908.pdf

Rutter, M. (1971). Parent-Child Separation: Psychological Effects on the Children. *Journal of Child Psychology and Psychiatry*, *12*(4), 233–260.

Rutter, M. (1979). Maternal Deprivation, 1972-1978: New Findings, New Concepts, New Approaches. *Child Development*, 50 (2) 283–305.

Rutter, M., Giller, H., and Hagell, A. (1998). *Antisocial Behavior by Young People: A Major New Review*. Cambridge: Cambridge University Press.

Sampson, R. J. (2006). How Does Community Context Matter? Social Mechanisms and the Explanation of Crime. In P.-O. H. Wikström and R. J. Sampson (Eds.), *The Explanation of Crime: Context, Mechanisms and Development* (pp. 31–60). Cambridge: Cambridge University Press.

Sampson, R. J. (2008a). Collective Efficacy Theory: Lessons Learned and Directions for Future Inquiry. in Cullen, F. T., Wright, J. P. and Blevins., K. (Eds*) Taking Stock: The Status of Criminological Theory. Advances in Criminological Theory Vol. 15*, 149. New Brunswick, NJ: Transaction Publishers.

Sampson, R. J. (2008b). Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure. *American Journal of Sociology*, *114*(1), 189–231.

Sampson, R. J. (2012). *Great American City: Chicago and the Enduring Neighborhood Effect*. Chicago; London: The University of Chicago Press.

Sampson, R. J., and Groves, W. B. (1989). Community Structure and Crime: Testing Social-Disorganization Theory. *American Journal of Sociology*, 94 (4) 774–802.

Sampson, R. J., and Laub, J. H. (1993). *Crime in the Making: Pathways and Turning Points Through Life*. Cambridge, MA: Harvard University Press.

Sampson, R. J., and Laub, J. H. (2005). A Life-Course View of the Development of Crime. *The Annals of the American Academy of Political and Social Science*, *602*(1), 12–45.

Sampson, R. J., Morenoff, J. D., and Earls, F. (1999). Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children. *American Sociological Review*, 64 (5) 633–660.

Sampson, R. J., Morenoff, J. D., and Gannon-Rowley, T. (2002). Assessing "Neighborhood Effects": Social Processes and New Directions in Research. *Annual Review of Sociology*, 28 443–478.

Sampson, R. J., and Raudenbush, S. W. (1999). Systematic Social Observation of Public

Spaces: A New Look at Disorder in Urban Neighborhoods. *American Journal of Sociology*, *105*(3), 603–651.

Sampson, R. J., and Raudenbush, S. W. (2001). *Disorder in Urban Neighborhoods: Does it Lead to Crime*. US Department of Justice, Office of Justice Programs, National Institute of Justice.

Sampson, R. J., Raudenbush, S. W., and Earls, F. (1997). Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. *Science*, *277*(5328), 918–924.

Scarborough, B. K., Like-Haislip, T. Z., Novak, K. J., Lucas, W. L., and Alarid, L. F. (2010). Assessing the Relationship Between Individual Characteristics, Neighborhood Context, and Fear of Crime. *Journal of Criminal Justice*, *38*(4), 819–826.

Schafer, J. A., Huebner, B. M., and Bynum, T. S. (2006). Fear of Crime and Criminal Victimization: Gender-Based Contrasts. *Journal of Criminal Justice*, *34*(3), 285–301.

Schreck, C. J. (1999). Criminal Victimization and Low Self-Control: An Extension and Test of a General Theory of Crime. *Justice Quarterly*, *16*(3), 633–654.

Schweinhart, L. J., Barnes, H. V., and Weikart, D. P. (1993). *Significant Benefits: The High/Scope Perry Preschool Study Through Age 27*. Ypsilanti, MI: High/Scope Press.

Seale, C. (2003). *Social Research Methods: A Reader*. London; New York: Routledge.

Sellin, T. (1931). The Basis of a Crime Index. *Journal of Criminal Law and Criminology (1931-1951)*, 22 (3) 335–356.

Shannon, L. W. (1982). *Assessing the Relationship of Adult Criminal Careers to Juvenile Careers: A Summary.* U.S. Dept. of Justice, Office of Juvenile Justice and Delinquency Prevention, National Institute for Juvenile Justice and Delinquency Prevention. http://purl.fdlp.gov/GPO/gpo10929.

Sharp, C., Aldridge, J., Medina, J., and Britain, G. (2006). *Delinquent Youth Groups and Offending Behaviour: Findings from the 2004 Offending, Crime and Justice Survey*. London: Home Office.

Sharp, C., Budd, T., and Britain, G. (2005). *Minority Ethnic Groups and Crime: Findings*

from the Offending, Crime and Justice Survey 2003. London: Research Development and Statistics Directorate.

Short, J. F., and Nye, F. I. (1957). Reported Behavior as a Criterion of Deviant Behavior. *Social Problems*, *5*(3), 207–213.

Simons, R. L., WU, C.-I., Johnson, C., and Conger, R. D. (1995). A Test of Various Perspectives on the Intergenerational Transmission of Domestic Violence. *Criminology*, *33*(1), 141–172.

Simpson, E. H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society, Series B (Methodological)*, 13 (2) 238–241.

Sivo, S. A., Fan, X., Witta, E. L., and Willse, J. T. (2006). The Search for "Optimal" Cutoff Properties: Fit Index Criteria in Structural Equation Modeling. *The Journal of Experimental Education*, *74*(3), 267–288.

Skardhamar, T., and Skirbekk, V. (2013). Relative Mortality among Criminals in Norway and the Relation to Drug and Alcohol Related Offenses. *PloS One*, *8*(11),

Smith, D. J., and McVie, S. (2003). Theory and Method in the Edinburgh Study of Youth Transitions and Crime. *British Journal of Criminology*, *43*(1), 169–195.

Sobel, M. E. (1986). Some New Results on Indirect Effects and their Standard Errors in Covariance Structure Models. *Sociological Methodology*, *16*, 159–186.

Solon, G., Haider, S. J., and Wooldridge, J. (2013). *What are we Weighting for?* NBER Working Papers 18859, National Bureau of Economic Research, Inc.

Spearman, C. (1904). The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, *15*(1), 72–101.

Stanford, M. S., Mathias, C. W., Dougherty, D. M., Lake, S. L., Anderson, N. E., and Patton, J. H. (2009). Fifty Years of the Barratt Impulsiveness Scale: An Update and Review. *Personality and Individual Differences*, *47*(5), 385–395.

Steiger, J. H., and Lind, J. C. (1980). Statistically Based Tests for The Number of Common Factors. in *Annual Meeting of the Psychometric Society, Iowa City, IA* (Vol. 758).

Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*, New Series, 103, (2684) 677-680

Stevens, S. S. (1951). Mathematics, Measurement, and Psychophysics. In *Handbook of Experimental Psychology* (pp. 1–49). Oxford: Wiley.

Suppes, P., Krantz, D. H., Luce, R. D., and Tversky, A (2006). *Foundations of Measurement Volume III: Representation, Axiomatization, and Invariance*. Mineola, NY: Dover Publications.

Sutherland, A., Brunton-Smith, I., and Jackson, J. (2013). Collective Efficacy, Deprivation and Violence in London. *British Journal of Criminology* 53 (6) 1050-1074

Sutherland, E. H. (1949). *White Collar Crime*. New York: Holt, Rinehart & Winston

Suttles, G. D. (1972). *Social Construction of Communities*. Chicago: University of Chicago Press.

Tabachnick, B. G., and Fidell, L. S. (2005). *Using Multivariate Statistics*. Boston: Pearson.

Tangney, J. P., Baumeister, R. F., and Boone, A. L. (2004). High Self-Control Predicts Good Adjustment, Less Pathology, Better Grades, and Interpersonal Success. *Journal of Personality*, *72*(2), 271–324.

Taylor, R. B. (1997). Social Order and Disorder of Street Blocks and Neighborhoods: Ecology, Microecology, and the Systemic Model of Social Disorganization. *Journal of Research in Crime and Delinquency*, *34*(1), 113–155.

Thomas, W., and Znaniecki, F. (1920). *The Polish Peasant in Europe and America* (Student edition). Urbana, IL: University of Illinois Press.

Thornberry, T. P., and Farnworth, M. (1982). Social Correlates of Criminal Involvement: Further Evidence on the Relationship Between Social Status and Criminal Behavior. *American Sociological Review*, *47*(4), 505–518.

Thornberry, T. P., and Krohn, M. D. (2000). The Self-Report Method for Measuring Delinquency and Crime. *Criminal Justice*, *4*(1), 33–83.

Tittle, C. R. (1995). *Control Balance: Toward a General Theory of Deviance*. Boulder, CO:

Westview Press.

Tseloni, A., and Zarafonitou, C. (2008). Fear of Crime and Victimization: A Multivariate Multilevel Analysis of Competing Measurements. *European Journal of Criminology*, *5*(4), 387–409.

Tucker, L. R., and Lewis, C. (1973). A Reliability Coefficient for Maximum Likelihood Factor Analysis. *Psychometrika*, *38*(1), 1–10.

Turner, M. G., and Piquero, A. R. (2002). The Stability of Self-Control. *Journal of Criminal Justice*, *30*(6), 457–471.

Tu, Y.-K., Gunnell, D., and Gilthorpe, M. S. (2008). Simpson's Paradox, Lord's Paradox, and Suppression Effects are the Same Phenomenon – the Reversal Paradox. *Emerging Themes in Epidemiology*, *5*, 2.

Twigg, L., Taylor, J., and Mohan, J. (2010). Diversity or Disadvantage? Putnam, Goodhart, Ethnic Heterogeneity, and Collective Efficacy. *Environment and Planning. A*, 42 (6), 1421.

Ullman, J. B. (2006). Structural Equation Modeling: Reviewing the Basics and Moving Forward. *Journal of Personality Assessment*, *87*(1), 35–50.

Valeri, L., and VanderWeele, T. J. (2013). Mediation Analysis Allowing for Exposure–Mediator Interactions and Causal Interpretation: Theoretical Assumptions and Implementation with SAS and SPSS Macros. *Psychological Methods*, *18*(2), 137.

Van Der Linden, D., Nijenhuis, J., and Bakker, A. B. (2010). The General Factor of Personality: A Meta-Analysis of Big Five Intercorrelations and a Criterion-Related Validity Study. *Journal of Research in Personality*, *44*(3), 315–327.

Vollmer, T. R., Borrero, J. C., Lalli, J. S., and Daniel, D. (1999). Evaluating Self-Control and Impulsivity in Children With Severe Behavior Disorders. *Journal of Applied Behavior Analysis*, *32*(4), 451–466.

Wang, J., and Wang, X. (2012). *Structural Equation Modeling: Applications Using Mplus*. Chichester, West Sussex: John Wiley & Sons.

Webster, C. J. (2015). Youth Crime and Justice. In B. Goldson and J. Muncie (Eds.), *Youth*

Crime and Justice (2nd Ed., pp. 31–48). Thousand Oaks, CA: SAGE Publications Ltd.

Weerman, F. M. (2011). Delinquent Peers in Context: A Longitudinal Network Analysis of Selection and Influence Effects. *Criminology*, *49*(1), 253–286.

Weerman, F. M., Maxson, C. L., Esbensen, F.-A., Aldridge, J., Medina, J., and van Gemert, F. (2009). *Eurogang Program Manual. University of Missouri-St. Louis., Online Report*. available at https://www.umsl.edu/ccj/Eurogang/EurogangManual.pdf

Wells, W., Schafer, J. A., Varano, S. P., and Bynum, T. S. (2006). Neighborhood Residents' Production of Order: The Effects of Collective Efficacy on Responses to Neighborhood Problems. *Crime & Delinquency*, *52*(4), 523–550.

Welsh, B. C., and Farrington, D. P. (2012). *The Oxford Handbook of Crime Prevention*. New York: Oxford University Press.

Whiteside, S. P., and Lynam, D. R. (2001). The Five Factor Model and Impulsivity: Using a Structural Model of Personality to Understand Impulsivity. *Personality and Individual Differences*, *30*(4), 669–689.

Whitworth, A. (2012). Inequality and Crime Across England: A Multilevel Modelling Approach. *Social Policy and Society*, *11*, 27–40.

Wikström, P.-O. H. (2009). Crime Propensity, Criminogenic Exposure and Crime Involvement in Early to Mid-Adolescence. *Monatsschrift Für Kriminologie Und Strafrechtsreform*, *92*(2/3), 253–266.

Wikström, P.-O. H. (2012). *Breaking Rules: The Social and Situational Dynamics of Young People's Urban Crime*. Oxford: Oxford University Press.

Wikström, P.-O. H., Ceccato, V., Hardie, B., and Treiber, K. (2010). Activity Fields and The Dynamics Of Crime. *Journal of Quantitative Criminology*, *26*(1), 55–87.

Wikström, P.-O. H., and Loeber, R. (2000). Do Disadvantaged Neighborhoods Cause Well-Adjusted Children to Become Adolescent Delinquents? A Study of Male Juvenile Serious Offending, Individual Risk and Protective Factors, and

Neighborhood Context. *Criminology*, *38*(4), 1109–1142.

Wikström, P.-O. H., and Sampson, R. J. (2003). Social Mechanisms of Community Influences on Crime and Pathways in Criminality. in Lahey, B. B., Moffitt, T.E. and Caspi, A. (Eds) *The Causes of Conduct Disorder and Serious Juvenile Delinquency*, 118–148. New York: Guildford Press

Wikström, P.-O. H., and Treiber, K. (2007). The Role of Self-Control in Crime Causation Beyond Gottfredson and Hirschi's General Theory of Crime. *European Journal of Criminology*, *4*(2), 237–264.

Wikström, P.-O. H., Treiber, K., Hardie, B., and Oberwittler, D. (2015). Felson's Review of "Breaking Rules": Smoke and Mirrors. *Journal of Criminal Justice Education*, *26*(1), 115–116.

Wilson, D., Sharp, C., and Patterson, A. (2006). *Young People and Crime: Findings from the 2005 Offending, Crime and Justice Survey.* Home Office Statistical Bulletin-Research Development And Statistics Directorate available online at http://dera.ioe.ac.uk/6902/1/hosb1706.pdf

Wilson, J. Q., and Kelling, G. L. (1982). Broken Windows. *Atlantic Monthly*, *249*(3), 29–38.

Winship, C., and Radbill, L. (1994). Sampling Weights and Regression Analysis. *Sociological Methods & Research*, *23*(2), 230–257.

Wood, P. B., Pfefferbaum, B., and Arneklev, B. J. (1995). Risk-Taking and Self-Control: Social Psychological Correlates of Delinquency. *Journal of Crime and Justice*, *16*(1), 111–130.

Wood, R. (1978). Fitting the Rasch Model—A Heady Tale. *British Journal of Mathematical and Statistical Psychology*, *31*(1), 27–32.

Wooldridge, J. (2008). *Introductory Econometrics: A Modern Approach*. Cengage Learning.

Wright, J. P., Tibbetts, S. G., and Daigle, L. E. (2008). Criminals in the Making: Criminality Across the Life Course. Los Angeles: SAGE Publications, Inc.

Wright, S. (1918). On the Nature of Size Factors. *Genetics*, *3*(4), 367.

Wright, S. (1921). Correlation and Causation. *Journal of Agricultural Research*, *20*(7), 557–585.

Xie, Y. (2007). Otis Dudley Duncan's legacy: The Demographic Approach to Quantitative Reasoning in Social Science. *Research in Social Stratification and Mobility*, *25*(2), 141–156.

Xu, Y., Fiedler, M. L., and Flaming, K. H. (2005). Discovering the Impact of Community Policing: The Broken Windows Thesis, Collective Efficacy, and Citizens' Judgment. *Journal of Research in Crime and Delinquency*, *42*(2), 147–186.

Young, J. (2003). Merton with Energy, Katz with Structure: The Sociology of Vindictiveness and the Criminology of Transgression. *Theoretical Criminology*, *7*(3), 388–414.

Yu, C. Y. (2002). *Evaluating Cutoff Criteria of Model Fit Indices for Latent Variable Models with Binary and Continuous Outcomes*. Los Angeles: University of California Press.

Zara, G., and Farrington, D. P. (2009). Childhood and Adolescent Predictors of Late Onset Criminal Careers. *Journal of Youth and Adolescence*, *38*(3), 287–300.