Risk Factors for the Development of Sexually Abusive Behaviour in Sexually Victimised Males: A Catch-up Longitudinal Design

Dean McMillan

Institute of Child Health
University College London

Ph.D. Thesis
**Abstract**

Male children who are sexually abused may be at heightened risk of sexually abusing others. However, by no means every victim becomes a perpetrator, which suggests that other risk factors must be involved. The first aim of the study was to identify the additional risk factors that may be causally related to subsequent abusive behaviour. A second aim was to establish whether these risk factors could be used to predict which sexually victimised males were at increased risk of perpetration before the onset of this behaviour.

The sample consisted of sexually victimised males referred to Great Ormond Street Hospital (N = 104). Participants were divided into those who had gone on to sexually perpetrate (N = 21) and those who had not done so (N = 83). The study used a catch-up longitudinal design. Each participant’s experience of the key risk factors was assessed using contemporaneous case material. Evidence of subsequent sexual perpetration was gathered from the case material and police records.

Three risk factors differentiated between the two groups: sexual victimisation by a female, neglect (failure to provide) and neglect (lack of supervision). Witnessing intrafamilial physical abuse fell just short of significance. A logistic regression examined the independent contribution of these risk factors to sexual perpetration. In this regression, only neglect (failure to provide) remained a significant predictor of sexual perpetration (Adjusted OR = 3.73; 95% CI 1.12 to 12.42). Four risk indexes were generated to assess the predictive capacity of the risk factors. The area under the curve in a series of ROC-curve analyses indicated that all of the indexes successfully discriminated between sexually victimised males who had gone on to perpetrate and those who had not done so. The methodological, theoretical, and policy and practice implications of these findings are discussed.
Acknowledgements

I would like to thank my mum and dad for their warmth, love and encouragement. Specific thanks go to my mum for walking me to the library once a week throughout my childhood, in every variety of Yorkshire weather; and to my dad for introducing me, memorably, to the concept of probability. I would also like to thank Damian, my brother and my friend, for his ability to make me laugh, and, along with Rachel, for providing me with the best incentive to get the PhD finished: the birth of their baby. It would be impossible to thank all of my relations separately; I can only say that I have appreciated feeling part of this vast and ever increasing extended family. My Nan and aunt Kath do require separate mention, and thanks go to them.

I also thank Paul Arundel for fifteen years of friendship, and for providing the comfort of knowing that someone else will see it as I do. I would also like to thank the rest of the Arundel family, and my friends Melvin Rickarby, Paul Lehman and Tom Ashden.

This thesis would not have been completed without the firm grasp of methodology, unflagging support and subtle positive reinforcement schedule of my supervisor, Richard Hastings. No matter how much I thought I had thought about a research problem Richard always managed to provide a fresh insight. Most of all I would like to thank him for providing an entertaining learning environment. I am proud to consider myself a product of the Behavioural Sciences Unit, and it is only upon leaving it that I have come to fully appreciate the debt my research and clinical thinking owes to the Unit. For that I thank my other supervisor and Head of Unit, David Skuse. I would like to thank the other members of the steering committee, namely Jim Stevenson (particularly for his advice to 'read some Meehl'), Jill Hodges (a model clinician-researcher) and Amon Bentovim (for encyclopaedic knowledge of the area). Thanks are also due to the other members of the research team: Mark Richards for using his adhesive climbing abilities to stick it out and see it through, and Tiffany Talbot for commitment beyond the call of duty and the ability to perform computer magic. Special thanks go to Daniel Salter for sharing in
the excitement of new ideas, and his ability to provide the soundest of sounding boards. I would also like to mention Jennifer Smith and Vivien Prior for making my time at the Behavioural Sciences Unit that much more enjoyable.

Finally I would like to thank Shubhu, my partner and friend. Thank you for enduring the long, dull sacrifice of evenings and weekends when I have been wedded to the laptop. And thank you for having the knack of always being able to cheer me up. Thanks also go to Shubhu’s mum for the food, Shubhu’s dad for the wise words, and Shubhu’s sister for her dry humour.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 Introduction</strong></td>
<td>21</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>21</td>
</tr>
<tr>
<td>1.2 Overview of the thesis</td>
<td>21</td>
</tr>
<tr>
<td>1.3 Summary</td>
<td>22</td>
</tr>
<tr>
<td><strong>2 Introduction to literature review</strong></td>
<td>24</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>24</td>
</tr>
<tr>
<td>2.2 Limits of the review</td>
<td>24</td>
</tr>
<tr>
<td>2.3 Overview of the review</td>
<td>24</td>
</tr>
<tr>
<td><strong>3 The extent of the problem</strong></td>
<td>28</td>
</tr>
<tr>
<td>3.1 Definitions of child sexual abuse</td>
<td>28</td>
</tr>
<tr>
<td>3.1.1 Research definitions</td>
<td>29</td>
</tr>
<tr>
<td>3.1.1.1 The type of activity involved</td>
<td>29</td>
</tr>
<tr>
<td>3.1.1.2 The age of the victim</td>
<td>29</td>
</tr>
<tr>
<td>3.1.1.3 Victim and perpetrator age</td>
<td>30</td>
</tr>
<tr>
<td>differential</td>
<td>30</td>
</tr>
<tr>
<td>3.1.1.4 Perpetrator-victim relationship</td>
<td>30</td>
</tr>
<tr>
<td>3.1.1.5 Consent and power</td>
<td>30</td>
</tr>
<tr>
<td>3.1.2 Legal definitions</td>
<td>31</td>
</tr>
<tr>
<td>3.1.3 Social work definitions</td>
<td>31</td>
</tr>
<tr>
<td>3.1.4 Discussion</td>
<td>33</td>
</tr>
<tr>
<td>3.1.5 Summary</td>
<td>34</td>
</tr>
<tr>
<td>3.2 Definitions of sexual perpetration by young people</td>
<td>34</td>
</tr>
<tr>
<td>3.2.1 Introduction</td>
<td>34</td>
</tr>
<tr>
<td>3.2.2 Definition of sexual behaviour in young people</td>
<td>34</td>
</tr>
<tr>
<td>3.2.2.1 Berliner, Manaoi and Monastersky's (1986) classification system</td>
<td>35</td>
</tr>
<tr>
<td>3.2.2.2 Johnson and Feldmeth's (1993) continuum</td>
<td>35</td>
</tr>
</tbody>
</table>
3.2.3 Core features of definitions

3.2.3.1 Characteristics of the sexual behaviour

3.2.3.2 Motives for the behaviour

3.2.3.3 Self-control

3.2.3.4 Emotions of the perpetrator

3.2.3.5 Perpetrator-victim relationship

3.2.4 Difficulties in defining sexual perpetration by young people

3.2.5 Summary

3.3 Prevalence

3.3.1 Introduction

3.3.2 Sources of variation in prevalence rates

3.3.2.1 Method of enquiry

3.3.2.2 Sampling procedures

3.3.2.3 Definition

3.3.3 Retrospective underreporting

3.3.3.1 Male victims and underreporting

3.3.4 Summary

3.4 The effects of child sexual abuse

3.4.1 Introduction

3.4.2 Psychological problems associated with child sexual abuse

3.4.3 Causal status of the effects

3.4.3.1 Association

3.4.3.2 Non-spuriousness

3.4.3.3 Temporal precedence

3.4.3.4 Causal mechanism
3.4.4 Summary 54
3.5 Strategies for the prevention of child sexual abuse 54
3.5.1 Introduction 54
3.5.2 Primary prevention of child sexual abuse 55
3.5.3 Secondary prevention of child sexual abuse 56
3.6 Summary 60

4 Causality and sexual perpetration 62
4.1 Introduction 62
4.2 Preliminary philosophical issues 63
4.2.1 Logical positivist andrealist conceptions of causality 63
4.2.2 Physicalist and constructivist views of mental causation 64

4.3 Causal criteria 66
4.3.1 Association 66
4.4.3.1 Deterministic vs. probabilistic associations 66
4.4.3.2 Statistical significance of an association 67
4.4.3.3 Size of the association 68
4.4.3.4 Summary 69
4.4.3.5 Factors that obscure an association 69
4.4.3.6 Designs to establish an association 72
4.3.2 Non-spuriousness 72
4.3.2.1 Retrospective reporting bias 76
4.3.3 Temporal precedence 80
4.3.4 Causal mechanism 81
4.4 Summary 84

5 Child sexual abuse as a risk factor for sexual perpetration 86
5.1 Introduction 86
6.9.1 Summary 117
6.10 Non-sexual antisocial behaviour 118
  6.10.1 Summary 119
6.11 Psychopathology 120
  6.11.1 Summary 121
6.12 Conclusion 122

7 Theories of sexual perpetration 124
7.1 Introduction 124
  7.1.1 Limits of review 125
  7.1.2 Structure of review 126
7.2 Feminist theories 126
7.3 Pornography theories 127
7.4 Psychoanalytic and psychodynamic theories 128
7.5 Family systems theories 131
7.6 Behavioural and learning theories 132
7.7 Cognitive theories 135
7.8 Attachment theory 136
7.9 Biological theories 138
7.10 Multifactor theories 139
  7.10.1 Finkelhor's Four-Factor Model 140
  7.10.2 Marshall and Barbaree's Integrated theory 141
  7.10.3 Hall and Hirschman's Quadripartite Theory 142
  7.10.4 Ward and Siegert's Pathways Model 143
  7.10.5 Developmental psychopathology 143
    7.10.5.1 Complex causal models 145
    7.10.5.2 Risk and protective factors 145
    7.10.5.3 Developmental psychopathology and sexual perpetration 145
    7.10.5.4 Multifactor theories and risky theoretical predictions 146
7.11 Summary 147
8 Summary and conclusion of literature review

9 Methodology: The catch-up longitudinal design

9.1 Introduction

9.2 The retrospective design

9.3 The longitudinal design

9.4 The catch-up longitudinal design

9.4.1 Advantages of the catch-up longitudinal design

9.4.2 Limitations of the catch-up longitudinal design

9.4.2.1 Random error

9.4.2.2 Measurement of risk factors

9.4.3 Generalisability

9.4.3.1 Random error

9.4.2.4 Researcher bias

9.4.2.5 Unidentified perpetration

9.4.2.6 Causal mechanism

9.5 Summary

10 Methodology: Actuarial prediction of sexual perpetration

10.1 Introduction

10.2 Actuarial predictions of sexual recidivism

10.3 Factors determining clinical applicability

10.3.1 Base rate of sexual perpetration

10.3.2 Utility decisions

10.3.2.1 Cost of a false alarm (false positive)

10.3.2.2 Cost of a miss (false negative)

10.3.2.3 Benefits of correct rejections (true negatives) and hits (true positives)

10.3.2.4 Balancing utilities

10.3.3 Treatment factors

10.3.3.1 Effectiveness of treatment
10.3.3.2 The proportion of the target population selected for treatment

10.3.3.3 Dose-response relationships in treatment

10.3.3.4 Assessing treatment factors

10.4 Summary

11 Method

11.1 Introduction

11.2 Aim and hypotheses

11.2.1 Hypotheses

11.3 Design

11.4 Ethical approval

11.5 Sample

11.5.1 Inclusion-exclusion criteria

11.5.1.1 Stage one

11.5.1.2 Stage two

11.5.1.3 Stage three

11.5.1.4 Stage four

11.5.1.5 Stage five

11.5.1.6 Stage six

11.5.1.7 Stage seven

11.5.2 Comparison of larger samples with final sample

11.5.3 Description of final sample

11.5.3.1 Demographic factors

11.5.3.2 Professional involvement

11.6 Instruments and measures

11.6.1 Development of the Risk Factor Manual

11.6.1.1 Definitions

11.6.1.2 Perpetrator criteria

11.6.1.3 Correct factor
11.6.1.4 Potential difficulties

11.6.1.5 Dichotomous measurement

11.6.1.6 Severity levels

11.6.1.7 Duration

11.6.2 The Computerised Extraction Program

11.6.2.1 Cover sheet

11.6.2.2 Quote screen

11.6.2.3 Review screen

11.6.2.4 Care table

11.6.2.5 Example

11.6.3 Definition of child sexual abuse

11.6.3.1 Sexual contact

11.6.3.2 Physical coercion

11.6.4 Definition of sexual perpetration

11.6.4.1 Criterion one

11.6.4.2 Criterion two

11.6.4.3 Criterion three

11.6.4.4 Determining the onset of perpetration

11.6.5 Measurement of the risk factors

11.7 Reliability

11.8 Structure of the study

11.8.1 Extraction of identifying information from clinical files

11.8.2 Search of criminal records

11.8.3 Tracing of social service files

11.8.4 Extraction and scoring of case material

11.7 Reliability

11.8 Structure of the study

11.8.1 Extraction of identifying information from clinical files

11.8.2 Search of criminal records

11.8.3 Tracing of social service files

11.8.4 Extraction and scoring of case material

12
11.9  Statistical analysis 209
  11.9.1  Data entry 209
  11.9.2  Data cleaning 209
  11.9.3  Assessing normality and homogeneity of variance 209
  11.9.4  Bivariate analysis of descriptive variables 210
  11.9.5  Bivariate analysis of risk factors 210
  11.9.6  Multivariate analysis of risk factors 211
  11.9.7  Construction of the Risk Indexes 211
    11.9.7.1  Construction of the Seven-Factor Dichotomous Risk Index 211
    11.9.7.2  Construction of the Seven-Factor Severity Risk Index 212
    11.9.7.3  Construction of shortened forms of the risk indexes 213
  11.9.8  Quantifying predictive validity 214
    11.9.8.1  Further analysis of the Seven-Factor Dichotomous Risk Index 216

11.10  Summary 220

12  Reliability 223
  12.1  Introduction 223
  12.2  Sources of potential unreliability 223
  12.3  Method 224
    12.3.1  Statistical analysis of inter-rater reliability 224
  12.4  Reliability of extracting case material 226
  12.5  Reliability of scoring case material 229
  12.6  Summary 231

13  Results 232
  13.1  Introduction 232
  13.2  Description of social service and clinical files 232
    13.2.1  Summary 234
13.8.5 Summary

13.9 Summary of results

14 Discussion

14.1 Introduction

14.2 Limitations

14.2.1 Introduction

14.2.2 Limitations of the catch-up design

14.2.2.1 Problems of using contemporaneous case material

14.2.2.2 Generalisability

14.2.2.3 Researcher bias in coding

14.2.3 Use of police records

14.2.4 Definition of sexual perpetration

14.2.5 Limitations of the risk indexes

14.2.6 Summary

14.3 Summary and interpretation of results

14.3.1 Summary

14.4 Methodological implications

14.4.1 Introduction

14.4.2 The catch-up longitudinal design

14.4.3 Computerised extraction program and Risk Factor Manual

14.4.4 Minimising researcher bias in the catch-up longitudinal design

14.4.5 The measurement of maltreatment

14.4.6 Summary

14.5 Theoretical considerations

14.5.1 Introduction

14.5.2 The importance of studying causal mechanisms

14.5.3 The importance of risky predictions

14.5.4 Testing causal mechanisms
List of tables

3.1 Legal definitions of sexual abuse (Glazebrook, 1997) 32
3.2 Effect sizes (r) for self-esteem, depression and other symptoms reported by Jumper (1995) 49
3.3 Effect sizes (r) and confidence intervals for the relationship between child sexual abuse and psychological outcomes reported by Rind, Tromovitch and Bauserman (1998) 50
5.1 Proportions of sexual perpetrators reporting sexual victimisation 88
5.2 Case-control studies examining sexual victimisation among sexual perpetrators 89
5.3 Cross-sectional studies examining sexual victimisation among sexual perpetrators 93
6.1 Proportions of sexual offenders reporting physical abuse 104
6.2 Non-sexual antisocial behaviour among sexual perpetrators 118
9.1 Differences between the real-time longitudinal design, the catch-up longitudinal design and the retrospective design 157
9.2 Comparison of contemporaneous and retrospective recall of dichotomous measures of maltreatment (Wolfe, McGee, Yuen, Wilson, & Camochan, 1995) 160
10.1 Predictive validity of sexual recidivism risk assessment instruments 168
10.2 Utility of four possible classification outcomes 170
11.1 Demographic information for victim-only and victim-perpetrator group 187
11.2 Professional involvement for victim-only and victim-perpetrator group 189
11.3 Definition of risk factors 199
11.4 Summary of risk factor measures 200
12.1 Extraction reliability for outcome and risk factor variables 228
12.2 Scoring reliability for each pairing of four researchers 230
13.1 Mean decimal years between birth date and end of usable information 233
13.2 Mean decimal years between start date and end of usable information 234
13.3 Summary of the perpetration characteristics of the victim-perpetrator group 237
13.4 Summary of child sexual abuse characteristics of the victim-only and victim-perpetrator groups 238
13.5 Summary of physical abuse characteristics of the victim-only and victim-perpetrator groups 239
13.6 Summary of witnessing intrafamilial physical abuse characteristics of the victim-only and victim-perpetrator groups 240
13.7 Summary of neglect (failure to provide) characteristics of the victim-only and victim-perpetrator groups 241
13.8 Summary of neglect (lack of supervision) characteristics of the victim-only and victim-perpetrator groups 242
13.9 Summary of rejection by carer characteristics of the victim-only and victim-perpetrator groups 243
13.10 Summary of care history characteristics of the victim-only and victim-perpetrator groups 243
13.11 Logistic regression examining four dichotomous risk factors 245
13.12 Adjusted odds ratios and 95% CI from logistic regression examining dichotomous risk factors 245
13.13 Logistic regression examining severity and duration measures of the risk factors 246
13.14 Adjusted odds ratios and 95% CI from logistic regression examining severity and duration measures of risk factors 247
13.15 Sensitivity, specificity, positive predictive values and negative predictive value of the Seven-Factor Dichotomous Risk Index

13.16 Sensitivity, specificity, positive predictive value and negative predictive value of the Seven-Factor Severity Risk Index

13.17 Sensitivity, specificity, positive predictive value and negative predictive value of the Four-Factor Dichotomous Risk Index

13.18 Sensitivity, specificity, positive predictive value and negative predictive value of the Four-Factor Severity Risk Index

13.19 Decision analysis: True positive given twice the weighting of other utilities (Base rate 3%)

13.20 Decision analysis: True negatives given twice the weighting of the other utilities (Base rate 3%)

13.21 Treatment efficacy of 25% (Base rate 3%)

13.22 Treatment efficacy of 50% (Base rate 3%)

13.23 Treatment efficacy of 75% (Base rate 3%)

13.24 Decision analysis: True positive given twice the weighting of other utilities (Base rate 10%)

13.25 Decision analysis: True negatives given twice the weighting of other utilities (Base rate 10%)

13.26 Treatment efficacy of 25% (Base rate 10%)

13.27 Treatment efficacy of 50% (Base rate 10%)

13.28 Treatment efficacy of 75% (Base rate 10%)

13.29 Decision analysis: True positives given twice the weighting of other utilities (Base rate 17%)

13.30 Decision analysis: True negatives given twice the weighting of other utilities (Base rate 17%)

13.31 Treatment efficacy of 25% (Base rate 17%)

13.32 Treatment efficacy of 50% (Base rate 17%)

13.33 Treatment efficacy of 75% (Base rate 17%)
### List of figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Finkelhor’s Four-Factor Model</td>
<td>140</td>
</tr>
<tr>
<td>7.2</td>
<td>Marshall and Barbaree’s Integrated Theory</td>
<td>141</td>
</tr>
<tr>
<td>7.3</td>
<td>Hall and Hirschman’s Quadripartite Model</td>
<td>143</td>
</tr>
<tr>
<td>11.1</td>
<td>Design of the study</td>
<td>179</td>
</tr>
<tr>
<td>11.2</td>
<td>Inclusion-exclusion criteria</td>
<td>184</td>
</tr>
<tr>
<td>11.3</td>
<td>ROC-curve with 50% of the area under the line</td>
<td>215</td>
</tr>
<tr>
<td>11.4</td>
<td>ROC-curve indicating prediction better than chance</td>
<td>215</td>
</tr>
<tr>
<td>11.5</td>
<td>Equation for decision analysis</td>
<td>217</td>
</tr>
<tr>
<td>11.6</td>
<td>Equation for proportion treated</td>
<td>218</td>
</tr>
<tr>
<td>11.7</td>
<td>Equation for proportion perpetrating if treatment population selected at random</td>
<td>218</td>
</tr>
<tr>
<td>11.8</td>
<td>Equation for proportion perpetrating if risk index used to select treatment population</td>
<td>219</td>
</tr>
<tr>
<td>11.9</td>
<td>Equation for improvement of selection using risk index over selection at random</td>
<td>220</td>
</tr>
<tr>
<td>11.10</td>
<td>Structure of the study</td>
<td>222</td>
</tr>
<tr>
<td>13.1</td>
<td>Distribution of Seven-Factor Dichotomous Risk Index scores for victim-only and victim-perpetrator group</td>
<td>249</td>
</tr>
<tr>
<td>13.2</td>
<td>ROC-curve for Seven-Factor Dichotomous Risk Index</td>
<td>250</td>
</tr>
<tr>
<td>13.3</td>
<td>Distribution of Seven-Factor Severity Risk Index scores for victim-only and victim-perpetrator group</td>
<td>251</td>
</tr>
<tr>
<td>13.4</td>
<td>ROC-curve for Seven-Factor Severity Risk Index</td>
<td>253</td>
</tr>
<tr>
<td>13.5</td>
<td>Distribution of Four-Factor Dichotomous Risk Index scores for victim-only and victim-perpetrator group</td>
<td>255</td>
</tr>
<tr>
<td>13.6</td>
<td>ROC-curve for Four-Factor Dichotomous Risk Index</td>
<td>256</td>
</tr>
<tr>
<td>13.7</td>
<td>Distribution of Four-Factor Severity Risk Index scores for victim-only and victim-perpetrator group</td>
<td>257</td>
</tr>
<tr>
<td>13.8</td>
<td>ROC-curve for Four-Factor Severity Risk Index</td>
<td>258</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Introduction

Child sexual abuse has become a central concern of policy makers, mental health practitioners and society as a whole. One of the reasons for this concern is the recognition that the experience is associated with a range of subsequent psychological problems. The hypothesis that today's victim of sexual abuse may be at increased risk of becoming tomorrow's sexual perpetrator is of particular concern. The status of this cycle-of-victimisation hypothesis is controversial, but at least two points are clear. There is no simple one-to-one mapping between victimisation and perpetration; by no means every victim of sexual abuse becomes a perpetrator. It is also apparent that if victimisation increases the risk of perpetration, this risk is likely to be greater for males than females.

Given these two points it may be useful to identify the additional risk factors for abusive behaviour among male children who have been sexually abused. The search for these additional risk factors is the central aim of this thesis, and it addresses this aim by dividing it into two questions. First, what is the causal status of these additional risk factors? Secondly, is it possible to use these risk factors to predict which sexually victimised males are at most risk?

1.2 Overview of the thesis

Chapters 2 to 8 review the literature relating to these questions. The methodology section, chapters 9 to 12, opens with a discussion of the catch-up longitudinal design. This design is seldom used in psychological research, but, as the chapter will

---

1 A summary of the author’s independent contribution to the research reported in this thesis is given in appendix 1.
argue, it is well placed to examine the two questions (cause and prediction) addressed in this thesis. In this study the catch-up design involves using contemporaneous clinical and social service case material to code information on sexual victimisation and additional risk factors. Sexual perpetration is coded from this case material, as well as police caution and conviction records. As chapter 9 argues, a crucial feature of this design is that information on risk factors is coded entirely on the basis of information gathered before the onset of sexual perpetration. The use of information coded prior to the outcome helps us to draw causal inferences from the catch-up design. As with the real-time longitudinal design, temporal order can be established and the outcome cannot affect the coding of the risk factor. The question of prediction is also well served by this methodological feature. If we use this design, we can examine whether it is possible to identify those sexually victimised males that are at risk of perpetration before the onset of this behaviour. Chapter 10 examines in greater detail the methodological issues involved in assessing the predictive capacity of the additional risk factors. Chapter 11 outlines the method used in the study and chapter 12 provides details of reliability estimates.

The results section, chapter 13, examines separately the questions of cause and prediction. Odds ratios and logistic regression are used to investigate the causal status of the risk factors. The issue of prediction is assessed using a number of techniques including ROC-curves and a decision analysis.

The final chapter discusses the results of the study, and covers areas such as the methodological, theoretical, and policy and practice implications of the results.

1.3 Summary

The aim of this thesis is to examine the additional risk factors for sexual perpetration among males who have experienced sexual victimisation. This is divided into two separate question: what is the causal status of these risk factors and can the risk
factors predict which sexually victimised males are at increased risk of perpetration? The thesis uses a catch-up longitudinal because it is well placed to investigate these two questions.
Introduction to literature review

2.1 Introduction

The vast literature on the sexual victimisation of children is one indicator of the concern evoked by child sexual abuse. There is, however, little consistency in the findings of research studies, and many of the studies lack methodological sophistication. This literature review assesses the current status of research findings that are relevant to the questions examined in this thesis.

2.2 Limits of the review

The research discussed in this thesis examines risk factors for sexually abusive behaviour in which a child is the victim. It is assumed that the aetiology of sexual perpetration against children differs from that against adults, and therefore the review will focus on research examining the abuse of children. Research focussing exclusively on perpetration against adults (rape) will not be reviewed. Many studies do not report separate results for the two types of offenders; a decision was to include these studies in the review. There are a large number of these mixed studies, and if they were excluded this would omit a substantial part of the literature. The mixed groups typically contain a high proportion of perpetrators against children, and this is another reason for including them. These studies will be clearly distinguished from research focussing exclusively on sexual abusers of children.

2.3 Overview of the review

In a study focussing on the relationship between sexual victimisation of children and sexual perpetration against children the definition of the terms victimisation and
perpetration is important. It has proved difficult to develop an adequate definition of these terms, and the literature review opens by discussing these difficulties.

A basic assumption of the thesis is that childhood sexual abuse is a serious societal problem, and this is addressed next in the literature review by examining prevalence rates and the effects of sexual abuse. A remarkably wide range of prevalence rates have been reported, but even the most conservative estimates suggest a substantial proportion of both males and females are sexually victimised during childhood. These experiences have been linked to a wide range of mental health problems, but the methodology of much of the research is weak. There is, however, sufficient evidence of an association between sexual victimisation and a number of mental health problems; there is also some limited evidence that the association cannot be attributed to potential confounds.

The prevalence of childhood sexual abuse and the possible long-term effects suggest a need for effective prevention strategies. Currently two strategies are widely used, and these are described and evaluated. Primary prevention strategies aim to teach children how to avoid potentially dangerous situations; secondary prevention strategies attempt to decrease the risk of re-offending among identified sexual abusers. The available research raises serious questions about the effectiveness of both of these strategies, and secondary strategies are inherently limited in that they focus on people who have already victimised children. These limitations suggest there is a need to focus on potential perpetrators before the onset of abusive behaviour.

The success of this approach is dependent on researchers identifying those factors that are predictive of offending; this task is made easier for us if we understand the causes of sexual perpetration. The need to establish a causal relationship is therefore central to this thesis. Chapter 4 outlines the criteria that need to be met to demonstrate causality.
A first step in the identification of causal antecedents is to identify risk factors for sexual perpetration. A risk factor can be defined as a characteristic, experience or event that, if present, is associated with an increase in the probability (risk) of a particular outcome over the base rate of the outcome in the general (unexposed) population (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997). Two chapters (chapters 5 and 6) of the literature review discuss putative risk factors for sexually abusive behaviour.

Chapter 5 examines the experience of sexual victimisation as a risk factor for perpetration. This thesis assumes that the sexual victimisation of males increases the risk of sexual perpetration, but that other risk factors are also involved. The possibility of prevention is increased if these additional risks can be identified, because sexually victimised males are routinely seen by social workers and mental health professionals. In the literature review, the evidence for the victim-to-perpetrator link among males will be evaluated. While the available research has not convincingly demonstrated a causal relationship, there is some evidence of an association between the two, and some limited evidence that this association cannot be attributed to other variables. There is, however, evidence that if there is a link between victimisation and perpetration it is likely to be stronger for male than female victims. The available evidence also supports the hypothesis that sexual victimisation is unlikely to be a sufficient cause of sexual victimisation. This suggests that other risk factors further increase the risk. The evidence for a causal relationship between the additional risk factors and sexual perpetration is reviewed in chapter 6. In most cases the evidence for these risk factors is inconclusive, which suggests there is a need to conduct further research on these putative risks.

The majority of research discussed in chapters 5 and 6 has used a retrospective design, and this has seriously limited the capacity to identify causal relationships. In particular a retrospective design has difficulty in addressing the causal criteria of non-spuriousness and temporal precedence. It will be argued that there is an urgent need to assess these risk factors in a longitudinal design. It will also be argued that
there are considerable ethical difficulties in conducting a longitudinal study to examine these risk factors, but the ‘catch-up’ longitudinal design does offer a potential solution.

Chapter 7 reviews the available theories of sexual perpetration against children. These theories provide hypotheses about the mechanism linking risk factors to sexual perpetration, and, in doing so, address the final causal criterion – the demonstration of a causal mechanism. This chapter will argue that these theories are limited because the evaluation of their hypotheses has relied on null hypothesis significance testing, which can provide only weak corroboration of a theory.
This chapter reviews the literature on the prevalence and long-term effects of child sexual abuse to argue that sexual abuse is a serious societal problem and that there is a need to identify risk factors for this behaviour. In the section on prevalence it is argued that methodologically robust studies suggest a substantial proportion of both males and females are sexually abused during childhood. These experiences have been linked to a range of mental health problems and the evidence for this link is examined in the section on long-term effects. Although few of the causal criteria have been successfully examined, there is sufficient evidence to conclude that there could be a relationship between sexual victimisation and later psychological problems.

In the next section the prevention of sexual abuse is discussed. It is argued that existing prevention strategies are limited, and that there is a need to develop strategies that focus on potential perpetrators before the onset of perpetration. This requires the identification of risk factors for sexually abusive behaviour.

It is first necessary to examine the definition of child sexual abuse and sexual perpetration. The definition of these concepts is central to this study, but there is some disagreement about how these terms should be defined. The next section will examine this disagreement.

3.1 Definitions of child sexual abuse

There are three main types of definitions of child sexual abuse and sexual perpetration: those used by researchers, those used by the legal profession and those used in social work. Each of these definitions reflects the particular purpose for
which it has been developed. The definitions used by these three professions are described in more detail below.

### 3.1.1 Research definitions

Research definitions of child sexual abuse are designed to help answer a particular research question, such as the prevalence rate or the long-term effects of abusive experiences.

Different researchers have used very different definitions of child sexual abuse (Haugaard, 2000; Mannon & Leitschuh, 2002). Ghate and Spencer (1995) suggest that the definitions used by researchers vary according to five assumptions; Goldman and Padayachi (2000) propose a similar list. The assumptions discussed by Ghate and Spencer (1995) include: the type of activity involved, the age of the victim, the age differential between victim and perpetrator, the nature of the relationship between victim and perpetrator, and issues of consent and power.

#### 3.1.1.1 The type of activity involved

Some definitions of child sexual abuse include as abusive non-contact behaviours such as exhibitionism and voyeurism (e.g. Finkelhor, 1984), whereas others include only those behaviours involving physical contact between the victim and perpetrator (e.g. Briere & Runtz, 1988). The exclusion of non-contact behaviours is presumably based on the assumption that these experiences, while improper, are not psychologically harmful, and that their inclusion might mask a relationship between sexual abuse and long-term psychological effects. Those researchers who have included these behaviours presumably do so on the basis that the question of whether they are harmful or not is one that requires empirical resolution.

#### 3.1.1.2 The age of the victim

The upper threshold of childhood is often specified in definitions of child sexual abuse. The victim must be below this threshold at the time of the experience for it to
be considered abusive. The majority of researchers specify the age of the victim in years, with 16 years considered the cut off point (Ghate & Spencer, 1995). Occasionally the threshold is determined by developmental stage, such as pubertal status (e.g. Fritz, Stoll, & Wagner, 1981).

3.1.1.3 Victim and perpetrator age differential
Definitions of child sexual abuse often specify that the perpetrator of the abuse must be older than the victim by a specified number of years. A commonly used criterion is an age difference of five or more years (Watkins & Bentovim, 1992); other researchers have used a difference of at least two years (e.g. Johnson, 1988). In some cases the age differential varies according to the age of the victim or perpetrator at the time of the event. Cantwell (1988) dispensed with an age difference altogether and focused exclusively on the nature of the behaviour.

There is no theoretical justification for a definition that excludes non-consensual sex with a same age peer, and yet the reliance on age differentials has led most definitions to exclude these experiences (Roosa, Reyes, Reinholtz, & Angelini, 1998). One advantage of an age differential approach is that it provides a clear, simple operationalisation of a power relationship, but it does seem necessary to include non-consensual sex with someone who is close in age.

3.1.1.4 Perpetrator-victim relationship
The historical emphasis on incest, especially that occurring between father and daughter, has led some definitions to focus on this type of abuse or, more broadly, on intrafamilial abuse. Most recent definitions encompass any abusive relationship regardless of victim and perpetrator gender or the nature of a familial relationship.

3.1.1.5 Consent and power
Issues such as the degree of consent and the abuse of power relationships are the most controversial elements of any definition of child sexual abuse. Some definitions require a lack of consent to be present if the experience is to be
considered abusive (e.g. Bagley, Wood, & Young, 1994). However, others argue that children do not have the capacity to give consent for sexual experiences, and the definition must focus on the abuse of a power relationship between a more and a less mature person (e.g. Kelly, Regan, & Burton, 1991). The exclusion of consensual acts might be based on the assumption that such experiences are not psychologically harmful, while their inclusion reflects an assumption that an act may be harmful regardless of consent at the time.

3.1.2 Legal definitions

The purposes of legal definitions of child sexual abuse are twofold. They are used to classify a person as perpetrator, which can lead to the impositions of sanctions. The definitions also serve to classify a child as victimised, which will determine professional response to the child, such as removal into care. Table 3.1 summarises the main categories of sexual offending embodied in the Criminal Law of England and Wales (Glazebrook, 1997).

3.1.3 Social Work definitions

Definitions used in social work are designed to determine professional responses to suspected abuse. These responses can include the registration of a child on the Child Protection Register, the instigation of legal proceedings and the removal of the child from the home. The definitions of child sexual abuse used in social work in the UK differ between departments, but all currently derive from the definition provided by Working Together Under the Children Act (Home Office, Department of Health, Department of Education and Science, Welsh Office, 1991): ‘Actual or likely sexual exploitation of a child or adolescent. The child may be dependent and / or developmentally immature.’
Table 3.1 – Legal definitions of sexual abuse (Glazebrook, 1997)

<table>
<thead>
<tr>
<th>Sexual crime (Year of Act)</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rape (1956)</td>
<td>It is an offence for a man to rape a woman or another man(^a). A man commits rape if: he has sexual intercourse(^b) with a person (whether vaginal or anal) who at the time of the intercourse does not consent; at the time he knows that the person does not consent to the intercourse or is reckless as to whether that person consents to it.</td>
</tr>
<tr>
<td>Buggery (1956)</td>
<td>It is an offence for a person to commit buggery with another person otherwise than in the circumstances described in subsection 1A... (1A) The circumstances referred to in subsection (1) are that the act of buggery takes place in private and both parties have obtained the age of eighteen.</td>
</tr>
<tr>
<td>Intercourse with girl under 13 (1956)</td>
<td>It is a felony for a man to have unlawful sexual intercourse(^b) with a girl under the age of thirteen.</td>
</tr>
<tr>
<td>Intercourse with girl under 16 (1956)</td>
<td>It is an offence... for a man to have unlawful sexual intercourse(^b) with a girl... under the age of sixteen.</td>
</tr>
<tr>
<td>Indecent assault on a woman (1956)</td>
<td>It is an offence... for a person to make an indecent assault(^c) on a woman. A girl under the age of sixteen cannot in law give any consent that would prevent an act from being an assault...</td>
</tr>
<tr>
<td>Indecent assault on a man (1956)</td>
<td>It is an offence... for a person to make an indecent assault(^c) on a man. A boy under the age of sixteen cannot in law give any consent that would prevent an act from being an assault...</td>
</tr>
<tr>
<td>Indecent conduct towards a young child (1960)</td>
<td>Any person who commits an act of gross indecency with or towards child under the age of fourteen, or who incites a child that age to such an act with him or another shall be liable on conviction on indictment to imprisonment...</td>
</tr>
<tr>
<td>Incest by a man (1956)</td>
<td>It is an offence for a man to have sexual intercourse(^b) with a woman whom he knows to be his granddaughter, daughter, sister, or mother.</td>
</tr>
<tr>
<td>Incest by a woman (1956)</td>
<td>It is an offence for a woman of the age of sixteen or over to permit a man whom she knows to be her grandfather, father, brother, or son to have sexual intercourse(^b) with her by her consent.</td>
</tr>
</tbody>
</table>

\(^a\) For the victim, the term man and woman apply regardless of age (1956).

\(^b\) Sexual intercourse is defined as the penetration of the penis into the vagina or anus; ejaculation is not required; the offence is committed as soon as penetration takes place (1956).

\(^c\) Acts of indecent assault range from unwanted kisses to forced oral sex, but exclude vaginal and anal intercourse (Ashworth, 1991).
3.1.4 Discussion

A definition of maltreatment, such as child sexual abuse, can be based on one of two conceptual premises (Barnett, Manly, & Cicchetti, 1991; Haugaard, 1991). A Developmental Psychopathological approach defines child sexual abuse as those acts that are demonstrably harmful to a child, for example, experiences associated with long-term psychological difficulties. In contrast a sociological approach basis its definition on a societal consensus that an act is improper (Haugaard, 1991).

Social work and the legal profession use a societal consensus definition. This is appropriate for these disciplines, because they are social institutions designed to reflect and enforce societal values.

For a research definition the Developmental Psychopathological approach, based on the concept of harm, is to be preferred. The objective of childhood sexual abuse research is to identify the causes and effects of this behaviour. This is hampered by a sociological definition because societal consensus rather than knowledge about the harm of the experience may obscure these causes and effects (Rind, Tromovitch, & Bauserman, 1998). Take the hypothetical example that non-contact experiences do not increase the probability of mental health problems, but contact abuse does. If people who have been exposed to non-contact experiences were classified as sexually abused, this would attenuate the association between abuse and mental health problems.

Current research definitions of child sexual abuse are based on a societal consensus that an act is improper, rather than demonstrated harm, but this is because there is a lack of methodologically robust research on the long-term effects of child sexual abuse. A developmental psychopathology approach, using the concept of harm, will involve the constant refinement of a definition as research on the effects of maltreatment accumulates. Specifically there is a need for studies that systematically vary each part of a definition to examine the relationship between it and an outcome. This method will help identify those features of a behaviour that are harmful.
In the absence of this research current definitions, including the one used in this thesis, must rely on a societal consensus approach. Chapter 11 provides further information about the operational definition of child sexual abuse used in this study.

3.1.5 Summary

There are three types of definitions of child sexual abuse: research, legal and social work definitions. All are currently based on a societal consensus that an act is improper, rather than knowledge that the act is harmful for the victim. There is a need for research definitions to move beyond a societal consensus approach to a definition based on harm.

3.2 Definition of sexual perpetration by young people

3.2.1 Introduction

While there are disagreements about how to define child sexual abuse when the perpetrator is an adult, the definition of sexual perpetration by a young person introduces yet further problems. The main difficulty is the need to differentiate between sexual perpetration, normal sexual behaviour and forms of problematic sexual behaviour that are not sexual perpetration. This section will outline a number of approaches to this problem.

3.2.2 Definitions of sexual behaviour in young people

There is no shortage of definitions of problematic sexual behaviour by young people (e.g. Berliner, Manaois, & Monastersky, 1986; Cunningham & McFarlane, 1991; 1996; Crisci & Brown, cited in Kikuchi, 1995; Friedrich, 1990; Gil, 1993; Johnson
& Feldmeth, 1993; Pithers, Gray, Cunningham, & Lane, 1993; Rasmussen, Burton, & Christopherson, 1992; Ryan & Blum, 1994; Sgroi, Bunk, & Wabrek, 1988). Some of these definitions will be outlined and the points of similarities will then be given, based on Araji’s (1997) summary.

3.2.2.1 Berliner, Manaoi and Monastersky’s (1986) classification system

Berliner, Manaoi and Monastersky (1986) provide one of the earliest attempts at defining problematic sexual behaviour in young people. They use the term ‘disturbance’ to reflect sexual behaviours that deviate from those expected as part of normal sexual development. The authors do not provide a definition of normal, but do discuss three types of non-normal sexual behaviour.

Inappropriate sexual behaviour is categorised as the least disturbed. It includes behaviours such as persistent masturbation, touching the breasts and genitals of others, and excessive interest in sexual material. It is viewed as disturbed only if it interferes with the child’s development, the behaviour remains despite intervention, multiple sexual behaviours are reported, and the behaviours are accompanied by other forms of behaviour such as conduct disorder. The next level of behaviour is termed developmentally precocious. This includes simulated or completed intercourse. The behaviours are explicit and intentional, but no coercion is present. The final level consists of those behaviours involving aggression or coercion. Aggressive sexual contact is defined as physical force, involving injury. Coercion may involve either threats or social coercion.

3.2.2.2 Johnson and Feldmeth’s (1993) continuum

Johnson and Feldmeth (1993) propose a continuum of sexual behaviours, including normal, sexually reactive, extensive mutual sexual behaviours and molestation. The behaviour is considered normal if:

- It is viewed as information gathering or curiosity
- The children are similar in age, size and developmental stage
- Both or all children engage in the behaviour voluntarily
• The children have an amicable relationship outside the sexual interaction
• The behaviours are limited in terms of frequency and type, and the child is not preoccupied with sexual behaviour
• If the child is told to stop they do so
• The behaviour is spontaneous and light-hearted.

Sexually reactive is defined as behaviours that deviate from normal, as defined above. Typically children in this group have an increased focus on sexuality and exhibit a wider range of sexual behaviours compared to children in the normal group.

The next group of children are those exhibiting extensive mutual sexual behaviours. Children in this group engage in adult-like sexual behaviours with other children. A key feature of this group is a lack of affect about their sexual behaviours; they exhibit a blasé, ‘matter-of-fact’ attitude. Johnson and Feldmeth (1993) point out that many of these children use sex as a means to an end to help them cope with feelings of loss or abandonment.

The final group are defined as children who molest. Johnson and Feldmeth (1993) suggest that the sexual behaviour of children in this group has an impulsive and compulsive quality. Some form of coercion is always present and could include physical force or social and emotional threats. The motive for this behaviour, according to the authors, is to reduce negative feelings of fear, anger or loneliness that have become associated with sexuality.

3.2.2.3 Rasmussen, Burton and Christopherson’s (1992) definition
Rasmussen, Burton and Christopherson (1992) use a number of criteria to differentiate appropriate and inappropriate sexual behaviours, these include:
• The power differential between perpetrator and victim (perpetrator has greater age, size or mental capacity)
• Role differential (the perpetrator is in a position of authority over the child, e.g. babysitter)
• The perpetrator exhibits predatory patterns (perpetrator sets up the victim, e.g. grooming)
• Elements of coercion (e.g. tricks, bribes, threats and physical force).

Matsuda and Rasmussen (1990) use this definition to identify two types of children, aged 12 or younger, who display sexually inappropriate behaviour. The sexually reactive child is defined as a child aged 8 or younger who display inappropriate sexual behaviour towards another that is harmful or unlawful. The pre-adolescent sex offender is aged between 9 and 12, and display sexually inappropriate behaviour towards another that is harmful or unlawful.

3.2.2.4 Ryan and Blum’s (1994) typology
Ryan and Blum (1994) use the concept of problematic sexual behaviours. Behaviours are defined as problematic in three circumstances. In the first the behaviour is a problem for the child exhibiting the behaviour. For example, the behaviour could put him or her at risk, it could interfere with other developmental tasks or relationships, it could be self-abusive or the child could believe it is a problem. A second circumstance is when the behaviour is a problem for others. The behaviour could make others feel uncomfortable, it could occur at the wrong time or in the wrong place and it could conflict with family or community values. Finally the behaviour could be problematic for the child to whom the behaviour is directed. For example, it could involve other children without their consent, one child could be more powerful than the other or one child could be coerced by the other.

3.2.3 Core features of definitions
Araji (1997) provides a useful summary of the features of sexual perpetration (as opposed to other forms of sexually problematic behaviours) that were consistently identified by 11 definitions, including the ones reviewed above (Berliner, Manaouis,
& Monastersky, 1986; Cunningham & MacFarlane, 1991, 1996; Crisci & Brown, cited in Kikuchi, 1995; Friedrich, 1990; Gil, 1993; Hindman, 1994; Johnson & Feldmeth, 1993; Pithers, Gray, Cunningham, & Lane, 1993; Rasmussen, Burton & Christopherson, 1992; Ryan & Blum, 1994; Sgroi, Bunk, & Wabrek, 1988). The similarities are grouped into a number of categories, including: the characteristics of the sexual behaviour, the motive for the behaviour, self-control, emotions demonstrated by the perpetrator, abuse histories, abuser-victim relationships (equality, power and control), environments, and treatment outcomes.

3.2.3.1 Characteristics of the sexual behaviour
Araji (1997) suggests that the definitions agree upon a number of key features of the sexual behaviour characteristics. All definitions agree that the behaviours are far beyond those expected for the age of the abuser. The behaviours include oral copulation, vaginal and oral intercourse, and forcible penetration of anus or vagina with fingers or other objects. There is also agreement that the behaviours have an aggressive quality and involve the use of force, coercion, secrecy or all three. Other areas of agreement are that the behaviours increase over time, represent a pattern rather than an isolated event and have an obsessive and compulsive quality. Finally the sexual behaviours are planned, calculated and predatory.

3.2.3.2 Motives for the behaviour
Not all definitions refer to the motives for sexually aggressive behaviour. Those that do suggest the motive is either to achieve power or to reduce feelings of fear, anger or loneliness.

3.2.3.3 Self-control
A number of the definitions suggest that sexual perpetration by young people is characterised by a lack of self-control.
3.2.3.4 Emotions of the perpetrator
The definitions suggest that perpetrators exhibit feelings of anger, rage, fear, shame and loneliness. In addition the perpetrators lack empathy for their victims.

3.2.3.5 Perpetrator-victim relationship
The definitions also agree upon characteristics of the perpetrator-victim relationship. According to these definitions, perpetrators select victims who have less power and control than themselves. Defining characteristics include differences in age (2-5 years), size, status, intelligence, cognitive development and physical ability.

3.2.4 Difficulties in defining sexual perpetration by young people
As with the definition of child sexual abuse, the conceptual approach to the operational definitions of sexual perpetration by young people can be based on one of two assumptions. Perpetration can be based on a societal consensus definition that an act is improper; alternatively it can be based on a developmental psychopathology definition that an act is harmful to the victim. As argued in the previous section, a developmental psychopathology definition is to be preferred for research purposes. However, current research must rely on a social consensus definition, because the research needed to establish a developmental psychopathology definition has not been conducted as yet.

The literature suggests a number of additional approaches to defining problematic behaviour by young people. Problematic sexual behaviour can be defined as those behaviours that deviate from normal sexual development, but which may or may not be sexual perpetration. For example, Johnson and Feldmeth's (1993) approach provides a definition of normal sexual behaviours and defines problematic sexual behaviours as those that lay outside this domain. Other conceptual approaches include defining problematic sexual behaviour as those indicating that the young person has a psychological problem. As Araji (1997) notes, the Berliner, Manaoi and Monastersky (1986) definition emphasises a close relationship between sexually
disturbed behaviours and psychological disturbance in the person exhibiting those behaviours.

The conceptual approaches to defining these types of problematic sexual behaviours by young people are as underdeveloped as the definitions of sexual perpetration. For example, a definition based on behaviours that depart from normal sexual development necessarily requires a concept of normal sexual development. However, what constitutes normal sexual development is by no means clear (Araji, 1997). Definitions based on those behaviours that are indicative of psychological difficulties are also problematic. This is because few studies have examined the relationship between particularly sexual behaviours and psychological difficulties.

The situation is not promising. To define sexual perpetration by young people it is necessary to differentiate vague social consensus definitions of sexual perpetration from underdeveloped, under researched definitions of other types of problematic sexual behaviour. This thesis uses a social consensus definition of sexual perpetration, but the distinction between this and other types of problematic sexual behaviour may be less than perfect. The definition used in this research derives from a combination of the definitions already described and the discussion of similarities in the definitions given by Araji (1997). A full description of the operational definition of sexual perpetration by young people is given in chapter 11.

3.2.5 Summary

There are additional difficulties in defining sexual perpetration against children when the perpetrator is a young person. A particular difficulty is the need to differentiate sexual perpetration from other forms of sexually problematic behaviour that are not considered abusive.
This section has provided an overview of the definitions of child sexual abuse; it is now possible to use these definitions to assess the prevalence and long-term effects of these experiences.

3.3 Prevalence

3.3.1 Introduction

In an early review of North American research, Peters, Wyatt and Finkelhor (1986) found that prevalence rates for child sexual abuse varied from 3% to 31% for males and 6 to 62% for females. This led them to conclude that there is “...not yet any consensus among social scientists about the national scope of sexual abuse”. This conclusion still seems to hold: two subsequent reviews (Dhaliwal, Gauzas, Antonowicz, & Ross, 1996; Goldman & Padayachi, 2000), for example, report similar variation. Dhaliwal, Gauzas, Antonowicz and Ross (1996) found figures ranging between 3 and 37% for males, and 7 and 53% for females. Goldman and Padayachi (2000) report figures ranging from 4 to 30% for males and 7 to 62% for females.

Even among the handful of studies conducted in the UK (Baker & Duncan, 1985; Bifulco, Brown, & Adler, 1991; Oaksford & Frude, 2001; Nash & West, 1985; Kelly, Regan, & Burton, 1991) there is a lack of agreement. Two of the studies (Baker & Duncan, 1985; Kelly, Regan, & Burton, 1991) provide prevalence rates for males. Baker and Duncan (1985) report a rate of 9%; Kelly, Regan and Burton (1991) report figures of 7% using a narrow definition and 27% using a broad definition. All five studies provide prevalence rates for females, with figures ranging from 9% (Bifulco, Brown, & Adler, 1991) to 59% (broad definition of Kelly, Regan, & Burton, 1991).
3.3.2 Sources of variation in prevalence rates

A number of methodological differences between the studies are likely sources of the discrepancies (Goldman & Padayachi, 2000; Peters, Wyatt, & Finkelhor, 1986); these include the method of enquiry, the definition used and the sampling procedure.

3.3.2.1 Method of enquiry

Features of the method of enquiry that may affect the prevalence rate include the number of questions asked and the survey format (e.g. telephone interview, face-to-face interview, self-completion questionnaire) (Goldman & Padayachi, 2000; Peters, Wyatt, & Finkelhor, 1986). Several narrative reviews have drawn conclusions about the relationship between these and prevalence rates (e.g. Wyatt & Peter, 1986), but these reviews assessed research studies that differed on dimensions other than these features, which makes it difficult to assess the independent effect of a particular methodological characteristic. To avoid this difficulty one review used a multiple regression to assess the independent contribution of a study feature to the prevalence rates (Gorey & Leslie, 1997). These authors found that the interview method (coded as: face-to-face, mail survey or telephone survey) did not make a significant contribution to the prevalence rate once the definition and response rate had been controlled for.

3.3.2.2 Sampling procedures

Any specific population, whether classified by age, gender, ethnicity or socio-economic status, may have a prevalence rate of child sexual abuse different to the population as a whole. If the findings observed for a sample are to be generalised to the wider population, then the sample must mirror the characteristic of that wider population on any features that may affect the prevalence rate. Data from national probability samples are useful here because these are selected to match the characteristics of the general population. Rind and Tromovitch (1997) discuss the prevalence estimates from seven national probability samples, and report figures ranging from 6% to 36% for males and 14% to 53% for females.
Response biases may also affect the prevalence rate. In a sample the prevalence rate may vary between those people who do and do not respond. If this were so then the prevalence rate based on the respondent sample would not give an accurate representation of the true prevalence rate. Poor response rates are common in sexual abuse prevalence research; Gorey and Leslie (1997), for example, found that 12 out of 25 studies had response rates of below 60%.

Peters, Wyatt and Finkelhor (1986) identify two opposing hypotheses about the affect of non-response. The first maintains that victims of sexual abuse are less likely than non-victims to respond, because of the emotional difficulty of discussing the experience, which leads to artificially low prevalence rates. The second states that victims are more likely to respond, because recounting their experience is cathartic, which leads to artificially high prevalence rates. The available empirical evidence supports the hypothesis that victims of sexual abuse are more likely to take part in prevalence studies. Gorey and Leslie (1997) found that higher response rates were associated with lower prevalence rates at a statistically significant level (p < 0.05) in the female samples reviewed, and a level approaching significance in the male samples (p < 0.10). This suggests that when response rates are low those who have experienced sexual abuse are overrepresented in the respondent sample.

Given the possible bias introduced by low response rates, it may be useful to focus on the national probability samples with adequate rates of response. Of the seven studies discussed by Rind and Tromovitch (1997), three had response rates of 80% or more. Badgley et al. (1984) attained a response rate of 94% in a study of Canadians aged over 18 years of age. This study found rates of 31% for males (N = 1002) and 53% for females (N = 1006). The British national probability sample (Baker & Duncan, 1985) has already been referred to; it had a response rate of 87% and found figures of 9% for males and 14% for females. The study by Lopez, Carpintero, Hernandez, Martin and Fuertes (1995) used a Spanish sample (aged 18 to 60) and attained a response rate of 82%; this study found that 15% of males reported sexual abuse, as did 22% of females. Differences in the definition of child
sexual abuse might account for the discrepancy between the figures reported by Badgley et al. (1984) and the considerably lower figures reported by the two other studies.

3.3.2.3 Definition

The variability in research definitions of child sexual abuse, discussed earlier in this chapter, will also affect prevalence rates. This is most clearly illustrated when researchers apply different definitions of sexual abuse to the same sample. Roosa, Reyes, Reinholtz and Angelini (1998) in a study of adult women (N = 1806) obtained a rate of 18% when the definition excluded non-contact abuse and peer abuse. When peer abuse was included the prevalence rate increased to 39%, and when non-contact abuse was included this increased further to 59%. Gorey and Leslie (1997) in their review of 16 prevalence studies found that differences in definition made the largest contribution to variance in prevalence rates ($B = 0.63$, $p<0.05$).

Although there is little agreement about the definition of child sexual abuse, some approaches should be preferred over others. The definition of Badgley et al. (1994) included experiences after the age of 18 years if this was the earliest age at which an unwanted sexual experience occurred. The prevalence rates obtained in this study are more appropriately considered a rate of sexual abuse, rather than childhood sexual abuse. If this study is excluded, this leaves us with two national probability studies with adequate response rates, and a narrow band of prevalence figures (Females: 14 to 22%; Males: 9 to 15%).

3.3.3 Retrospective underreporting

All of the prevalence studies discussed above use a retrospective design. It is possible, however, that retrospective reports may lead to an underestimate of the occurrence of sexual abuse. Chapter 4 will discuss in detail this general methodological issue; here the focus is on retrospective reports of sexual abuse.
In a methodologically impressive study, Widom and Morris (1997) identified cases of sexually abused males \((N = 19)\) and females \((N = 75)\) processed through the courts in the USA between 1967 and 1971; twenty years later this sample was traced and interviewed about experiences of maltreatment. The authors used four approaches to measuring sexual victimisation. The first defined sexual abuse as a report by the subject of at least one sexual experience before the age of 12, on the basis of a list of behaviours presented to the subject ranging from sexual suggestions to intercourse. The second approach asked the participant if they themselves thought that the experience constituted sexual abuse. The third approach asked if the subject had experienced a sexual act prior to the age of 12 with a person who was at least ten years older. The final approach asked ‘has anyone ever bothered you sexually or tried to have sex against your will?’ The proportion of female victims who failed to report the experience at follow-up ranged from 32% to 60% depending on the measure of sexual victimisation used. For male victims the degree of underreporting, 67.9% to 100%, was even more striking.

Fergusson, Horwood and Woodward (2000) report data from a longitudinal study that also raises concerns about the accuracy of retrospective reports of childhood victimisation. The sample \((N = 983)\) were interviewed about child sexual abuse at ages 18 and 21, and the researchers found that the test-retest kappa for this report was 0.45, which suggests a marked discrepancy between the reports at the two time-points. A study by Sjoberg and Lindblad (2002) also questions the accuracy of self-reports of sexual abuse. They found that children whose sexual abuse was documented on videotape tended to either deny or minimise their experiences. There is, therefore, substantial evidence that child sexual abuse is underreported.

### 3.3.3.1 Male victims and underreporting

Almost without exception prevalence studies have found that females report higher rates of sexual victimisation than males (Peters, Wyatt, & Finkelhor, 1986; Dhaliwal, Gauzas, Antonowicz, & Ross, 1996). On the basis of the two national
probability samples discussed above, three females are victimised for every two males. This may reflect either a genuine difference in the prevalence rates, an unwillingness on the part of males to report these experiences or a combination of both of these reasons (Urquiza, & Keating, 1990; Watkins & Bentovim, 1992). The Widom and Morris (1997) study provides some evidence that males are more likely to underreport victimisation. In this study, whatever the measure of sexual victimisation used, males that were known to have been sexually victimised retrospectively underreported sexual abuse compared to females. However, the Widom and Morris (1997) study cannot address whether male underreporting can entirely account for the gender difference in prevalence rates.

3.3.4 Summary

On the basis of the most methodologically robust studies 9 to 15% of males and 14% to 22% of females are sexually abused as children. However, these figures must be viewed in the context of the respondent underreporting identified by Widom and Morris (1997). Females consistently report higher prevalence rates than males, but the extent to which this reflects male underreporting or genuine differences in the prevalence rates is unclear.

The evidence suggests that a substantial proportion of both males and females are sexually abused in childhood. Given the prevalence of child sexual abuse, it is important to establish the impact of this experience on development; this is examined in the next section.
3.4 The effects of child sexual abuse

3.4.1 Introduction

A large number of studies have examined the relationship between child sexual abuse and mental health problems, and these have been summarised in a number of narrative reviews (e.g. Bauserman & Rind, 1997; Beitchman, Zucker, Hood, DaCosta, & Akman, 1991; Beitchman, Zucker, Hood, DaCosta, Akman, & Cassiva, 1992; Briere & Runtz, 1993; Browne & Finkelhor, 1986; Kendall-Tackett, Williams, & Finkelhor, 1993; Urquiza & Capra, 1990; Watkins & Bentovim, 1992). However, this section will discuss the results of the five meta-analytic reviews that have been conducted to date (Jumper, 1995; Neuman, Houskamp, Pollock, & Briere, 1996; Paolucci, Genuis, & Violato, 2001; Rind & Tromovitch, 1997; Rind, Tromovitch, & Bauserman 1998). There are a number of reasons for this decision. These meta-analytic studies provide an estimate of the overall association in the form of an effect size between sexual abuse and a given mental health outcome; this helps when we examine the association criterion (see below). The effect sizes are based on the pooled results of different studies, which increases statistical power and helps identify small and medium effect sizes. Other reasons include the extensive, systematic and explicit approach the studies use to identify relevant studies, and the explicit use of inclusion and exclusion criteria in selecting studies for analysis.

3.4.2 Psychological problems associated with child sexual abuse

The review of Rind, Tromovitch and Bauserman (1998) illustrates the variety of mental health problems said to be linked to sexual abuse. This review discusses outcomes such as alcohol problems, anxiety, depression, eating disorders, obsessive-compulsive symptoms, psychotic symptoms, sexual adjustment, social adjustment, and suicidal ideation and behaviour.
These and other problems have been reported for both males (Dhaliwal, Gauzas, Antonowicz, & Ross, 1996; Watkins & Bentovim, 1992) and females (Beitchman, Zucker, Hood, DaCosta, Akman, & Cassavia, 1992), with effects reported over the course of childhood, adolescence and adulthood.

### 3.4.3 Causal status of the effects

Despite the considerable number of studies and the range of possible mental health problems, there remains a debate about the causal status of sexual abuse (e.g. Dallam et al., 2001; Rind, Tromovitch, & Bauserman, 2000; Rind, Tromovitch, & Bauserman, 2001). The criteria for a causal relationship between a risk factor, such as sexual abuse, and an outcome are discussed in chapter 4. For the moment it is sufficient to say that four criteria must be met:

- **Association:** There must be an association between childhood sexual abuse and the outcome. (E.g. The sexually abused group are more likely than a non-abused comparison group to report a particular problem.)

- **Non-spuriousness:** The association cannot be attributed to another variable or set of variables. (E.g. The sexually abused group may be more likely than a comparison group to report a particular problem, but they may also be more likely to have been exposed to physical abuse, and it may be this, rather than the sexual abuse, that causes the psychological problem.)

- **Temporal precedence:** The experience of childhood sexual abuse must precede the outcome. (It is important to exclude the alternative temporal relationship, in which a person with a particular psychological problem is at an increased risk of being sexually abused as a result of that psychological problem.)

- **Mechanism:** A logical or plausible mechanism must be identified to explain the relationship between sexual abuse and the outcome.

The next section uses these criteria to examine the evidence for a causal role of child sexual abuse in mental health problems.
3.4.3.1 Association

The five meta-analytic reviews have assessed the strength of the association between child sexual abuse and psychological problems by calculating an effect size (d or r) (Jumper, 1995; Neuman, Houskamp, Pollock, & Briere, 1996; Paolucci, Genuis, & Violato, 2001; Rind & Tromovitch, 1997; Rind, Tromovitch, & Bauserman, 1998). (Small, medium and large effect sizes correspond to d’s of 0.20, 0.50 and 0.80, and r’s of 0.10, 0.30 and 0.50; Cohen, 1988.)

Jumper (1995) examined 30 samples from 26 published studies consisting mainly of female subjects. The studies were divided into community, clinical and undergraduate student samples. The effect sizes for self-esteem, depression and other symptoms are summarised in table 3.2.

Table 3.2 – Effect sizes (r) for self-esteem, depression and other symptoms reported by Jumper (1995)

<table>
<thead>
<tr>
<th></th>
<th>Community samples</th>
<th>Clinical samples</th>
<th>Student samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-esteem</td>
<td>0.34</td>
<td>0.36</td>
<td>-0.02</td>
</tr>
<tr>
<td>Depression</td>
<td>0.17</td>
<td>0.34</td>
<td>0.09</td>
</tr>
<tr>
<td>Other symptoms</td>
<td>0.27</td>
<td>0.29</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Neuman, Houskamp, Pollock and Briere (1996) examined 38 published studies, consisting of samples of female subjects, and found a small effect size for non-clinical samples (d = 0.32) and a medium effect size for clinical samples (d = 0.50). Paolucci, Genuis and Violato (2001) used 37 published studies and calculated weighted effects sizes (Glass’s d) for a number of outcomes, including Post-Traumatic Stress Disorder (d = 0.40), depression (d = 0.44), suicide (d = 0.44), sexual promiscuity (d = 0.29) and poor academic performance (d = 0.19). Rind and Tromovitch (1997) calculated effect sizes for 7 male and 7 female national probability samples, and found small effect sizes for both males (r = 0.07) and females (r = 0.10). Rind, Tromovitch and Bauserman (1998) examined 59 published
and unpublished studies of college samples. Effect sizes were calculated for each of the outcomes listed above; the findings are summarised in Table 3.3.

Table 3.3 – Effect sizes (r) and confidence intervals for the relationship between child sexual abuse and psychological outcomes reported by Rind, Tromovitch and Bauserman (1998)

<table>
<thead>
<tr>
<th>Psychological outcome</th>
<th>Effect size and confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol problems</td>
<td>0.07 (0.02 - 0.12)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.13 (0.10 - 0.15)</td>
</tr>
<tr>
<td>Depression</td>
<td>0.13 (0.10 - 0.14)</td>
</tr>
<tr>
<td>Dissociation</td>
<td>0.09 (0.04 - 0.15)</td>
</tr>
<tr>
<td>Eating disorders</td>
<td>0.06 (0.02 - 0.10)</td>
</tr>
<tr>
<td>Hostility</td>
<td>0.11 (0.06 - 0.16)</td>
</tr>
<tr>
<td>Interpersonal sensitivity</td>
<td>0.10 (0.06 - 0.15)</td>
</tr>
<tr>
<td>Locus of control</td>
<td>0.04 (-0.02 - 0.09)</td>
</tr>
<tr>
<td>Obsessive compulsive symptoms</td>
<td>0.10 (0.06 - 0.15)</td>
</tr>
<tr>
<td>Paranoia</td>
<td>0.13 (0.07 - 0.16)</td>
</tr>
<tr>
<td>Phobia</td>
<td>0.12 (0.07 - 0.17)</td>
</tr>
<tr>
<td>Psychotic symptoms</td>
<td>0.13 (0.06 - 0.15)</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>0.04 (0.01 - 0.07)</td>
</tr>
<tr>
<td>Sexual adjustment</td>
<td>0.09 (0.07 - 0.11)</td>
</tr>
<tr>
<td>Social adjustment</td>
<td>0.09 (0.04 - 0.10)</td>
</tr>
<tr>
<td>Somatization</td>
<td>0.10 (0.06 - 0.12)</td>
</tr>
<tr>
<td>Suicidal ideation and behaviour</td>
<td>0.09 (0.06 - 0.12)</td>
</tr>
<tr>
<td>Wide adjustment</td>
<td>0.11 (0.06 - 0.15)</td>
</tr>
</tbody>
</table>

The results of the five meta-analyses are reasonably consistent: they have reported low to moderate, but predominantly low, effect sizes. It is important to assess whether methodological features of the studies reviewed may decrease the observed association between child sexual abuse and the psychological outcomes. All of the
studies reviewed in the meta-analyses use a retrospective measure of sexual victimisation, and this may be one such feature. On the basis of the Widom and Morris (1997) study, a minimum of a third of sexually victimised females and two thirds of victimised males do not report the experience retrospectively. This could attenuate the adjusted association between sexual abuse and psychological outcomes.

The psychometric properties of the psychological outcome measures may also affect the association; in particular, studies using measures with poor or unknown psychometric properties may decrease the effect sizes. The Rind, Tromovitch and Bauserman (1998) review describes the measures used for each of the eighteen types of psychological outcome they examine. Most of the outcomes include investigator authored items, which are likely to have unknown psychometric properties. Of the seven studies reviewed by Rind and Tromovitch (1997), two do not use instruments with recognised psychometric properties. It is possible that the inclusion of these studies may decrease the association between sexual abuse and the outcome measures. The other reviews do not provide details of the psychometric properties of the studies, and so this assessment cannot be made for the results of these meta-analyses.

In summary, although low effect sizes are reported, methodological features of the studies may have attenuated these associations.

3.4.3.2 Non-spuriousness

Of the studies reviewed in the five meta-analyses that find an association between sexual abuse and a psychological outcome, few address the criterion of non-spuriousness. Rind, Tromovitch and Bauserman (1998) make the most comprehensive assessment of this criterion. The authors identified thirteen studies that examined the contribution of sexual victimisation to psychological outcomes controlling statistically for a range of different variables, broadly classifiable as ‘family environment’. Of 34 significant associations between victimisation and
psychological outcomes, only 14 remained significant once the family environment was controlled for.

Just as methodological features of the study could artificially lower the association at the bivariate level, they could have the same effect on adjusted associations. Rind, Tromovitch and Bauserman (1998) discuss a number of these features, but find only limited evidence that they may have been in operation. For example, they report that in many of the studies there is a low base rate of sexual victimisation, which can lower an adjusted association. The authors dealt with this by using a formula that provides an estimate of the association under different assumptions about the base rate. This calculation suggested that if the base rate were 50% this would increase the association by 0.03 at most.

The statistical controls used in these studies are also problematic if the causal relationship between the two independent variables is not known. If child sexual abuse increases the likelihood of family problems, which in turn leads to psychological problems, the variance shared by the family environment and sexual abuse should be attributed to the abuse. An adjusted association for child sexual abuse in which this variance has been removed could lead to the false conclusion that sexual abuse is not causally implicated in psychological problems. Rind, Tromovitch and Bauserman (1999) suggest that this interpretation is unlikely. They point to a study by Ageton (1988), which found that family problems predated sexual victimisation, to support their view. However, Ageton (1988) uses a retrospective design, which makes it inappropriate to draw conclusions about temporal precedence (Menard, 1991). In the absence of a longitudinal assessment it is impossible to assess the temporal relationship between sexual abuse and family environment problems.

Rind, Tromovitch and Bauserman (1998) do mention a number of features of the studies that could have attenuated the adjusted association between child sexual abuse and the psychological outcomes examined in their review. The capacity of a
research study to identify a significant adjusted association for a variable is determined by the statistical power of the study. Low statistical power, resulting from an inadequate sample size, may be responsible for the non-significant adjusted associations. They also report that for male participants unwanted sexual abuse was associated with worse symptoms. If the studies had contrasted wanted and unwilling sexual experiences, they may have found that consent moderated the relationship with the outcome. An additional reason, not mentioned by Rind, Tromovitch and Bauserman (1998), is the effect of participants retrospectively underreporting the experience of sexual abuse. As discussed above, this may have attenuated the unadjusted associations in these studies and this would also affect any adjusted associations.

3.4.3.3 Temporal precedence
To establish a causal role for sexual abuse, it is also necessary to demonstrate that the experience of the abuse predates the onset of the psychological problem. Ageton (1988), as discussed above, and Burnam et al. (1988) report that sexually victimised participants tended to be symptomatic both before and after the experience of sexual victimisation. One possible implication of this finding is that the temporal relationship between sexual abuse and psychological outcomes is not consistent with a causal role for the abuse experience. However, the retrospective nature of this study seriously limits the faith that can be placed in this interpretation. Retrospective studies cannot be relied on to make accurate timings of temporal order, because participants may fail to accurately recall the order events occurred in. Longitudinal studies are needed to establish this criterion.

3.4.3.4 Causal mechanism
The association, non-spuriousness and temporal precedence criteria are essential in providing evidence for a causal hypothesis, but this is usually only a first step. The identification of a mechanism that connects the putative cause and the effect substantially strengthens the causal hypothesis, yet the five meta-analyses do not
address this criterion. But perhaps this is unsurprising – much more work is still needed on the basic causal criteria.

3.4.4 Summary

Childhood sexual abuse has been linked to a vast range of mental health problems, but few studies have examined the causal status of sexual abuse. This section has discussed the five meta-analytic reviews conducted in this area. These reviews have found a predominantly low association between victimisation and psychological problems, but it is possible that methodological features of the studies may have artificially decreased this association. Few studies have addressed the issue of non-spuriousness; of those that have, a number have found that the association is rendered non-significant once family environment has been controlled for. However, the validity of these statistical associations is open to question, and it would be inappropriate to conclude on the basis of the available research that the relationship between sexual victimisation and mental health problems is spurious. The criteria of temporal precedence and mechanism have not been comprehensively addressed in any of the studies reviewed.

It is important to recognise that the hypothesis that child sexual abuse is causally implicated in mental health problems still stands. An absence of methodologically robust research that has effectively addressed the causal criteria cannot be taken as evidence that there is no causal role for sexual victimisation.

3.5 Strategies for the prevention of child sexual abuse

3.5.1 Introduction

The substantial proportion of males and females that are sexually victimised during childhood, and the possibility that these experiences are linked to subsequent mental
health problems, suggests there is a need for strategies that reduce the occurrence of sexual abuse. Two prevention strategies have dominated these attempts: primary prevention strategies have focused on protecting potential victims from becoming victims or encouraging the victims to report their abuse, while secondary prevention strategies have attempted to ensure that perpetrators do not offend again. There is, however, a disagreement about how best to prevent the occurrence of child sexual abuse (Putnam, 2003). This chapter will review the effectiveness of the two traditional strategies, and argue that there is a need to develop new prevention techniques.

3.5.2 Primary prevention of child sexual abuse

Educational programs designed to teach children how to avoid potentially abusive situations have proliferated in the past two decades (MacMillan, MacMillan, Offord, Griffith, & MacMillan, 1994). These programs originated largely on the basis of clinical information (Reppucci & Haugaard, 1989), but more recent research has attempted to examine the effectiveness of these methods. MacMillan, MacMillan, Offord, Griffith and MacMillan (1994) review eighteen such studies. The interventions included various combinations of verbal instruction, videotaped presentations and behavioural training. Subjects taking part in the intervention program were compared with groups not receiving the intervention. The assessment of effectiveness fell into four categories: basic knowledge of prevention concepts, children's responses to hypothetical vignettes, behavioural responses of children under simulated conditions and child disclosure of sexual abuse.

A number of studies have clearly demonstrated that educational programs can increase children’s knowledge about prevention (MacMillan, MacMillan, Offord, Griffith, & MacMillan, 1994), but increased knowledge may not translate into effective behaviour (Reppucci & Haugaard, 1989). Only a small number of studies have attempted to assess the effects on behaviour of these interventions.
Two studies examined the rate of disclosure of sexual abuse (Kolko, Moser, Litz, & Hughes, 1987; Kolko, Moser, & Hughes, 1989), and one of them (Kolko, Moser, & Hughes, 1989) also examined reports of inappropriate touching. Some differences were found; for example, in the 1989 study the intervention subjects were more likely than the comparison group to report experiencing inappropriate touches \((p < 0.07)\). However, participants were not assigned at random to the two conditions, which limits the causal interpretation of these data.

Other studies used simulated abductions and found that children who have received educational programs were more likely than a comparison group to successfully respond to the situation (Fryer, Kraizer, & Miyoshi, 1987a; 1987b; Poche, Yoder, & Miltenberger, 1987). Still, a question remains about the extent to which these simulated situations bear a relationship to an actual sexual assault. In particular it may not be appropriate to extrapolate from simulations in which force was not used to situations that either involve force from the outset or may come to involve force if the child were to attempt to stop the abuse.

### 3.5.3 Secondary prevention of child sexual abuse

Secondary prevention aims to reduce re-offending in identified sex offenders. There is substantial disagreement about the effectiveness of these treatment strategies. The most striking example of this is the very public spat between Marshall and Quinsey (Marshall, 1993; Marshall, 1998; Marshall, Jones, Ward, Johnston, & Barbaree, 1991; Quinsey, 1998; Quinsey, Harris, Rice, & Lalumiere, 1993). Although these two psychologists are departmental colleagues (Department of Psychology, Queen’s University at Kingston, Canada) they disagree vehemently about the proven effectiveness of sex offender treatment strategies, with Marshall arguing that the available research suggests treatment is effective and Quinsey arguing that this conclusion is not justified.
There are some uncontroversial criteria that need to be met to draw quasi-definitive conclusions about treatment effectiveness; these will be outlined first and then the areas of disagreement will be discussed. An exhaustive literature review (Hanson & Bussiere, 1998) found a 13.4% recidivism rate over 4 to 5 years ($N = 23393$; 18.9% for 1839 rapists and 12.7% for 9603 child molesters). This low base rate of sexual re-offending — or at least the substantial underreporting of sexual offences — may make it difficult to detect a significant effect. The problem can be dealt with to some extent with a suitable duration of follow-up and the use of multiple sources of data (e.g. police, social service, probation, victim reports). Other strategies include the examination of court records for non-sexual offences, because a sexual crime may be reduced to a non-sexual one during plea-bargaining (Quinsey, Harris, Rice, & Lalumiere, 1993). Attention must also be given to whether the treatment condition may differentially affect the likelihood of detection for a particular source. For example, self-reports of the offenders must be used with caution, because treatment could affect the willingness to self-report sexual crimes. In addition, there is a need for cross-validation in terms of setting and offender characteristic (particularly pre-treatment level of risk of the sex offenders) to establish the treatment’s domain of generalisability.

There is general agreement about the necessity of meeting these methodological standards, but there is disagreement about two questions. First, do existing studies (most of which are non-randomised) provide evidence of treatment efficacy? Secondly, is randomisation needed to establish the efficacy of a sex-offender treatment program?

There are a large number of treatment studies, but these are difficult to interpret. Marshall (1993; Marshall, Jones, Ward, Johnston, & Barbaree, 1991) and Quinsey (Quinsey, Harris, Rice, & Lalumiere, 1993), for example, arrive at contradictory conclusions despite using the same set of research studies (Marshall, Jones, Ward, Johnston, & Barbaree, 1991). It is of note that Quinsey, Harris, Rice and Lalumiere (1993) end their critique with a plea for meta-analytic procedures. To date, however,
only one meta-analysis of sex offender treatment studies has been conducted (Hall, 1995). This review found a small effect for treatment ($r = 0.12; N = 1313$); the recidivism rate for treated sex offenders was 19% compared to 27% for the untreated offenders. Although the meta-analysis provides evidence from non-randomised studies of treatment efficacy, this may be a side issue. The second and more substantial source of disagreement is whether it is possible to make inferences about treatment efficacy in the absence of a randomisation procedure (McConaghy, 1999; Marshall, 1993; Quinsey, Harris, Rice, & Lalumiere, 1993).

The issue of randomisation will be discussed in detail in chapter 4. However, it is useful to briefly summarise the arguments here. There are two alternatives to randomisation to control for potential confounds: matching procedures and statistical controls. It is only practically possible to match on two or three potential confounding variables. However, of the 69 variables (all independent of treatment) examined in the meta-analysis of Hanson and Bussiere (1998) approximately one-third of these were significantly related to sexual recidivism ($p < 0.05$) with correlations ($r$) of 0.10 or greater. Given the small effect sizes found for treatment (Hall, 1995) even a small number of these variables non-randomly distributed between the treatment and comparison group could account for the observed group differences. The other alternative to randomisation is to use statistical controls. In this approach the potential confounds between the two groups are measured and then controlled for statistically. As Cook and Campbell (1979) point out, there are substantial difficulties with this approach. In a non-experimental design, two criteria must be met to infer a causal relationship after statistically controlling for potential confounds: all possible confounds must be identified and there must be no error in measurement. The point does not need to be laboured that these two criteria are difficult to meet, particularly in this area of research.

Inferences about causality using non-experimental data are inevitably problematic, and data based on randomisation procedures should be preferred if such data exist. Although the use of randomised procedures for treatment of sexual offending
necessarily introduces a number of ethical issues (Marshall, 1993), three studies have used random assignment procedures to assess the effects of treatment on re-offending rates; these are summarised in the Cochrane review of White, Bradley, Ferriter and Hatzipetrou (2001).

Marques, Nelson, West and Day (1993) randomly assigned child molesters to either a cognitive-behavioural treatment (N = 71) or non-treatment group (N = 79), and assessed recidivism using arrests for a sex crime over a mean of 2.7 years. Of the treatment group, 7.9% had re-offended compared to 10% of the non-treatment group. The difference is not statistically significant, but the authors argue that a longer follow-up time is needed to assess the efficacy of the treatment program.

Romero and Williams (1983) randomly assigned sex offenders (sexual assaulters, child molesters, exhibitionists) to either group therapy and probation (N = 148) or probation alone (N = 83). Assessment of recidivism was made over a ten-year period and consisted of rearrest for a sexual offence. The study found that 13.51% of those subjects who received treatment subsequently committed a sexual offence, and this is in fact higher than the group that had not received treatment (7.23%).

McConaghy, Blaszczynski and Kidson (1988) randomly assigned 30 sex offenders to one of three treatment conditions: Medroxyprogesterone alone, imaginal desensitization alone, and Medroxyprogesterone and imaginal desensitization. There were no significant differences between the three treatments. However, the small sample size and the low base rate of re-offending (N = 3) compromises the conclusions of this study. In addition, the absence of a no-treatment or placebo condition means that it is unclear whether the treatments are equally effective or equally ineffective.

The question of whether sex offender treatment programs are effective remains controversial. Although there is some supportive evidence from studies that have not used randomisation methods to assign treatment, the studies that have used
randomisation have not found treatment to be effective. To complicate matters these studies may fail to meet a number of the other criteria of the ideal study outlined earlier. For example, the follow-up period of the Marques, Nelson, West and Day (1993) study is less than three years. In addition, three studies hardly constitutes extensive cross-validation.

One final point is worth making. Whether or not these programs do reduce recidivism, there is an inevitable limitation of these approaches: such programs can only hope to stop abuse from reoccurring; they cannot stop it happening in the first place.

3.6 Summary

The prevalence literature suggests that a substantial proportion of both males and females are sexually abused during childhood. Research examining the long-term effects of these experiences is limited in that it has not addressed the issue of causality. Despite this, there is some evidence that sexual abuse is likely to be implicated in the development of psychological problems.

The doubts that have been raised about the two existing prevention strategies highlight the need to consider other possible approaches aimed at the prevention of child sexual abuse. One approach is to identify people at risk of perpetrating and divert them from this trajectory. With the exception of a comment by Ryan (1997a) that the strategy is worthy of research, this approach has received no attention in the literature. The aim of this thesis is to take the first step in identifying those people at most risk.

This requires the identification of risk factors for sexual perpetration, and the assessment of the capacity of these factors to discriminate between those people who do and do not perpetrate. This process is helped if we can establish the causes of
sexual perpetration, but few studies have systematically attempted to do this. Subsequent chapters will review the causal status of a number of putative risk factors for sexual perpetration. But first it is necessary to discuss the concept of cause, the criteria for establishing a causal relationship, and the methodological and statistical implications of these criteria.
4 Causality and sexual perpetration

4.1 Introduction

Chapter 3 argued for the need to develop primary prevention strategies that identify and treat potential sexual perpetrators before the onset of abusive behaviour. While it is not necessary to identify the causal antecedents of a class of behaviour to predict that behaviour, the identification of the causes of sexual perpetration would greatly facilitate the development of actuarial instruments to predict sexual perpetration. Aside from this actuarial issue, the question of the causes of sexual perpetration is an important question in its own right. Most risk factor research, for example, aims to establish the causal status of the risk factors (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997).

Much of this chapter examines how claims about causal relations are to be validated or under what circumstances a relationship can be inferred. The chapter will outline a set of criteria necessary for inferring a causal relationship, and it will discuss the difficulties of meeting these criteria in psychological research in general and sexual perpetration research in particular.

This discussion serves three purposes. First, the criteria are used to highlight the limitations of previous research on risk factors for sexually abusive behaviour (chapters 5 and 6). Although there is a large literature on risk factors for sexual offending, little attention has been paid to these causal criteria. To date, only a brief invited commentary has addressed some of the issues raised in this chapter (Cannon, 2001). Secondly, the criteria are used to evaluate the extent to which the empirical work described in this thesis can overcome these problems (chapters 9 and 14). Finally, they are used to suggest future steps for increasing our understanding of the causes of sexual perpetration (chapter 14).
4.2 Preliminary philosophical issues

As White (1990) points out, theories of causation form part of or can be derived from a philosophical position. The concept of cause derived from this doctrine will determine the criteria we have for inferring a causal relationship. Before we move on to discussing these criteria, it is necessary to at least briefly mention a number of fundamental issues relating to these philosophical positions.

4.2.1 Logical positivist and realist conceptions of causality

Hume (1739; 1777) provides probably the most famous analysis of cause, an analysis that is a consequence of his positivist philosophy. For Hume only the experience of the senses can be known, and therefore all unobservables, including causes, have no status. Hume argued that all we observe in an apparent instance of causation is one event followed by another; we do not observe one event causing another event. To use his example, if a moving billiard ball strikes a stationary billiard ball, and that second ball begins to move, no cause is observed; all we observe is one event followed by another. From this perspective causality is reduced to being nothing more than the occurrence of observable regularities.

This approach finds it modern form in logical positivism. A logical positivist analysis, such as that of Russell (1913), attempts to replace the notion of cause with that of functional relationship of a mathematical form. Rather than attempt to identify the cause of a phenomenon, the aim of science is to derive functional relations that can accurately predict a phenomenon. In psychology this approach is the philosophical basis of Behaviourism (Locke, 1972). To quote Skinner (1953): 'The old 'cause-and-effect connection' becomes a 'functional relation'. 'The new terms do not suggest how a cause causes its effect; they merely assert that different events tend to occur together in a certain order.' As applied to the study of sexual perpetration, the aim would be to develop better predictors of sexual perpetration, but abandon an attempt to delineate the causal mechanism.
The motivation behind Hume, Russell and Skinner's rejection of the concept of cause is the difficulty or impossibility of making valid inferences about an unobservable concept. However, there are unfortunate consequences of replacing the concept of cause with that of a functional relationship. This approach allows of no distinction between high correlation and causality. To use Cook and Campbell's (1979, p. 10) example, the high correlation between hair length and opposition to the Vietnam War would be given a causal status. It also fails to acknowledge the asymmetry of causal relationships. For example, the position of the sun can be predicted from the shadow cast by a sundial, and the position of the shadow can be predicted from the position of the sun. However, a causal relationship can only operate in one direction – the sundial cannot cause the position of the sun.

The replacement of positivism with realism as a philosophical basis for science provides a way out of this problem (Manicas & Secord, 1983). This approach assumes the existence of unobservables, such as cause, but recognises that we may be fallible in explaining phenomena. At the basis of this approach is a combination of ontological realism (causes exist as part of the world) and epistemological fallibilism (we may be fallible in identifying them). The note of epistemological fallibilism does not mean that it remains impossible to identify what the cause of a phenomenon is, and this is precisely because of the ontological realism. The assumption that causes do exist as part of the world constrains what we identify or reject as a possible cause. The realist approach has replaced logical positivism as the dominant theory in the philosophy of science, and this is reflected in the move from behaviourism to cognitive approaches as the dominant conceptual framework in psychology (Manicas & Secord, 1983).

4.2.2 Physicalist and constructivist views of mental causation

One further issue that needs to be briefly considered is the question of whether cognitive states can be assigned a causal role. As mentioned, the dominant paradigm
of psychology is a cognitive one. A fundamental assumption of the cognitive paradigm is that cognitive states have a causal role in behaviour (Bolton & Hill, 1996; Bolton, 1997).

As applied to the development of sexual perpetration, the assumption is that exposure to a particular risk factor brings about a change in cognitive functioning and this alteration in cognitive functioning leads to sexual perpetration. To describe the assumption in more detail, there will be a complex set of transactions between the child (biology and previous experience) and environment (risk factors). The transactions at one stage will determine the cognitive state of the child at a subsequent stage (Sameroff, 1995). Over time this complex causal chain will culminate in sexually abusive behaviour. However, in both its simple and more complicated form there is the assumption that the cognitive states are causally efficacious. This assumption is in-keeping with common sense or folk psychology views of cognition, but it is philosophically problematic (Bolton, 1997; Heil & Mele, 1995; Bem & de Jong, 1997).

The problem relates to the distinction between reasons for behaviour (e.g. disliking a person) and causes of behaviour (e.g. a neurophysiological process). If it is maintained that a reason is a cause, because a reason state is underpinned by a neurophysiological state, then we face the problem of reduction. Why cannot the reason state be reduced to the neurophysiological state? If it is reduced, then the reason state has no causal role; it is just an epiphenomenon. This is the position of the physicalists. The constructivists draw a distinction between reasons and causes. Reasons represent discursive processes embedded in a social context; they are separate from neurophysiological states. In drawing this distinction the constructivists argue that cognitions cannot be viewed as causes.

Many philosophers try and find a middle way between physicalist and constructivist positions. This approach attempts to give cognitions a causal status, without reducing cognitions to mere epiphenomena of a neurophysiological process (Bolton
This debate is important once we address the issue of psychological mechanisms, in which cognitions are typically given a causal role. As the thesis is not directly testing psychological mechanisms, the debate will not be addressed in any more detail.

4.3 Causal criteria

This section will suggest that four criteria must be met before a causal relationship can be inferred between a risk factor and sexual perpetration:

- There must be an association between the variables
- That association must be non-spurious
- Temporal precedence must be established
- A logical or plausible mechanism must be identified.

In psychology and psychiatry similar criteria have been outlined by researchers such as Haynes (1992), Kazdin, Kraemer, Kessler, Kupfer and Offord (1997), and Rutter (1995). The following section will discuss the methodological and statistical implications of these criteria for sexual perpetration research.

4.3.1 Association

The association criterion is the most basic one for establishing a causal relationship. It states that there must be an association between the occurrence of the putative cause and the effect.

4.3.1.1 Deterministic vs. probabilistic associations

Some definitions of cause demand a deterministic association between a cause and an effect. This definition demands constant conjunction (if Y occurs, X must have preceded it) and invariable succession (if X occurs, Y must follow it). In other words, the cause must be both necessary and sufficient for the effect.
This definition is too restrictive for psychology. The problem of adequate measurement prevents the identification of a deterministic association (Haynes, 1992), even if one were to exist. Most—possibly all—behavioural outcomes are the result of an INUS condition (Mackie, 1974), which stands for Insufficient Necessary Unnecessary Sufficient. For a particular outcome, such as sexual perpetration, a particular risk factor will be an insufficient (I) but necessary (N) part of a condition that is itself unnecessary (U) but is sufficient (S) for that outcome. A risk factor may increase the risk of an outcome, but that relationship will be neither necessary nor sufficient for its expression. There are yet more problems. Some events can have a substantial effect on the probability of a behavioural outcome, but for the purpose of prediction they are effectively random (Meehl, 1954; 1978). These random but important events may be large as a class, and will affect the relationship between a particular variable and an outcome. This means it will be impossible to state the conditions under which an association will be perfect for a particular risk factor. The association criterion for psychology is, therefore, much more modest: a risk factor need have only a probabilistic association with an outcome.

4.4.3.2 Statistical significance of an association
According to this definition, there must be a statistically significant association between the putative cause and the effect. All that is required by the criterion defined in this way is a statistically significant non-zero relationship between the two variables, and a prediction of the direction of that relationship (either positive or negative). Given this, it is worth noting the critique of null hypothesis significance testing in non-experimental psychology (Cohen, 1994; Judd, McClelland, & Culhane, 1995; Lykken, 1968; Meehl, 1967; 1978; 1990a; Morrison & Henkel, 1970). Meehl's (1990a) version of the critique will be briefly outlined. He begins by pointing out that in psychology everything correlates with everything else to a not insubstantial degree, a phenomenon he terms the 'crud factor'. The existence of this crud factor means that in non-experimental research there will (almost) always be a non-zero relationship between any two variables, and given adequate statistical
power and adequate measurement this will be detected in the form of a statistically significant association. It is important to emphasise that this statistical significance does not reflect a type I error: the correlation is reliable and reflects real facts about the world.

It is usually a straightforward matter to establish a statistically significant non-zero relationship in non-experimental psychology as long as certain methodological requirements (adequate power, adequate measurement) are met. The association criterion, therefore, does nothing more than establish the direction (positive or negative) of the putative causal relationship. This does not mean to say that the criterion is pointless. It is a necessary step to demonstrate that the association is in the predicted direction (Judd, McClellan, & Culhane, 1995), but doing so plays only a very minor role in demonstrating a causal relationship (Winch & Campbell, 1969). For example, of the many threats to validity discussed by Cook and Campbell (1979), only one of these is dealt with by the statistical test for a non-zero relationship.

4.4.3.3 Size of the association

The presence of a ubiquitous and not insubstantial crud factor in psychological research has led to an increased emphasis on effect sizes (Cohen, 1994; Judd, McClelland, & Culhane, 1995). For the association criterion this leads to an additional difficulty; namely, should the criterion state that an association should be of a particular size? Although the arguments have been made about theory testing, they apply just as well to causal relationships. Lykken (1968) argued that on empirical grounds this demand should not made, because the size of an association is determined by the number of factors causally implicated in a particular behavioural outcome. In contrast Swoyer and Monson (1975) argue that the larger the association between two variables the smaller the number of alternative explanations for that variable; therefore the size of the association is important.
The points made by these authors are not mutually inconsistent. On empirical grounds there is no reason why a small association between two variables excludes the possibility of a causal relationship. However, a small association between two variables in non-experimental research may just reflect Meehl’s crud factor. A large association is much harder to dismiss in this way. What this means for defining the association criterion is that a larger association is to be preferred, but a causal relationship should not be ruled out on the basis of a small association.

4.4.3.4 Summary
The association criterion has now been defined. The association can be probabilistic; it need not be deterministic. There must be a non-zero statistically significant association between the variables. The association can be of any size, but the larger the association the better.

4.4.3.5 Factors that obscure an association
The association criterion as just defined means that a lack of a significant association in the predicted direction is consistent with there being no causal role for the risk factor. However, before settling on this conclusion it is important to examine methodological features of the study that could obscure a significant association.

A causal relationship will only operate under certain conditions; therefore the absence of an association under some conditions does not necessarily imply the absence of an association under others (Haynes, 1992). An association may be apparent only in certain environmental contexts, at particular developmental stages, within limited values of the risk factor, for some but not other features of an outcome (e.g. onset, duration, frequency) and when other variables are present or absent (Haynes, 1992).

The developmental stage at which both the risk factor and the outcome are measured can affect the observed association. This has implications for both retrospective and
longitudinal research strategies. For both, a risk may be implicated in sexual perpetration, but this may not be detected if the measurement of the outcome is made before its onset. The same difficulty applies to longitudinal research if the measurement of the risk occurs prior to exposure to that risk. In addition, a factor may operate as a risk only if it is experienced at a certain developmental stage, and may have no affect on the outcome if experienced outside this time. These difficulties can be addressed by using time sampling strategies that measure risk factors and outcomes at a range of time-points across development (Haynes, 1992).

A variable may be causally implicated in the development of an outcome, but this relationship may occur only if the group with the outcome were exposed to an additional risk or were not exposed to a protective factor. This has implications for assessing the results of studies discussed in the next two chapters. A number of studies have concluded that a factor is not implicated in the aetiology of sexual offending because the study failed to find a higher rate of exposure in the perpetrator group than a comparison group. For example, Benoit and Kennedy (1992) found comparable rates of sexual victimisation among adolescent child molesters, aggressive non-sex offenders and non-aggressive non-sex offenders. On the basis of this the authors disputed the assumption that sexual victimisation is a risk factor for sexual perpetration. While this conclusion is consistent with this finding, it is equally consistent with a causal role for sexual victimisation, but one that occurs only in the presence of an additional risk or the absence of protective factors (Garber & Hollon, 1991).

A risk factor may be causal in some but not other features of an outcome (e.g. onset, duration, severity), and it is necessary to specify these features (Haynes, 1992). For example, a risk factor may be causally implicated in the onset of sexual perpetration but not in the severity or the duration of the sexual offending. In addition to these quantitative features the relationship may hold only for certain types of sexual perpetration (e.g. rapist; child molester; perpetrator against males, females or both).
In this study the outcome of interest is restricted to the probability of the onset of sexual perpetration.

Other methodological features of a study could also obscure a substantial and significant association. The accuracy with which perpetration is measured is one feature. The integrity of any comparison of sexual perpetrators with non-perpetrators depends on the correct assignment of the subjects to the two groups. It is likely that the most common problem will be the assignment of people who have committed abusive acts to a non-perpetrator group. Many studies rely on conviction data to assign sexual perpetration status. However, studies that have used more detailed assessments suggest that these comparison groups may contain a substantial proportion of sexual perpetrators. Spaccarelli, Bowden, Coatsworth and Kim (1997), for example, examined the relationship between a number of risk factors and sexual aggression in a sample of 210 delinquents, of which 24 had been arrested for sexual offences. The authors also obtained self-reports of sexual aggression using a sexual assault item from a delinquency measure, and six abuse questions embedded in a thirty-item measure of sexual experiences. Of the 186 subjects arrested for non-sexual crimes – initially classified as non-sex offenders – 14% reported sexual perpetration.

In the same way, a measure of the risk factor that has poor psychometric properties may attenuate an association. Few of the studies reviewed in the subsequent chapters provide details of reliability or validity of the risk factor measures.

The statistical power of a research study can also obscure a significant finding (Cohen, 1988). The capacity of a research study to reject the null hypothesis, when the null hypothesis is false, is determined by the power of the statistic being used and the sample size. Few of the studies discussed in the next two chapters provide details of power calculations.
4.4.3.6 Designs to establish an association

Researchers have used three different approaches to examining an association between risk factors and sexual perpetration (a full review of these approaches is given in the next two chapters). The case-control design compares sexual perpetrators with a non-sexually perpetrating comparison groups (e.g. Benoit & Kennedy, 1992). Both groups are typically sampled from forensic or clinical populations. In this design a significantly higher rate of exposure in the perpetrator than non-perpetrator group, ascertained by retrospective self-report, provides evidence of an association. In a cross-sectional design the researcher compares the proportion of self-reported sexual perpetrators reporting a particular risk factor with those who do not self-report this behaviour (e.g. Borowsky, Hogan, & Ireland, 1997). These studies typically use a non-forensic population, such as university students. Evidence of an association in this design requires a significant relationship between a self-report of exposure to the risk and a self-report of sexual perpetration. A longitudinal study follows a group exposed to the risk and one or more non-exposed groups to establish the proportion of each that perpetrates (e.g. Widom & Ames, 1994). In this design a higher rate of sexual perpetration in the exposed group provides evidence of an association. It is worth mentioning at this point that a genuine longitudinal study carried out in real time would be unethical. A real-time longitudinal design would require researchers to follow a group of subjects, some of which were hypothesised to be at high risk of sexual perpetration because of their exposure to key risk factors. If the hypotheses about these risk factors were to be assessed, it would be necessary to follow the sample without intervening. Other longitudinal strategies such as the catch-up longitudinal design are needed to avoid this ethical problem.

4.3.2 Non-spuriousness

The criterion of non-spuriousness states that the association between the risk factor and perpetration cannot be attributed to other variables (confounds). This is a crucial
step in establishing the role of a risk factor (Rutter & Maughan, 1997), but it is one
that can prove difficult to meet. In the two retrospective designs (case-control and
cross-sectional), even if perpetrators are more likely than non-perpetrators to have
experienced this risk, it is possible that other differences between the groups account
for the association. In the same way in a longitudinal study people exposed to the
risk may be more likely than a comparison group to sexually perpetrate, but this
association could be a result of group differences other than the putative risk.

A convincing demonstration of this criterion requires an experimental design
(Holland, 1986; Rubin, 1974). As Cook and Campbell (1986) point out, there are
two senses in which the term experimentation is used. There is the isolation
experiment in which a laboratory setting is used to control potential confounds.
There is also the random assignment experiment in which confounds are controlled
by the act of randomly assigning experimental units to either an experimental
condition or a control condition (Fisher, 1925). It is in this second sense that the
term experimentation is typically used in psychology.

The act of randomisation helps limit the differences between the experimental and
control group to the causal variable of interest. Cook and Campbell (1979) provide
an exhaustive list of threats to non-spuriousness, and they note that the
randomisation experiment is able to address many of these.

Although the random assignment experiment is the gold standard method for
establishing this criterion the technique cannot be used in this area of research. It
would be unethical to assign subjects to the type of conditions identified by the
literature as possible risk factors for sexual perpetration (e.g. sexual victimisation,
physical abuse, emotional maltreatment). This would be unethical not only because
exposure to the risk would harm the subjects, but also because it is hypothesised that
the risk increases the probability of sexual perpetration, which would harm others.
One important question that must be addressed is whether the non-spuriousness criterion can be met, and therefore whether causal inferences can be made, in the absence of experimentation. This question has led to considerable controversy among both social scientists and statisticians (e.g. Briere, 1988; 1992; Briere & Elliot, 1993; Cook & Campbell, 1979; Lord, 1969; Tabachnick & Fidell, 1989).

A number of methods for making causal inferences in the absence of experimentation have been suggested, such as matching procedures, stratification and statistical controls, natural experiments (Rutter, 1994), and quasi-experiments (Cook & Campbell, 1979). The research literature on risk factors for sexual perpetration has, without exception, used stratification and statistical procedures; the next section will discuss the problems of using this method to make causal inferences.

Stratification involves the selection of a comparison group that has not been exposed to a risk factor, but is broadly comparable to the exposed group on potential confounds. Demographic features such as age, ethnicity and socio-economic status often moderate the relationship between psychological variables, and therefore much of the sexual perpetration research has stratified on these features. With a statistical control the shared variance between the risk and one or more potential confounds is removed from the association between the risk and the outcome. This adjusted association represents the unique contribution of the risk to the outcome.

It can be difficult to interpret these adjusted associations (Briere, 1988, Briere & Elliot, 1993; Cook & Campbell, 1979; Lord, 1969; Tabachnick & Fidell, 1989). A convincing demonstration of a non-spurious relationship using this approach requires the identification of all potential confounds and the perfect measurement of each of these (Cook & Campbell, 1979). Both of these criteria are near impossible to meet, but if not met, a spurious association may remain substantial and statistically significant.
The possible confounds for any putative causal relationship are always, in some sense, infinite – even in an experimental design (Cook & Campbell, 1979). We typically ignore this possibility, and focus on only those potential confounds identified as plausible on the basis of either background knowledge or current theory (Cook & Campbell, 1979). However, in non-experimental psychology the number of plausible alternative explanations for any putative relationship is typically very large. To quote Meehl (1990b): ‘For most statistical findings in soft psychology, I daresay a group of faculty or graduate students could come up with a dozen plausible alternatives to the theory of interest if allowed a morning’s conversation over coffee and Danish’. The requirement that we identify all relevant confounds is going to be difficult to meet.

The perfect measurement of each of the confounds also poses substantial problems for psychology. The validity and reliability of the covariates is often not reported in the sexual perpetration literature. When this information is reported, it serves only to increase our fears. In psychology a claim about validity is usually substantiated by reference to some form of validity coefficient. As Meehl (1990b) points out, a supposedly respectable coefficient does not look that impressive when translated into the amount of shared variance, but despite this the measure is typically classed as valid. If a risk factor remains significant once this ‘valid’ measure of a potential confound is statistically controlled, this is taken as evidence for the causal role of that risk factor.

The requirement that we identify and measure perfectly all relevant potential confounds presents grave problems for the use of statistical controls.

To complicate matters, a small and non-significant adjusted association may in fact represent a causal relationship (Briere, 1988; Briere & Elliot, 1993). If the risk variable of interest and the control variable are highly intercorrelated (multicolinearity), and each correlates with the predictor variable it is difficult to assess the unique contribution of the risk factor to the outcome (Briere, 1988). A
number of methodological features can exacerbate this problem. Just as poor psychometric properties of the measures and a lack of statistical power can attenuate an unadjusted association, they can have the same effect on an adjusted association (Briere, 1988). As Briere and Elliot (1993) note, a lack of adequate statistical power is often combined with the use of many variables; this combination capitalises on error variance, and can lead to erroneous conclusions.

The control variable can also affect the conclusion. If the risk factor is causally implicated in the variable that is being controlled for, then any shared variance between the two should be attributed to the risk. If the statistical control procedure removes the shared variance from the association between the risk and the outcome, then this could lead to the false conclusion that the risk is not implicated in the development of the outcome.

There is a long list of difficulties in using statistical controls to establish a causal relationship between two variables in non-experimental research. Given the ability of the random assignment experiment to successfully address many of these threats to valid causal inference, this method should be used if ethically possible. If the experimental method cannot be employed, as in sexual perpetration research, then we must be modest in any claim we make about a causal relationship (Cook & Campbell, 1979).

4.3.2.1 Retrospective reporting bias

The vast majority of studies use a design in which perpetrators of sexual abuse are asked to retrospectively recall experiences of a particular risk factor, often without any form of corroboration. This raises an additional challenge to the non-spuriousness criterion because the recall of risk factors could be biased by the outcome (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997).

Two reviews (Brewin, Andrews, & Gotlib, 1993; Maughan & Rutter, 1997) have examined in detail the reliability and validity of retrospective reports, and both
conclude that the limitations of retrospective reports have been overstated. Maughan and Rutter (1997) discuss three sources of evidence that can be used to assess the accuracy of these reports: test-retest reliability, independent corroboration using sources such as family members and independent corroboration using contemporaneous written documents. Recall appears to be generally stable according to test-retest studies, and the independent corroboration methods suggest moderate to good agreement with retrospective reports (e.g. Bifulco, Brown, Lillie, & Jarvis, 1997). Brewin, Andrews and Gotlib's (1993) review also suggests that retrospective reports are less susceptible to current mood state than is often supposed. They also suggest that the accuracy of retrospective reports can be improved by the use of techniques such as providing adequate recognition cues.

There are, however, some limitations of retrospective data. For example, although there is no evidence that people with poor outcome tend to overreport negative experiences, underreporting does occur in people with good outcomes (Maughan & Rutter, 1997; Rutter, Maughan, Pickles, & Simonoff, 1998). There is also a notable reduction in adult recall of incidents taking place before the age of 5 (Maughan & Rutter, 1997). Evidence from cognitive psychology also suggests that autobiographical memories are to some extent reconstructed in the light of current experience (e.g. Radke-Yarrow, Campbell, & Burton, 1970).

Although retrospective reports appear to be accurate in many non-clinical and clinical populations, the reports of sexual offenders may be a special case. Sexual perpetrators might invent or elaborate experiences of risk factors to explain or excuse their abusive behaviour (Hindman & Peters, 2001). Any observed association between a risk factor and sexual perpetration may be a result of the invention of an experience, rather than the actual experience of the risk. This threat to the non-spuriousness criterion is extremely problematic for sexual perpetration research, and yet it is rarely mentioned in the literature.
A series of studies have examined this possible source of bias (Hindman & Peters, 2001). The first study, originally reported in Hindman (1988), obtained self-reports of sexual histories from child molesters referred to a clinic between 1980 and 1988. The clinic introduced polygraph testing in 1982, and the study found that 67% of those subjects referred prior to its introduction reported sexual victimisation (N = 40) compared to 29% of those referred after its introduction (N = 129). A second study, conducted in 1994, compared adult sex offenders (N = 76) who self-reported their sexual history with a group receiving polygraph testing (N = 152). By this time all offenders referred to the clinic were polygraphed, but before polygraphing took place the non-polygraph group were either sentenced or excluded from the program at the clinic. The results were similar to the 1988 study, with the non-polygraph group reporting a higher rate of child sexual abuse (65%) than the polygraph group (32%). The final study used a repeated measures design, in which the researchers compared the self-reports of child sexual abuse made with and without polygraph testing in 173 adult offenders. The results were broadly comparable with the earlier studies (61% without testing vs. 30% with testing). As Hindman and Peters (2001) point out, the base rates of victimisation appear to reverse when subject to the scrutiny of a polygraph test, a phenomenon they term ‘Magical X’.

These findings are striking. Thirty years of research – what must amount to hundreds of studies – has relied on the reports of sexual perpetrators, and yet this research might tell us nothing. But ‘Magical X’ has explosive properties only if the Hindman and Peters (2001) studies are methodologically unimpeachable, and not everyone is convinced they are. Lee, Jackson, Pattison and Ward (2001), in their discussion of the Hindman (1988) study, point out that the polygraph and non-polygraph groups ‘were not properly matched on a number of relevant variables’. Pre-existing differences between the two groups could account for the observed differences in the prevalence rates. This same criticism also applies to the second study reported in Hindman and Peters (2001), because it too uses a quasi-experimental design, and the authors fail to report details of the comparability or otherwise of the two groups. This is dealt with to some extent in the final study,
which compares the prevalence rates made by the same sample under testing and no-
testing conditions. However, in this third study all participants received the no-test
condition first and then the test conditions. In Cook and Campbell’s (1979)
terminology this is an AB design, and, as they point out, this is unable to rule out a
number of threats to validity. For example, events taking place between the two
conditions could have led the offenders to alter their reports, even if the first report
was accurate. The use of a counterbalancing procedure would have greatly helped
the interpretability of this third study.

The construct validity of ‘faking’ is perhaps the main weakness of all three studies.
Changes in proportions are taken as an indicator of faking, but there are alternative
explanations of any observed differences. For example, the polygraph group might
have denied a history of childhood sexual abuse to avoid re-imprisonment. Neither
abuse was defined for the sexual perpetrators. If the participants were not given a
definition or were provided with only a vague one, then they had to effectively rely
on their own idiosyncratic definition. They may have tended to use a more
restrictive definition in the test condition, because of the consequences they faced if
they failed the test.

The Hindman and Peters (2001) studies are the only ones to have examined the
validity of sexual offenders’ retrospective reports, but firm conclusions are difficult
to make because of the methodological limitations. However, this particular threat to
non-spuriousness remains until further studies demonstrate that once these
methodological limitations are addressed there is no evidence of perpetrator
invention. Until that time, the threat looms large. The methodological integrity of
any retrospective study is limited by this threat. The longitudinal design avoids the
possibility of invention by sexual perpetrators, because in this design the risk factor
is coded prior to the onset of perpetration.
4.3.3 Temporal precedence

The temporal precedence criterion states that the experience of a risk factor must precede the onset of perpetration. Retrospective research (case-control and cross-sectional) faces difficulties in meeting this criterion. It would be possible to obtain retrospective reports of the onset of both the risk factor and sexual perpetration, but this relies on the accuracy of retrospective recall of temporal order. A more fundamental problem is that even if perpetration did precede the risk, this order could be switched in retrospective recall, because of the perpetrators desire to excuse their behaviour.

This criterion usually requires a longitudinal design (Menard, 1991). In this design, subjects are assessed at a minimum of two points in time: at the first one exposure to the risk factor is assessed and it is established that the outcome is not present; at the second the presence or absence of the outcome is determined (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997). The longitudinal strategy helps to clarify the temporal order of events; however, even this approach faces problems when examining risk factors for sexual perpetration. As Rutter (1989; 1992; Rutter & Sroufe, 2000) points out, many psychopathological outcomes are heterotypic, which means that the particular expression of the cognitive, affective and behavioural features of the outcome is determined by a person's developmental stage. In longitudinal research it may be difficult to identify the onset of the psychopathological outcome if the expression of the outcome is variable across development. This problem is reflected in the discussion in the sexual perpetration literature about how to define sexual perpetration in adolescents and children (Araji, 1997).

An additional problem for longitudinal research is that acts of sexual perpetration are usually committed in private and are surrounded by secrecy. Given this secrecy it may be difficult to establish an accurate date of onset.
4.3.3 Causal Mechanism

The final criterion states that a logical (Haynes, 1992) or plausible mechanism (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997) must be identified to explain the relationship between the risk and the outcome. Stated in this way the criterion is not difficult to meet. In psychology it is usually a straightforward matter to generate a plausible explanation of how a risk factor could lead to an outcome. In fact, it is usually easy to generate a fair number of different, equally plausible explanations (Meehl, 1990b). Given the amount of possible explanations, it is usually necessary to provide evidence of a mechanism (Haynes, 1992).

The demonstration of a causal mechanism requires the identification of mediating variables, which can be defined as processes or mechanisms through which a risk factor leads to an outcome (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997). Theories of sexual abuse provide hypotheses about these mechanisms; chapter 7 will examine these theories.

The requirement that we provide evidence of the causal mechanism leads to the question of what constitutes compelling evidence. This in turn leads back to the critique of null hypothesis significance testing first discussed in the association criterion section. One of the main motivations behind this critique is that the null hypothesis approach is used to corroborate theories about psychological mechanisms, but the corroboration offered by this technique is typically feeble (Cohen, 1994; Judd, McClelland, & Culhane, 1995; Lykken, 1968; Meehl, 1967; 1978; 1990a; Morrison & Henkel, 1970).

In non-experimental psychology, such as sexual perpetration research, theories are typically corroborated using a directional test of significance. The theory states that there will be a correlation between two variables in a certain direction (either positive or negative). Alternatively the theory may state that one group (e.g. the sexual perpetrators) will have a higher mean on a particular variable than another group (e.g. non-sexual perpetrator comparisons). If we accept that everything
correlates with everything else in psychology, as argued above, the relationship between two variables in a non-experimental study will (almost) always be non-zero. In other words there will almost always be either a positive or negative relationship between the two variables, and the mean of one group will almost always be either higher or lower than the mean of another group. As statistical power improves, and as measurement improves, there will be approaching a 1 in 2 chance of corroborating a theory in a non-experimental design, even if the theory is drivel. The problem is that in non-experimental psychology two variables may correlate with each other for any number of reasons, not just the one that is consistent with the theory. This approach to corroboration makes it difficult to distinguish a half decent theory from a poor one, because even poor theories have a good chance of clearing the low hurdle set by a directional prediction. If a theory is corroborated using this method it does very little to convince us of its verisimilitude (truth-likeness).

A theory about a psychological mechanism that makes a directional prediction about the relationship between two variables does not provide compelling evidence favouring that psychological mechanism. We are back to the question of what does constitute compelling evidence.

A number of solutions have been proposed in the literature (Leventhal, 1994; Lykken 1968; Lykken, 1991; Serlin & Lapsley, 1985; Swoyer & Monson, 1975; Woodson, 1969), and even the sketch of an answer, such as the one offered by Meehl (1990a), can be deeply complicated. Meehl’s (1990a) proposed solution and the discussion it produced (Campbell, 1990; Chow, 1990; Dar, 1990; Fiske, 1990; Humphreys, 1990; Kimble, 1990; Kitcher, 1990; Kukla, 1990; McMullin, 1990; Maxwell & Howard, 1990; Meehl 1990c; Rorer, 1990; Serlin & Lapsley, 1990) range over such fundamental issues as the legacy of Fisherian statistics in psychology, the applicability of Bayes’ Theorem to theory corroboration, and Post-Popperian approaches to the philosophy of science.
Fortunately at the centre of most proposed solutions is a nugget of scientific common sense; namely the riskier the prediction made by a theory, the more convinced we are by it when the risky prediction is borne out by the data (Meehl, 1993). A risky prediction is one that has a low prior probability of being correct if the theory making the prediction is false. Take Meehl's Theory of Climate as an example (Meehl, 1978). If Meehl's theory predicts that it will rain sometime next April and the prediction turns out to be correct, the scientific community will be unimpressed. However, if the theory predicts that on April 4th it will rain 1.7 inches, 2.3 inches on 5th April, and so forth, and the theory makes seven correct predictions, then the theory will be considered impressive. A compelling demonstration of a psychological mechanism linking a risk factor to sexual perpetration will involve a corroborated prediction that is in some sense riskier than the usual directional prediction.

'Riskiness' can be increased in a number of ways. Ideally, instead of making a directional prediction (e.g. the theory predicts there will be a positive correlation), a theory should make a point prediction (e.g. the theory predicts there will be a Pearson's r of 0.67) (Meehl, 1978). However, this is beyond the scope of the current, rather primitive state of theorising in this area. In fact, the theories are not able to generate even narrow range predictions (e.g. the theory predicts a Pearson's r between 0.6 and 0.7), a technique recommended by Serlin and Lapsley (1985).

There are other methods for increasing the 'riskiness' of predictions that sexual perpetration theories should be capable of making. For example, most of the theories should be able to make predictions about the temporal relationship between events (e.g. risk factor precedes psychological mechanism, psychological mechanism precedes onset of perpetration). Empirical evidence of the temporal relationship would put 'money in the bank' for a theory. Lykken (1991) suggests a number of techniques for theory testing in psychology that all address the 'riskiness' criterion. If a theory makes a series of quasi-independent directional predictions and these are all borne out then this constitutes a riskier test. As Lykken (1968; 1991) points out, a
theory making one directional prediction in a non-experimental design has approaching a 1 in 2 chance of being corroborated as statistical power and measurement improves. However, if a theory makes five directional predictions this probability drops to $0.5^5$, and with ten such predictions the probability becomes $0.5^{10}$. Another suggestion made by Lykken (1991) draws on the Popperian emphasis on pitting theories against each other (Popper, 1959). If an empirically corroborated prediction flows naturally from a theory and this prediction requires *ad hoc* adjustments from competing theories, then this places ‘money in the bank’ for that theory.

### 4.4 Summary

Convincing evidence of a causal link between a risk factor and sexual perpetration requires an association between the two variables, a non-spurious relationship, a demonstration of temporal precedence, and the identification of a plausible or logical mechanism.

The discussion of the causal criteria emphasises that the experimental method can provide substantial leverage on the causal criteria (Cook & Campbell, 1979), as does the longitudinal method (Rutter, 1995). However, it would be unethical to conduct a genuine random assignment experiment of the causes of sexual perpetration. It would also be unethical to carry out a genuine longitudinal study in real time. Without experimentation and longitudinal research there are considerable difficulties in making causal inferences. This does not mean to say that attempts at causal explanation should be abandoned. The prediction of sexual perpetration would be greatly facilitated by a causal explanation of this behaviour. What it does mean is that addressing causal criteria will require methodological ingenuity and even then causal claims will be modest.
The causal criteria outlined in this chapter will be used in the next two chapters to assess the causal status of a number of putative risk factors for sexual perpetration. The causal mechanism criterion will be discussed separately in chapter 7, which examines theories of sexual perpetration. This is because the theories provide hypotheses about these mechanisms, but each theory often applies to more than one putative risk factor. The next chapter examines the claim that sexual victimisation is a risk factor for sexual perpetration.
5 Child sexual abuse as a risk factor for sexual perpetration

5.1 Introduction

One of the most popular explanations of sexual perpetration is the victim-to-perpetrator hypothesis, which proposes that victims of sexual abuse are at heightened risk of sexually abusing others (Fagan, Wise, Schmidt, & Berlin, 2002; Garland & Dougher, 1990; Hanson & Slater, 1988; Starzyk & Marshall, 2003; Veneziano & Veneziano, 2002). This chapter assesses the evidence for a causal relationship between victimisation and perpetration. Although there is evidence of an association, few studies have addressed the additional causal criteria. The association does appear to be stronger for male than female victims, which suggests that male victims may be at particularly heightened risk. It is clear that by no means every victim develops an abusive pattern of behaviour, and this suggests that other risk factors must be involved.

The studies discussed in this chapter are grouped by research design. The first set of studies provides information about the proportion of sexual perpetrators from a forensic population retrospectively reporting sexual abuse in childhood; these studies do not include a comparison group. The second set of studies use a case-control design in which a forensic population of sexual perpetrators is compared to a forensic population that have not committed a sexual offence (e.g. violent or property offenders). The third set of studies uses a cross-sectional design. In this design a comparison is made of the proportion of self-reported sexual perpetrators reporting sexual victimisation and those who do not self-report committing these acts. These studies typically use a non-forensic population, such as university students. The final design is a longitudinal one in which a sexually victimised group and a non-victimised group are followed to establish the proportion that subsequently commit a sexual offence.
The causal criteria outlined in the previous chapter are used to evaluate each of the designs. The studies are first divided into those that do and do not find evidence of an association. Those studies that do report an association are then examined in more detail to assess whether they demonstrate that the observed relationship is non-spurious and whether temporal precedence has been demonstrated. The criterion of establishing a psychological mechanism is discussed in chapter 7.

5.2 Studies without a comparison group

There is no shortage of evidence to indicate that a substantial proportion of sexual perpetrators report being sexually victimised in childhood (see table 5.1). Prevalence figures for adult males who have abused children range between 19% (Prentky, Knight, Straus, Rokou, Cerce, & Sims-Knight, 1989) and 75% (Romano & De Luca, 1997). The one study of adolescent males who abuse children found a figure of 8% (Adler & Schutz, 1995). A study of heterogeneous adolescent sex offenders (abused either children, peers or adults) reports a rate of 41% (James & Neil, 1996). The two studies of female perpetrators report figures of 38.5% (Allen & Lee, 1992) and 90% (Hunter, Lexier, Goodwin, Browne, & Dennis, 1993). Although these high rates of sexual victimisation appear to hold promise for the abused-abuser hypothesis, in the absence of a comparison group to provide an indication of the association, they do not meet even the most basic criterion for establishing a causal relationship.
Table 5.1 – Proportions of sexual perpetrators reporting sexual victimisation

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adler &amp; Schutz (1995)</td>
<td>Male sibling incest offenders</td>
<td>8</td>
</tr>
<tr>
<td>(Mean age = 16; range = 13 to 19) (N = 12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Mean age = 41) (N = 75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female sex offenders</td>
<td>(Mean age = 33) (N = 65)</td>
<td>38.5</td>
</tr>
<tr>
<td>(Range = 17 to 60+) (N = 36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenberg et al. (1993)</td>
<td>Male child molesters (victim aged &lt;13)</td>
<td>42</td>
</tr>
<tr>
<td>(Mean age = 32; sd = 13.1) (N = 135)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male hebephiles (victim aged 13 to 15)</td>
<td>(Mean age = 41; sd = 14.3)</td>
<td>44</td>
</tr>
<tr>
<td>Hanson et al. (1994)</td>
<td>Male intrafamilial child molesters</td>
<td>39.2</td>
</tr>
<tr>
<td>(Mean age: 34.8; sd = 5.86) (N = 74)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hunter et al. (1993)</td>
<td>Female sexual perpetrators against children</td>
<td>90</td>
</tr>
<tr>
<td>(Mean age = 15; range = 13 to 17) (N = 10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>James &amp; Neil (1996)</td>
<td>Male sex offenders</td>
<td>35.4</td>
</tr>
<tr>
<td>(Mean age = 14.8; sd = 1.7) (N = 31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prentky et al. (1989)</td>
<td>Male sex offenders</td>
<td>19</td>
</tr>
<tr>
<td>(Mean age = 30; range = 18 to 68) (N = 82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Richardson et al. (1995)</td>
<td>Male sex offenders</td>
<td>41</td>
</tr>
<tr>
<td>(Mean age = 15; sd = 1.4; range 11 to 18) (N = 100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Romano &amp; De Luca (1997)</td>
<td>Male child molesters</td>
<td>75</td>
</tr>
<tr>
<td>(Mean age = 41; range = 17 to 67) (N = 24)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.3 Case-control studies

A number of studies use a comparison design in which the rate of sexual victimisation among known sexual perpetrators, such as incarcerated sex offenders, is compared with that found among a group who are not sexual offenders. Table 5.2 summarises these studies.
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Sexual perpetrators (% victimised)</th>
<th>Comparison group (% victimised)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awad &amp; Saunders (1991)</td>
<td>Male adolescents</td>
<td>Child molesters (N = 45) = 21%</td>
<td>Delinquents (N = 24) = 0%</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sexual assaulters (N = 49) = 4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benoit &amp; Kennedy (1992)</td>
<td>Male adolescents</td>
<td>Child molesters of females (N = 25) = -</td>
<td>Non-aggressive offenders (N = 25) = -</td>
<td>ns (p &gt; 0.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Child molesters of males and females (N = 25) = -</td>
<td>Aggressive offenders (N = 25) = -</td>
<td></td>
</tr>
<tr>
<td>Craissati &amp; McClurg (1996)</td>
<td>Male adults</td>
<td>Child molesters (N = 80) = 51%</td>
<td>Violent offenders (N = 20) = 15%</td>
<td>p = 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rapists (N = 29) = 62%</td>
<td>Property offenders (N = 20) = 5%</td>
<td></td>
</tr>
<tr>
<td>Dhawan &amp; Marshall (1996)</td>
<td>Male adults</td>
<td>Child molesters (N = 16) = 50%</td>
<td>Non-sexual offenders (N = 20) = 20%</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rapists (N = 29) = 62%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fagan &amp; Wexler (1988)</td>
<td>Male adolescents</td>
<td>Sex offenders (N = 34) = 9%</td>
<td>Violent offenders (N = 208) = 1%</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td>Ford &amp; Linney (1995)</td>
<td>Male adolescents</td>
<td>Child molesters (N = 21) = 52.17%</td>
<td>Violent offenders (N = 26) = 17.39%</td>
<td>p &lt; 0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rapists (N = 14) = 17.39%</td>
<td>Status offenders (N = 21) = 13.04%</td>
<td></td>
</tr>
<tr>
<td>Kaplan &amp; Green (1995)</td>
<td>Female adults</td>
<td>Child molesters (N = 11) = 81.8%</td>
<td>Non-sex offenders (N = 11) = 45.5%</td>
<td>-</td>
</tr>
<tr>
<td>Lightfoot &amp; Evans (2000)</td>
<td>Male and female children</td>
<td>Coercive sexual behaviour group: (N = 20) = 15%</td>
<td>Non-sexually abusive group: (N = 20) = 15%</td>
<td>-</td>
</tr>
<tr>
<td>Marshall &amp; Mazzucco (1995)</td>
<td>Male adults</td>
<td>Child molesters (N = 24) = 41.7%</td>
<td>Non-sex offenders (N = 23) = 8.7%</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td>Zgourides et al. (1997)</td>
<td>Male adolescents</td>
<td>Sex offenders (N = 80) = 57%</td>
<td>Non-sex offenders (N = 96) = 19%</td>
<td>p &lt; 0.001</td>
</tr>
</tbody>
</table>

"-" Indicates data is not given in the article
5.3.1 Association

There is reasonably conclusive evidence of an association between sexual victimisation and child molestation. Three studies focus on male adults who have abused children (Craissati & McClurg, 1996; Dhawan & Marshall, 1996; Marshall & Mazzucco, 1995) and two find a significantly higher rate of sexual abuse among the molesters than the comparison groups (Craissati & McClurg, 1996; Marshall & Mazzucco, 1995). The study of Dhawan and Marshall (1995) failed to find a difference. This found that a group of child molesters and rapists report a significantly higher rate of victimisation than non-sexual offenders, but, when the groups are considered separately, the association for child molesters is non-significant.

Four studies compare sexual victimisation rates among children or adolescent abusers of children with other offenders (Awad & Saunders, 1991; Benoit & Kennedy, 1992; Ford & Linney, 1995; Lightfoot & Evans, 2000), and two find significantly higher rates among the sexual perpetrators (Awad & Saunders, 1991; Ford & Linney, 1995). Two studies use a heterogeneous group of adolescent perpetrators (Fagan & Wexler, 1988; Zgourides, Monto, & Harris, 1997). Both find evidence of an association between sexual victimisation and sexual perpetration.

The one study that focuses exclusively on females compared a small number of child molesters (N = 11) and non-sex offenders (N = 11) (Kaplan & Green, 1995). The study found higher rates of sexual victimisation among the molesters (81.8%) than the non-perpetrators (45.5%); no statistical analysis was carried out.

5.3.2 Non-spuriousness

It is important to assess whether the association reported above are non-spurious. Most of these studies ensure that the perpetrator and comparison group are the same gender and are similar in terms of age. Most also ensure that the groups are
comparable on an additional variable, such as socio-economic status and incarceration status. But few make further attempts to address this criterion.

A number of studies, for example, identify significant differences on potential confounds, but fail to control for them when assessing the association between sexual victimisation and perpetration. The study by Craissati and McClurg (1996) compared child molesters (N = 80), non-sexual violent offenders (N = 20) and property offenders (N = 20). The molesters were significantly more likely to report sexual victimisation (p = 0.001), but were also more likely to report physical abuse (p = 0.001), being bullied (p = 0.001) and having no friends (p = 0.001). The article does not attempt to statistically control for these differences. The same criticism applies to other studies (Dhawan & Marshall, 1996; Ford & Linney, 1995).

Some studies have statistically controlled for other putative risk factors, but interpretation is still difficult. Zgourides, Monto and Harris (1997) found that adolescent sex offenders (N = 80) were more likely than non-offenders (N = 96) to report sexual and physical abuse (p < 0.001). Although the authors used a logistic regression, sexual abuse was entered before the entry of physical abuse, and so the independent contribution of sexual abuse was not assessed.

Two studies provide some evidence of non-spuriousness because these found comparability between the groups on a number of other putative risk factors. In the Awad and Saunders (1991) study, child molesters (N = 45), sexual assaulters (N = 49) and juvenile delinquents (N = 24) were compared on measures of physical abuse, social isolation, and antisocial behaviour; on these variables the groups did not differ significantly from each other (p > 0.05) Marshall and Mazzucco’s (1995) finding that the rates of sexual victimisation are higher among child molesters (N = 24) than non-perpetrators (N = 23) is strengthened because the two groups were found not to differ on a range of other maltreatment variables, including family violence and parental rejection.
5.3.3 Temporal precedence

None of the studies enquired about the temporal relationship between risk and perpetration. It is worth re-emphasising that even if they had done so the evidence for this criterion would be weak, because of the reliance on retrospective reports of temporal order.

5.3.4 Summary

The case-control studies have found evidence of an association between the experience of sexual victimisation and sexually perpetration against children. The most basic requirement of non-spuriousness is also met by most of the studies, but beyond this the criterion has not been comprehensively addressed. The studies rely on the uncorroborated self-reports of sexual victimisation made by identified perpetrators, and this leads to additional difficulties for the non-spuriousness criterion. It remains possible that any difference is a result of the perpetrator group inventing experiences of sexual abuse to explain their behaviour. The retrospective nature of these studies also mean that the criterion of temporal precedence cannot be addressed; indeed no study enquired about the relationship between the experience of the risk factor and the onset of perpetration.

5.4 Cross-sectional studies

The cross-sectional studies establish the proportion of a sample self-reporting sexual victimisation, the proportion self-reporting sexual perpetration (or sexual interest in children), and the association between the two. These studies tend to use non-forensic populations, such as university students. The results of the eight studies that have used this design are summarised in table 5.3. An additional cross-sectional study used a forensic population and so relied on official reports rather than self-
reports of sexually abusive behaviour (Glaser et al., 2001); this study is also summarised in table 5.3.

Table 5.3 – Cross-sectional studies examining sexual victimisation among sexual perpetrators

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Definition of victimisation and perpetration</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bagley et al.</td>
<td>Males 18 - 27 yrs old</td>
<td>Sexual abuse before age 16 and: Sexual interest/activities with children aged &lt; 13</td>
<td>p &lt; 0.01</td>
</tr>
<tr>
<td>(1994)</td>
<td>Non-students N = 750</td>
<td>Sexual interest/activities with males aged 13-15</td>
<td>p &lt; 0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sexual interest/activities with females aged 13-15</td>
<td>p &gt; 0.01</td>
</tr>
<tr>
<td>Borowsky et al.</td>
<td>Male &amp; female Students N = 665842</td>
<td>Males subjects who forced sex on another and were sexually abused by a family member:</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>(1997)</td>
<td></td>
<td>Male subjects sexually who forced sex on another and were sexually abused by a non-family member:</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female subjects who forced sex on another and were sexually abused by a family member:</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female subjects who forced sex on another and were sexually abused by a non-family member:</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Duane et al.</td>
<td>Male &amp; female Students N = 958</td>
<td>Childhood rape or oral sex and: Forced oral sex on another:</td>
<td>p &lt;= 0.0001</td>
</tr>
<tr>
<td>(1997)</td>
<td></td>
<td>Forced sexual intercourse on another:</td>
<td>p &lt;= 0.0001</td>
</tr>
<tr>
<td>Fischer (1992)</td>
<td>Male &amp; female Students N = 796</td>
<td>Perpetration: Sexual acts involving physical force, the threat of physical force or verbal coercion.</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Using original definition of victimisation:</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Using altered definition of victimisation:</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td>Fromuth et al.</td>
<td>Males Students N = 582</td>
<td>Relationship between child sexual abuse and incest or paedophilia</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td>(1991)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glaser et al.</td>
<td>Males Forensic group N = 747</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stevenson &amp; Gajarsky (1992)</td>
<td>Male &amp; female Students N = 209</td>
<td>Males:</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females:</td>
<td>ns</td>
</tr>
</tbody>
</table>

5.4.1 Association

The Bagley, Wood and Young (1994) study obtained self-reports of sexual interest or activities involving children in a sample of 750 young males aged 18 to 27. An
association between reports of unwanted sexual acts (none, single, multiple) prior to 17 years of age was found to be significantly associated with sexual interest in or activity with females. This study can only provide an indirect test of the victim-to-perpetrator hypothesis, because it combines sexual interest with activities involving children.

Fromuth, Burkhart and Webb Jones (1991) provide a more direct test. In this study the outcome measure is a self-report by male adolescents (N = 582) of contact or non-contact sexual abuse of a child. This study provides partial support for an association. No difference was found in the proportion of perpetrators and non-perpetrators reporting sexual victimisation when the original definition of sexual victimisation was used. This definition, based on the one developed by Finkelhor (1979), excluded the possibility of sexual abuse by an individual aged less than 16 years. Many of the perpetrators in the sample were abused at a young age, and so they could have been abused by people who were substantially older than them, but who were not sixteen. When the researchers altered the definition to permit abuse by individuals younger than sixteen, the rate of sexual victimisation among perpetrators and non-perpetrators was significant (Yates's correction $\chi^2 = [1, N = 582] = 4.42, p < 0.05$). Similar definitions to this have been used in other studies, but the findings must be treated with some caution, because the definition was altered post hoc.

Glaser et al. (2001) examined the relationship between child sexual abuse and incest or paedophilia in a large sample (N = 747) of males referred to a forensic psychotherapy service. Of the 612 males who did not report a history of child sexual abuse, 76% (N = 466) were not perpetrators and 24% (N = 146) were perpetrators. In comparison, of those who were victimised, 41% (N = 56) were not perpetrator and 59% (N = 79) were. This produced an odds ratio of 4.5 (95% CI = 3.05 to 6.55).

In four of the studies the definition of perpetration does not specify the age of the victim; instead perpetration is defined solely in terms of a coercive sexual act. Stevenson and Gajarsky (1991) used a self-report measure of sexual perpetration
designed by Koss and Oros (1982), but they adapted the measure to make it applicable to both male and female students, and, it appears, perpetration against victims of any age. This measure defined sexual perpetration as acts involving the use of drugs, alcohol, continual arguments, a position of authority, threat of force or use of force. Sexual victimisation was defined as experiences ranging from non-contact behaviours (invitation to do something sexual, exhibitionism) to sexual intercourse before the age of 16. The study found that male college students (N = 93) who reported committing sexually abusive acts were significantly more likely to have been sexually abused than those who did not self-report perpetration ($\chi^2 [1, N = 93] = 11.75, p < 0.001$). The relationship between sexual victimisation and perpetration was not significant for females, but details of the analysis are not reported.

Duane, Stewart and Bridgeland (1997) reported similar results. The study found that male college students (N = 375) who reported experiencing forced oral or genital sexual intercourse, were significantly more likely to report forcing oral sex (correlation = 0.32, $p \leq 0.0001$) and forcing genital sexual intercourse on another person (correlation = 0.43, $p \leq 0.0001$). Female students (N = 583) who reported sexual victimisation were significantly more likely to report forcing oral sex on another ($p < 0.001$), but the correlation was only 0.13; this group was not more likely to report forcing genital intercourse on another (correlation = 0.07, ns). Caution is needed, however, because the authors fail to state which statistic was used to provide the measures of association, apart from stating it was a correlation.

The Borowsky, Hogan and Ireland (1997) study uses data from the Minnesota Student Survey, a self-administered anonymous questionnaire completed by male and female students (N = 71594). In the survey, perpetration was measured by the question: ‘Have you ever forced someone into a sexual act with you?’ The study examined intrafamilial and extrafamilial sexual abuse separately, and both were found to be associated with reports of sexual perpetration in males (intrafamilial: $\chi^2 = 962.99, \text{df} = 33669, p < 0.001$; extrafamilial: $\chi^2 = 896.26, \text{df} = 33699, p < 0.001$)
and females (intrafamilial: $X^2 = 129.73, df = 35081, p < 0.001$; extrafamilial: $X^2 = 149.42, df = 35062, p < 0.001$).

Of the studies to have used this design, only Fischer's (1992) study of male and female college students ($N = 796$) does not find any evidence of an association. It may be worth examining whether this study suffers from those limitations that could decrease the likelihood of finding a significant association. One possible limitation is the failure to calculate separate statistics for males and females. Studies that have reported separate analyses have tended to find a substantially stronger association for male victims than females. It is possible that the non-significant finding could have obscured a significant relationship for male victims of sexual abuse.

### 5.4.2 Non-spuriousness

As with the case-control design few of the studies make an attempt to establish whether the observed relationship is non-spurious. In this respect the studies by Bagley, Wood and Young (1994), and Borowsky, Hogan and Ireland (1997) are important. The Bagley, Wood and Young (1994) study obtained self-reports of sexual victimisation and emotional abuse, and these were found to correlate ($0.41, p < 0.001$). Regression was used to examine whether sexual victimisation predicted sexual interest or activity with children once emotional abuse was controlled for. In this regression the duration of sexual abuse did predict the outcome ($0.15, p < 0.05$), but the measure of its severity did not ($0.05, p > 0.05$). As the authors recognise, the amount of variance explained by sexual victimisation is small. In addition, the outcome measure in this study included sexual interest in children, rather than sexual activities exclusively.

The study by Borowsky, Hogan and Ireland (1997) used a logistic regression to control for a number of other putative causes, including physical abuse, witnessing abuse, antisocial behaviour (measured by gang membership) and discontinuity of care (measured by family structure: one parent vs. two parents). The relationship
between sexual victimisation and perpetration remained significant, even when these variables had been controlled for. Male subjects who had being sexually victimised by an adult family member were more likely than non-victimised males to report perpetration (Adjusted OR = 2.71, 95% CI = 1.95 to 3.78), as were those who were victimised by a non-family adult (Adjusted OR = 2.36, 95% CI = 1.82 to 3.06). The same was true of females abused by an adult family member (Adjusted OR = 1.53, 95% CI = 1.11 to 2.10) and those abused by non-family members (Adjusted OR = 1.51, 95% CI = 1.16 to 1.97).

5.4.3 Temporal precedence

As with the case-control design, the cross-sectional studies that have examined the relationship between sexual victimisation and sexual perpetration are retrospective, and therefore cannot effectively address the temporal precedence criterion.

5.4.4 Summary

Of the seven cross-sectional studies to focus on acts of perpetration only one fails to find any evidence of an association (Fischer, 1992). Few studies have attempted to establish that the association is non-spurious. The Borowsky, Hogan and Ireland (1997) study is an exception; it controls for a number of other putative causes and found that the association still remains. It is impossible to make firm any conclusions about this criterion on the basis of one study. The cross-sectional study design is retrospective, and this leads to a number of inevitable limitations. The measure of sexual victimisation in these studies relies on the uncorroborated reports of perpetrators, and it remains possible that they invent these experiences to explain their current functioning. In addition, the design cannot address the criterion of temporal precedence. A more compelling assessment of the causal status of the victim-to-perpetrator hypothesis requires a longitudinal design.
5.5 Longitudinal studies

In a longitudinal design, a group of sexually victimised people and one or more comparison groups are followed prospectively to identify the proportion of each group that perpetrates sexual abuse.

Widom and Ames (1994) used a catch-up longitudinal design (Robins, 1966) in their study of the criminal consequences of childhood sexual victimisation. (Chapter 9 provides a detailed discussion of the catch-up design.) In this study, three groups of maltreated children (sexually abused, physically abused and neglected) were identified using cases processed through the courts between 1967 and 1971. All of the subjects were 11 years or younger at the time of the maltreatment. A control group, matched on sex, age, race and socio-economic status, was also recruited. Official arrest records for rape or sodomy obtained in adolescence or adulthood were used to assess subsequent perpetration. Only male subjects were used in the assessment, because only males can be convicted of these offences.

The catch-up design does much to avoid the possibility of perpetrators inventing sexual victimisation. In this design, the report of victimisation is likely to be made in childhood. Unless we assume that a perpetrator began abusing in childhood, and, while still a child, invented an experience of sexual victimisation to explain their perpetration, the catch-up design does avoid this confound. The level of cognitive sophistication required makes this scenario unlikely. The catch-up design cannot conclusively avoid the possibility of invention, but it is an improvement on the retrospective approach.

The same is true of the temporal precedence criterion. Unlike the retrospective approach, this design can provide an accurate timing of sexual victimisation. It does remain possible that unidentified perpetration preceded the victimisation, but in many cases this would require a very early onset in childhood.
Widom and Ames (1994) used a logistic regression to examine whether sexually victimised males were at heightened risk of arrest for rape or sodomy. In the regression, race, sex and type of maltreatment (sexual, physical and neglect) were the explanatory variables. Sexual abuse did not predict arrest for a violent sex crime (Adjusted OR = 7.6, p > 0.05).

One interpretation of this finding is that if a design removes the possibility of perpetrator invention then the association disappears. However, a closer look at this study indicates that this conclusion may be unjustified. As described above, the victimisation-to-perpetration link was examined for male subjects only. The study does not provide details of the number of males in the sample who were sexually victimised, nor the number of sexually victimised males arrested for a sex crime. However, an estimate of these two numbers can be made on the basis of other information provided in the paper. The authors report that the sexual abuse group (N = 153) was predominantly female with less than a fifth being male. If less than one fifth of the 153 are male, then at most there are 30 sexually victimised males (30 / 153 = 19.6%, but 31 / 153 = 20.3%). The authors also state that 0.7% of the sexually victimised subjects (N = 153) had a conviction for a violent sexual crime (rape or sodomy), which means that only one of the sexually victimised subjects had committed a sex crime. This can only be a male victim, and therefore 1 out of the 30 (3.3%) sexually victimised males had such a conviction. Given the small number of sexually victimised males in the sample and the low base rate of convictions for rape or sodomy, it would be inappropriate to conclude that sexual victimisation does not have a causal role in sexual perpetration solely on the basis of this study.

5.6 Characteristics of sexual victimisation

The term child sexual abuse covers a variety of experiences, and it is possible that some types increase the risk of perpetration more than others. The literature on child maltreatment suggests that maltreatment varies along a number of dimensions,
including the severity of the act, frequency, duration, age at onset, relationship to perpetrators and number of perpetrators (e.g. Barnett, Manly, & Cicchetti, 1993). For child sexual abuse the gender of the perpetrator has also been suggested as a possible moderator of long-term outcomes (e.g. Burton, Miller, & Shill, 2002).

The relationship between these victimisation characteristics and the risk of perpetration has remained largely unexplored. Skuse et al. (1998) compared these characteristics in a group of sexually victimised adolescent males who had sexually abused others (N = 11) with a sexually victimised but non-perpetrating group (N = 14). This study found no difference between the two groups in terms of severity (measured by presence or absence of penetration) ($\chi^2 = 0.00$, df = 1, p = 0.97), duration, (Mann-Whitney U test $z = 0.71$, p = 0.48) whether the abuse was within or outside the family ($\chi^2 = 0.00$, df = 1, p = 0.97), and the number of perpetrators involved (Mann-Whitney U test $z = 0.40$, p = 0.69). However, the small sample size might account for these findings.

Burton, Miller and Shill (2002) examined a number of victimisation characteristics in 216 sexually victimised adolescent sexual offenders and 93 sexually victimised adolescents who were non-sexual offenders. The sexual offenders were more likely to report being sexually victimised by a family member (relative or parent) ($\chi^2 = 24.04$, df = 3, p < 0.001). A large number of the sample was abused by both males and females, and the sexual offenders were more likely to be abused by both genders ($\chi^2 = 39.19$, df = 2, p < 0.001). The sexual offenders were also more likely than the non-sexual offenders to report that their perpetrators used physical force ($\chi^2 = 15.9$, df = 2, p < 0.001). They also reported a longer duration of victimisation (mean = 5.6 years) than the non-sexual offenders (mean = 3.9 years) ($t = -2.13$, p = 0.034). The age at onset did not differ between the two groups. The sexual offenders reported a mean age of onset of 8.8 years (sd = 3.97) compared to 7.6 years (sd = 3.40) for the non-sexual offenders (details of the statistical test are not given in the article). In a logistic regression, in which all the significant variables were entered in one step,
two variables remained significant predictors: gender of perpetrators being both male and female, and being victimised by force.

5.7 Continuity between victimisation and perpetration

All of the studies reviewed in this chapter identify victims of sexual abuse who do not appear to have committed sexual offences. Although it remains possible that there is underreporting of perpetration in the non-perpetrating groups, the findings do suggest that sexual victimisation alone is unlikely to be a sufficient cause of perpetration; other risk factors must be involved.

The hypothesised continuity between victimisation and perpetration is thought to be greater in male than female victims of sexual abuse. Those studies that have compared the strength of association in both types of victim provide support for this assumption. The correlations reported by Duane, Stewart and Bridgeland (1997) for male subjects are 0.32 (forced oral sex on another) and 0.43 (forced rape on another), compared to 0.13 (forced oral sex on another) and 0.07 (forced rape on another) for females. In the Stevenson and Gajarsky (1992) study, the relationship between sexual victimisation and sexual perpetration was significant for males ($\chi^2 [1, N = 93] = 11.75, p < .001$), but not for females (ns, $\chi^2$ statistic not reported). The Borowsky, Hogan and Ireland (1997) study finds a significant adjusted association ($p < 0.001$) for both male and female victims, but the adjusted odds ratios are larger for male (2.36 and 2.71) than female victims (1.51 and 1.53). These findings suggest that if there is continuity between victimisation and perpetration it is likely to be stronger for males than females.

5.8 Summary

The available studies fall short of establishing a causal relationship between sexual victimisation and perpetration, but there is at least some evidence for some of the causal criteria.
The case-control studies and cross-sectional studies do provide some evidence for the association criterion, with the majority of studies finding evidence of a significant association between victimisation and perpetration. However, few studies have attempted to establish that any observed relationship is non-spurious. The possibility that perpetrators invent sexual victimisation to explain their current behaviour is a serious threat to the non-spurious criterion that cannot be addressed by either of these designs. A longitudinal design is needed to address this problem and to establish temporal precedence. The one study to use this design failed to find a relationship, but it is difficult to make conclusions on the basis of this study because of the low base rate of sexual perpetration.

The studies that have compared the association between victimisation and perpetration in males and females provide suggests that the continuity is likely to be stronger for male than female victims. One important additional conclusion is that the relationship between victimisation and perpetration is not inevitable. A number of subjects report victimisation but have not developed an abusive pattern of behaviour. This suggests that other risk factors must be involved. The next chapter will address the question of what these additional risk factors might be.
6 Other risk factors

6.1 Introduction

Sexual victimisation may be a risk factor for sexually abusive behaviour, but, as the previous chapter suggested, it is unlikely to be a sufficient cause in itself. Research has identified a range of other possible risk factors, these include: physical abuse, domestic violence, emotional maltreatment, physical neglect, discontinuity of care, poor social relationships and antisocial behaviour (Fagan, Wise, Schmidt, & Berlin, 2002; Starzyk & Marshall, 2003; Veneziano & Veneziano, 2000).

The vast majority of previous research has not compared these additional risk factors in sexually victimised males who have sexually abused others with those who have not done so. Therefore, this chapter relies on studies examining additional risk factors of sexual perpetrators, regardless of the sexual victimisation status of the sample.

As in the previous chapter, studies examining a particular risk factor are categorised by their design: retrospective studies of forensic populations lacking a comparison group, retrospective case-control studies of forensic populations, cross-sectional studies of non-forensic populations and longitudinal designs. Many of the risk factors have received considerably less attention than sexual victimisation and few of these factors have been examined by all four of the designs.

The causal status of these risk factors is assessed using the evaluative framework outlined in chapter 4. The association criterion is examined for all of the studies. The criteria of non-spuriousness and temporal precedence are then examined for those studies that do report an association. The criterion for establishing a psychological mechanism is discussed in the following chapter.
6.2 Physical abuse

Physical abuse as a risk factor for sexual perpetration has received nearly as much attention as child sexual abuse. Studies that have not used a comparison group report prevalence rates of physical abuse among sexual perpetrators ranging from 22.2% to 92%. These studies are summarised in table 6.1.

Table 6.1 – Proportions of sexual offenders reporting physical abuse

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adler &amp; Schutz (1995)</td>
<td>Male sibling incest offenders (Mean age = 16; range = 13 to 19) (N = 12)</td>
<td>92</td>
</tr>
<tr>
<td>Glaser (1988)</td>
<td>Male child molesters (Range = 17 to 60+) (N = 36)</td>
<td>22.2</td>
</tr>
<tr>
<td>Graham (1996)</td>
<td>Male incest offenders (Natural father) (Mean age = 48) (N = 71)</td>
<td>46.8</td>
</tr>
<tr>
<td></td>
<td>Male incest offenders (Stepfather / relative) (Mean age = 42) (N = 74)</td>
<td>56.76</td>
</tr>
<tr>
<td></td>
<td>Paedophile (Mean age = 46) (N = 78)</td>
<td>43.59</td>
</tr>
<tr>
<td>James &amp; Neil (1996)</td>
<td>Male sex offenders (Mean age = 14.8; sd = 1.7) (N = 31)</td>
<td>41.9</td>
</tr>
<tr>
<td>Kaplan et al. (1991)</td>
<td>Male sex offenders (Mean age = 15.4) (N = 264)</td>
<td>42</td>
</tr>
<tr>
<td>Prentky et al. (1989)</td>
<td>Male sex offenders (Mean age = 30; range = 18 to 68) (N = 82)</td>
<td>66</td>
</tr>
<tr>
<td>Scheela &amp; Stern (1994)</td>
<td>Male incest offenders (N = 42.65; sd = 11.06) (N = 20)</td>
<td>75</td>
</tr>
</tbody>
</table>

Although these proportions are striking, studies that have used a case-control design, and therefore provide an assessment of the association criterion, provide only equivocal evidence of an association between physical abuse and sexual perpetration.
Lang and Langevin (1991) found higher rates of frequent physical beatings in adult incest offenders (N = 36; 55.3%), homosexual paedophiles (N = 29; 41.4%) and heterosexual paedophiles (N = 66; 33.8%), than a comparison group of community volunteers (N = 50; 0%) (p < 0.0001). In contrast Craissati and McClurg (1996) found no significant differences between the rates reported by adult child molesters (N = 80; 45%), property offenders (N = 20; 10%) and violent offenders (N = 20; 45%). Benoit and Kennedy (1992) also found comparable rates of physical abuse among adolescent abusers of female children (N = 25), abusers of males and females (N = 25), aggressive offenders (N = 25), and non-aggressive offenders (N = 25) (χ² = 1.336, df = 3, p > 0.01). Lightfoot and Evans (2000) found higher rates of physical abuse (60%) among a group of children and adolescents who had exhibited sexually coercive behaviour (N = 20) than a matched clinical group who had not exhibited this behaviour (35%; N = 20). However, the authors report that this difference was not statistically significant (details of the statistical analysis are not reported in the article).

Studies that use a mixed group of sex offenders have also produced equivocal results. Dutton and Hart (1992) compared physically abused and non-abused adult offenders, and found a significant relationship between these experiences for both sexual violence against strangers (OR = 3.30, p < 0.01) and sexual violence against family members (OR = 2.67, p < 0.05). Fagan and Wexler (1988) reported that juvenile sex offenders (N = 34) were more likely than violent offenders (N = 208) to report severe child abuse. However, the only example given is that 35% of sex offenders reported being hit by an object compared to 22% of the violent offenders; a test of statistical significance is not reported. The results of the study by Spaccarelli, Bowden, Coatsworth and Kim (1997) are also inconclusive. No differences were found between the proportion of physical abuse reported by adolescent sex offenders (N = 50) and a high-violence comparison group (N = 106) (details of the analysis are not reported).
Two studies have addressed the issue of non-spuriousness. Zgourides, Monto and Harris (1997) found that 56% of adolescent sex offenders (N = 80) reported being 'hurt by an adult for no reason' compared to 21% of adolescent non-offenders (N = 96) ($\chi^2 = 20.22$, df = 1, $p < 0.001$). In this study the physical abuse measure remained significant when it was entered into a logistic regression after the measure of sexual abuse (Adjusted OR = 2.09, $p < 0.01$). Skuse et al. (1998) compared a group of sexually victimised adolescents males who had sexually abused others (N = 11) with a group of sexually victimised adolescent males who had not done so (N = 14). 90.9% of the sexual perpetrator group had experienced physical abuse compared to 35.7% of the non-perpetrators (Unadjusted OR = 18.0; 95% CI = 1.8 to 184.7). However, the odds ratio was no longer significant when entered into a logistic regression along with witnessing intrafamilial violence and discontinuity of care.

One cross-sectional study has examined the relationship between physical abuse and sexual perpetration. The study by Borowsky, Ireland and Wood (1997) found a significantly higher rate of reports of physical abuse in participants who reported committing sexually abusive acts than those who had not done so (OR = 4.51; 95% CI = 4.03 to 5.09). However, when this variable was entered into a logistic regression along with a number of additional risk factors, the adjusted association was non-significant (Adjusted OR = 1.07; 95% CI = 0.87 to 1.32).

The catch-up longitudinal study of Widom and Ames (1994), discussed in the previous chapter, does find some evidence of a link between the experience of physical abuse and sexual perpetration in males. A logistic regression, controlling for race and age, was used to examine predictors of an arrest for rape or sodomy. Physically abused subjects were found to be at heightened risk (Adjusted OR = 7.6, $p < 0.05$) of an arrest for these crimes compared to controls.
6.2.1 Summary

The available research presents a confusing picture of the role of physical abuse in sexual offending. Studies have found that a high proportion of sexual offenders report experiencing physical abuse, but those studies that have used a comparison group have not consistently found a significant relationship. It should be noted that many of the comparison groups consist of violent offenders, and the experience of physical abuse may play a causal role in their offending behaviour. One possibility is that physical abuse has a non-specific causal role: it increases the likelihood of both violent and sexual offending, and so cannot be used to discriminate between the two types of offenders.

The longitudinal study of Widom and Ames (1994) is not dependent on retrospective reports. This makes it unlikely that the association is because of a retrospective reporting bias in which experiences of physical abuse are invented or elaborated in an attempt to explain or excuse sexual perpetration. The finding suggests that the link can be considered non-spurious in this sense. In addition, the longitudinal nature of the study means that there can be some confidence that physical abuse preceded the onset of sexual perpetration. Physical abuse is the only risk factor to have this level of support.

6.3 Witnessing intrafamilial physical abuse

Recent research has expanded the focus on physical abuse to include those events in which a person observes violence rather than directly experiences it. For example, a study of 100 adolescent sex offenders (Richardson, Graham, Bhaté, & Kelly, 1995) found high rates of parent-parent (53%) and parent-sibling physical violence (28%) in the families of participants.

Spaccarelli, Bowden, Coatsworth and Kim (1997) also reported high rates of parental violence. This study found that 50% of self-reported juvenile sex offenders
(N = 26) and 58% of juvenile arrested sex offenders (N = 24) reported exposure to adult violence in the home. The authors combined these two groups and compared the rates of parental violence with a group of low violence offenders (N = 54) and a group of high violence offenders (N = 106). The sex offender group reported a significantly higher rate of exposure to adult violence involving weapons than the low violent group ($X^2 = 10.78, p < 0.001$); however, the rate of exposure to serious adult violence did not differ between the groups ($X^2 = 1.83, ns$). There were no significant differences between the sex offenders and high violent offenders on these two measures. Fagan and Wexler (1988) found that juvenile sex offenders (N = 34) were less likely than violent offenders (N = 208) to report witnessing parental violence, but fail to report a statistical analysis of these differences. The Skuse et al. (1998) study, described above, compared sexually victimised boys who had sexually abused other children (N = 11) with sexually victimised boys who had not done so (N = 14). The boys who had sexually abused others were more likely to have witnessed intrafamilial violence (81.8%) than the non-perpetrators (35.7%). This remained significant when entered into a logistic regression along with discontinuity of care and physical abuse (Adjusted OR = 39.7; 95% CI = 1.1 to 1472.6).

The cross-sectional study of Borowsky, Hogan and Ireland (1997) finds evidence of a link between witnessing intrafamilial violence and sexual perpetration. A self-report of witnessing abuse predicted self-reports of forcing someone into a sexual act, even when a range of other risk factors were controlled for (Adjusted OR = 1.43; 95% CI = 1.18 to 1.72).

6.3.1 Summary

Compared to direct experiences of physical abuse, the effect of witnessing violence has received less research attention. Of the four studies to have used a comparison group, two found evidence of an association, and the relationship remained significant after controlling for other risk factors. However, one study is equivocal and one study fails to find evidence of an association. Further research is needed to
explore the relationship between this vicarious form of maltreatment and the risk of sexual perpetration.

6.4 Emotional maltreatment

All of the adult male incest offenders (N = 20) in Scheela and Stern’s (1994) study were rated as experiencing emotional maltreatment, and James and Neil (1996) found that 61.2% (19/31) of male juvenile sex offenders were emotionally abused. Although these findings appear to be striking, they are compromised because they do not provide a definition of emotional maltreatment and they do not use a comparison group. There is substantial disagreement about how emotional maltreatment should be defined (McGee & Wolfe, 1991b), and it is possible that these apparently striking findings are simply a result of an over-inclusive definition.

Studies that apply the same definition to a group of sexual perpetrators and a comparison group are better placed to address the role of emotional maltreatment; these studies have tended not to find elevated rates among the sexual perpetrators.

For example, five studies fail to find evidence of an association between emotional maltreatment and sexual perpetration. Ford and Linney (1995) use a measure of parental verbal aggression (Conflict Tactic Scale; Straus, 1979) to compare male adolescent rapists (N = 14), child molesters (N = 21), violent offenders (N = 26) and status offenders (N = 21). The child molester group reported a higher score on this measure (35.43) than the other groups (rapists = 25.86; violent non-sexual offenders = 28.04; status offenders = 26.38), but the difference was not statistically significant. The size of the subgroups indicate that the study may lack statistical power, and given this it would have been useful to have examined the effect size of the difference between the child molesters and the other groups. Unfortunately the authors do not provide details of the standard deviations of the mean scores, which make it impossible to calculate an effect size.
Marshall and Mazzucco (1995) used the Parental Acceptance - Rejection Questionnaire (Rohner, 1976) to obtain self-reports of maternal and paternal rejection experienced by adult child molesters (N = 24) and a comparison group of non-criminal subjects (N = 23). The two groups reported comparable rates on both the maternal and paternal measure; in fact the scores of the comparison group were somewhat higher than those of the molester group.

Craissati and McClurg (1996) compared rates of emotional neglect in adult molesters of children (N = 80), violent offenders (N = 20) and property offenders (N = 20). The authors report that the rates did not differ between the three groups, but fail to provide a definition of emotional neglect or details of the analysis.

Lightfoot and Evans (2000) did report higher rates of emotional maltreatment (60%) among a group of children and adolescents who had engaged in coercive sexual behaviour (N = 20) than a matched clinical group who were not sexually abusive (N = 20; 42%). However, the difference between the two groups was not significantly different.

6.4.1 Summary

The findings from research studies that have explored the relationship between experiences of emotional maltreatment and sexual offending mirror the results of other maltreatment risk factors. Studies that have not used a comparison group report apparently high rates of emotional maltreatment, but studies using a comparison group have produced equivocal findings. In addition, the studies have failed to convincingly address whether the relationship is non-spurious and whether the temporal order of event supports a causal role for emotional maltreatment.
6.5 Physical neglect

Physical neglect has received substantially less research attention than other types of maltreatment (Zuravin, 1991); only two studies have examined the relationship between this experience and sexual perpetration. James and Neil (1996) found that the majority of male juvenile sex offenders experienced neglect (61.2%; 19/31). However, this finding is difficult to interpret because the study lacks a comparison group and no definition of neglect is given. The study of Widom and Ames (1994) did not find that neglected males were at heightened risk of arrest for rape or sodomy, but it is unclear how the reliance on criminal records affected the findings. Given the hypothesised detrimental impact of physical neglect on development (Zuravin, 1991) and the lack of research examining this question, further research does seem warranted.

6.6 Intrafamilial sexual atmosphere

Given the popularity of the victim-to-perpetrator hypothesis, it is perhaps surprising that little attention has been paid to the potential role of vicarious experiences of sexual abuse. This can be defined as experiences such as intrafamilial sexual victimisation, sexual perpetration or other forms of sexual deviation not directly involving the subject as a victim.

A study of adult sex offenders (N = 82) found evidence of sexual deviance in the families of 31% of the offenders (Prentky et al., 1989). Sexual deviation included incestuous activities, promiscuity, sexual parties and the production of pornography, occurring in the immediate or extended family, but not involving the subject.

Studies of sexual victimisation in the parents of adolescent offenders provide another indication of the rate of vicarious sexual abuse among sexual perpetrators. Adler and Schutz (1995) in a study of 12 adolescent incest offenders found that 7
(58%) of the mothers reported a history of being sexually abused; the authors do not indicate whether these were child or adult experiences. Kaplan, Becker and Cunningham-Rathner (1988) interviewed 27 parents or parent surrogates (23 mother figures, 4 father figures) of adolescent incest perpetrators. Thirty percent (N = 8) were sexually abused in either childhood or adulthood. The abuse consisted of childhood victimisation of four parents, and five cases of abuse in adulthood (one parent was abused both as an adult and child). These high rates of sexual victimisation of parental figures may be restricted to incest perpetrators. A further study by Kaplan (Kaplan, Becker, & Martinez, 1990) found higher rates of sexual victimisation among mothers of incest perpetrators (34%, 16/47) than the mothers of non-incest perpetrators (13.6%, 11/81) ($\chi^2 = 6.3, p = 0.01$).

6.6.1 Summary

The possible contribution of sexual experiences other than direct sexual victimisation has not been explored until recently. While the proportion of sex offenders reporting these experiences appear to be high, their causal role has not been examined as yet.

6.7 Discontinuity of care

A detailed picture of the care histories (0 to 18 years) of adult sex offenders (N = 82) is provided by Prentky et al. (1989). The sex offenders in this study, consisting of child molesters and rapists, had been cared for by a mean of 3.17 (sd = 1.44) significant caregivers, and had experienced 2.10 (sd = 1.44) changes in carer. Significant caregivers were defined as adults who had primary responsibility for the person for a year or more; a change in carer was defined as an addition or subtraction of a significant person to or from the person’s life. From birth to age 18, the sex offenders had spent an average of 13.10 years (sd = 5.16) living with the biological mother and 10.16 years living with the biological father. Further
information is provided by Glaser (1988), who found that 17% of adult child molesters (N = 36) had experienced a history of parental separation before the age of 18, and a further 20% had experienced the death of a parent by this age.

A similar picture emerges for adolescent sex offenders. James and Neil (1996) found that 42% of a group of sexual perpetrators (N = 31) were either adopted, fostered, accommodated or in the care of social services. Richardson, Graham, Bhave and Kelly (1995) found that 54 out of 100 adolescent perpetrators no longer had any regular contact with their father, and 58% were from families in which their parents had separated. A small study of adolescent sibling incest offenders (N = 12) found that while most (83%) were living with their parents, there had been previous marital separations in 58% of cases (Adler & Schutz, 1995).

Fagan and Wexler (1988) found that over half of a group of adolescent sex offenders (N = 34) reported living with their biological parents, compared to only 18% of violent offenders (N = 208). In a study comparing adolescent rapists, child molesters, violent non-sexual offenders and status offenders, the living situations (both parents, single parent, parent/stepparent, grandparent/other) did not differ between the groups (Ford & Linney, 1995). In contrast, Lightfoot and Evans (2000) report a higher rate of multiple disruptions to attachment in a group of children and adolescents who had engaged in coercive sexual behaviour (N = 20; 65%) than a clinical comparison group who were not abusive (N = 20; 32%) (p < 0.05).

Skuse et al. (1998) found that sexually victimised boys who had sexually abused others (N=11) were more likely to have experienced discontinuity of care (54.5%) than a group of sexually victimised boys who were non-perpetrators (N=14; 14.3%). However, this variable fell short of significance when entered into a logistic regression along with witnessing intrafamilial violence and physical abuse (Adjusted OR = 15.0; 95% CI = 0.9 to 245.2).
Two cross-sectional studies fail to find evidence of a link. Fromuth, Burkhart and Webb Jones (1991) identified by self-report 16 men who had sexually abused a child from a group of 582 college students. The subjects also reported whether before the age of 16 they had ever lived without their father for more than one year. The authors found that this measure did not differentiate between the perpetrators and non-perpetrators (no data or analysis are presented). In the study by Borowsky, Hogan and Ireland (1997), living with one parent, compared to living with both biological parents or any other living situation, did not predict sexual perpetration when other risk factors were controlled for (Females: Adjusted OR = 0.92, 95% CI = 0.69 to 1.23; Males: Adjusted OR = 0.96, 95% CI = 0.81 to 1.13).

### 6.7.1 Summary

The available research does not suggest that discontinuity of care is a risk factor for sexual offending. Only one study found a significant association, but this was rendered non-significant once other risks were controlled for. However, many of the studies had only small sample sizes; larger studies may be needed to identify significant differences.

### 6.8 Social rejection

A number of researchers have suggested that social rejection and isolation increases the likelihood of sexual perpetration (e.g. Marshall & Barbaree, 1990).

Of the two case-control studies to focus on adult sexual perpetrators, only one finds evidence of an association. Craissati and McClurg (1996) compared rates of contact with friends among adult child molesters (N = 80), violent offenders (N = 20), and property offenders (N = 20). 30% of the child molesters reported that they did not have contact with friends, this compares to figures of 5% for the violent offenders and 0% for the property offenders (p = 0.001). In addition, 45% of the perpetrators
reported being bullied, compared to 5% of the violent offenders and 10% of property offenders (10%) \((p = 0.001)\). The authors fail to report how being bullied was defined or measured. In another study of adult males who abused children (Haywood, Kravitz, Wasyliw, Goldberg, & Cavanaugh, 1996) a difference was not found. This study compared levels of social introversion among cleric child molesters \((N = 24)\), cleric comparisons \((N = 48)\), non-cleric child molesters \((N = 45)\) and non-cleric comparisons \((N = 40)\), using one of the subscales of the Minnesota Multiphasic Inventory (Greene, 1991). The proportions scoring in the clinical range did not differ significantly between the groups.

The two case-control studies of adolescent offenders also fail to find a difference. Ford and Linney (1995) compared male adolescent rapists \((N = 14)\), child molesters \((N = 21)\), violent offenders \((N = 26)\) and status offenders \((N = 21)\), using the Child's Assertive Behavior Scale (CABS; Michelson & Woods, 1982), an instrument designed to measure social skills and social behaviour. The study found no differences between the groups for perceived ability to establish peer relationships; the groups were as likely to report preferring friends of the same age, and the majority believed they had as many friends as most people. Fagan and Wexler (1988) also found similar rates of peer integration among juvenile sex offenders \((N = 34)\) and violent offenders \((N = 208)\). However, the measure was designed for the study and the authors fail to provide details of its psychometric properties, which limits the faith we can place in the non-significant finding.

The one cross-sectional study to address this question also fails to find evidence of a link. Fromuth, Burkhart and Webb Jones (1991) administered a questionnaire to 582 male students \((\text{mean age} = 20)\), of which 3% \((N = 16)\) self-reported sexual perpetration against a child. No relationship was found between sexual perpetration and social isolation, as measured by a self-report of the number of good friends at age 12. The study does not report the data on which the test of significance is made nor the results of the test.
6.8.1 Summary

Although social rejection is one of the most commonly cited explanations of sexual offending, the available research fails to provide strong evidence of a causal role. Only one study finds any evidence of an association, and none of the studies have addressed other causal criteria.

6.9 Intelligence

The intellectual functioning of sexual perpetrators has been assessed using a variety of intellectual or attainment indicators. Glaser’s (1988) study, for example, found that 5.5% of adult males who had abused children (N = 36) were rated as ‘mentally retarded’. Graham (1996) found that the average educational level of adult incest offenders (Natural father: N = 71; Stepfather/relative: N = 74) and rapists (N = 63) was 9th grade, and for paedophiles (N = 78) was 10th grade. Over half of the adolescent sex offenders (N = 34) in the James and Neil (1996) study were reported as being of below average ability; however, this figure appears to be based on the subjective view of social workers or clinicians rather than any formal testing. A substantial proportion (44%) of the adolescent offenders in the Richardson, Graham, Bhate and Kelly (1995) study were statemented under section V of the 1981 Education Act, typically for a combination of learning and behavioural difficulties.

Two case-control studies of adult males who abused children report lower rates of educational attainment among the abusers than the non-perpetrators. The study of cleric and non-cleric child molesters (Haywood, Kravitz, Wasyliw, Goldberg, & Cavanaugh, 1996) found that cleric comparisons (N = 48) had the highest educational level, followed by cleric child molesters (N = 24) and non-cleric normal comparisons (N = 40); non-cleric child molesters had the lowest level (N = 45). Lang and Langevin (1991) found that the comparison group of community volunteers (N = 50) had a significantly higher educational level than the groups of adult heterosexual paedophiles (N = 66), homosexual paedophiles (N = 29) and
incest offenders \(N = 36\) (\(F = 14.80, p < 0.0001\)). Sexual offenders against children \(N = 80\) were also more likely to have been to a special school than violent offenders \(N = 20\) or property offenders \(N = 20\) (Craissati & McClurg, 1996). 15% of the child molesters had been to such a school compared to 5% of the violent offenders and none of the property offenders \(p = 0.005\).

Awad and Saunders (1989) report that 48% of adolescent males who had abused children had a diagnosed learning difficulty and 65% had failed a grade. Although over half (59%) were of average intelligence, 24% were of low average ability and 10% were of borderline intelligence (mean = 91, sd = 10) (WISC-R; Wechsler, 1974). The authors report that the perpetrators were more likely than a non-sexually abusive comparison group matched on age, socio-economic status and time of referral, to have had a severe learning disability (data or a statistical analysis are not reported).

Ford and Linney (1995) compared IQ levels of male adolescent rapists \(N = 14\), child molesters \(N = 21\), violent offenders \(N = 26\) and status offenders \(N = 21\). The groups did not differ in the proportions scoring in the low average range or the borderline/mentally retarded range; however, it is unclear whether a difference would have been found if IQ was treated as a continuous variable.

6.9.1 Summary

There is some evidence of an association between sexual offending and intelligence, but the causal role of such a variable remains unclear. Intelligence is a high order construct; it subsumes within it a range of other variables and mechanisms. It would be necessary to identify the mechanism through which intelligence served as a risk for subsequent sexual perpetration.
6.10 Non-sexual antisocial behaviour

A number of studies have examined the prevalence of antisocial behaviour in the background of sexual offenders. Table 6.2 summarises the findings from those studies that have not used a comparison group.

Table 6.2 – Non-sexual antisocial behaviour among sexual perpetrators

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adler &amp; Schutz (1995)</td>
<td>Male sibling incest offenders</td>
<td>Conduct disorder = 42</td>
</tr>
<tr>
<td></td>
<td>(Mean age = 16) (N = 12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Mean age = 14) (N = 29)</td>
<td>Delinquency record = 24.1</td>
</tr>
<tr>
<td></td>
<td>(Mean age = 41.3) (N = 80)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Range = 17 to 60+) (N = 36)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Mean age = 14.8; sd = 1.7) (N = 31)</td>
<td>Lying = 80.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stealing = 67.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fighting = 32.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arson = 6.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>None of above = 16.1</td>
</tr>
<tr>
<td>Richardson et al. (1995)</td>
<td>Male sex offenders</td>
<td>Theft = 55</td>
</tr>
<tr>
<td></td>
<td>(Age range = 11 to 18) (N = 100)</td>
<td>Burglary = 44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vandalism = 41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shoplifting = 40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arson = 30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Car theft = 27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Truancy = 59</td>
</tr>
</tbody>
</table>

Two case-control studies provide data on the level of antisocial behaviour among sexual perpetrators. Ford and Linney (1995) compared the number of school suspensions (scored: none, one, two, three or more) of male adolescent rapists (N = 14), child molesters (N = 21), violent offenders (N = 26) and status offenders (N = 21). The difference between the groups approached significance (p < 0.10); the
violent offenders had a higher proportion of three or more suspensions (74%) than the other groups (Rapists = 57%; Child molesters = 43%; Status offenders = 57%). Fagan and Wexler (1988) compared juvenile sex offenders (N = 34) and violent offenders (N = 208) and found no substantial differences between the groups in terms of juvenile court records. The authors also obtained self-reports of deviant acts committed in the year preceding interview. The sex offenders reported a mean of six acts, which is 50% fewer than the violent offenders.

The cross-sectional study of Borowsky, Hogan and Ireland (1997) did find some indirect evidence for an association between antisocial behaviour and sexual perpetration. The authors used self-reports of gang membership as a proxy measure of antisocial behaviour. A self-report of currently being a member of gang predicted sexual perpetration and this difference remained significant in the logistic regression controlling for other putative risk factors (OR = 2.20; 95% CI = 1.78 to 2.71).

6.10.1 Summary

Although the two case-control studies tended to find somewhat lower rates of antisocial behaviour committed by sexual perpetrators than the comparison groups, it should be noted that the rates are high for the sexual perpetrators. Over 40% of the child molesters had been suspended from school three times or more in the Ford and Linney (1995) study, and the sex offenders in the Fagan and Wexler (1988) study self-report committing a mean of six deviant acts in the preceding year. However, this high rate of non-sexual antisocial behaviour could have more to do with the samples being drawn from forensic populations, than an indication that generalised antisocial behaviour is a risk factor for sexual perpetration. The cross-sectional study of Borowsky, Hogan and Ireland (1997) can address this problem because it uses a non-forensic sample of sexual perpetrators. It found an association between self-reports of sexual perpetration and a proxy measure of antisocial behaviour. Although the authors controlled for a number of other putative risk factors, this association could still be spurious because of the reliance on retrospective self-reports. Those
subjects who are willing to self-report committing a sexually abusive act may also be more willing to self-report other antisocial characteristics such as being a gang member.

As with intelligence, antisocial behaviour is a high level construct. If antisocial behaviour is a risk factor, this leads to a further set of questions about the mechanisms involved. One potential use of the construct is as a moderator in a multicausal theory. A multicausal theory suggests that there are different pathways to sexual perpetration. Watkins and Bentovim (1992) argue that one pathway involves sexual perpetration as part of a wider pattern of antisocial behaviour, whereas other paths are unrelated to pervasive antisocial behaviour.

6.11 Psychopathology

Psychopathology has also been implicated in the development of sexual offending. Glaser (1988) in a study of 36 incarcerated adult child molesters found that 4 (11.1%) of the offenders suffered from schizophrenia and 4 (11.1%) suffered from an affective disorder. James and Neil (1996) report that only 25.8% of adolescent sex offenders (N = 31) reported no psychological problems; 38.4% suffered from anxiety, 29% suffered from depression and 12.9% had phobias.

Spaccarelli, Bowden, Coatsworth and Kim (1997) compared adolescent sex offenders (N = 50) to low violence (N = 54) and high violence (N = 106) groups on measures of depression (Children’s Depression Inventory; Kovacs, 1981), anxiety (Revised Children’s Manifest Anxiety Scale; Reynolds & Richmond, 1978) and PTSD-related symptoms of rumination (Trauma Symptom Checklist: Lanktree & Briere, 1990; Intrusive Thought Scale: Horowitz, 1979) and dissociation (Trauma Symptom Checklist; Lanktree & Briere, 1990). The sex offenders did not differ significantly from the two groups on these measures.
Haywood, Kravitz, Wasyliw, Goldberg and Cavanaugh (1996) compared rates of psychopathology (Minnesota Multiphasic Inventory; Greene, 1991) in cleric child molesters (N = 24), cleric comparisons (N = 48), non-cleric child molesters (N = 45) and non-cleric comparisons (N = 40). The cleric child molesters were significantly more likely than the cleric comparisons to have elevated rates on the scales measuring Depression, Hysteria, Psychopathic Deviate, Psychasthenia, and Schizophrenia (p <= 0.001). The non-cleric child molesters differed from the non-cleric comparisons in that they had elevated rates on the measures of Psychopathic Deviate and Schizophrenic Scales (p <= 0.001).

One cross-sectional study (Bagley, Wood, & Young, 1994) has examined the relationship between a range of mental health symptoms and sexual interest in or activities involving children. The study used the Mental Health Questionnaire (Crowne & Crisp, 1979), the CES-D Depression Scale (Radloff, 1977), the Trauma Symptom Checklist (Briere, 1989), and a measure of suicidal ideation and behaviour (Ramsay & Bagley, 1985). Each of the measure significantly predicted sexual interest or activities in a multiple regression (p < 0.05), but the amount of variance explained by these measures was small.

6.11.1 Summary

A variety of mental health symptoms have been found to be associated with sexual offending. However, as with the previous two risk factors, psychopathological constructs are higher order variables. It would be necessary to carefully delineate what aspect of a particular psychiatric diagnosis acted as the mechanism that increases the risk of sexual perpetration. In addition, the possibility that some of these psychological problems are a result of being identified as a sexual perpetrator cannot be addressed in retrospective designs.
6.12 Conclusion

Although the putative risk factors discussed in this chapter have often been implicated as causes of sexual offending, the available research provides at best very limited support for the causal role of these variables. The most basis causal criterion, the identification of an association between a risk and sexual perpetration, has not been consistently met for any of the putative risks. However, as discussed in chapter 4, the absence of an association does not necessarily imply that a variable is not causal. Methodological features of a study may mask a relationship, and the research studies discussed in this chapter often have methodological limitations.

One limitation is the use of dichotomous measures, particularly of maltreatment risk factors. A number of researchers have criticised this approach to measurement, because it could artificially deflate an association with the outcome variable (Cicchetti & Barnett, 1991; Newcombe & Locke, 2001; Wolfe & McGee, 1994). The research reported in this thesis compares dichotomous with ordinal and continuous measures of maltreatment.

The absence of an association between a risk factor and sexual offending could also be due to a third variable. For example, a study that compares the rates of exposure to domestic violence in a group of sex offenders and a group of violent offenders may find no difference between the rates of exposure. However, it is possible that domestic violence is causally implicated in both violent and sexual offending, and a third variable acting on either of the groups determines the difference in offending style. None of the studies have addressed this possibility by specifying these third variables.

The majority of studies have not examined the criterion of non-spuriousness. The possibility that an association between a putative risk and sexual perpetration is a result of a correlation between that putative risk and a genuine causal variable has not been systematically addressed. Few studies statistically control for other risk
factors, and in those studies that do this is typically limited to a small number of variables. An additional problem for the non-spuriousness criterion is that the vast majority of the studies have used a retrospective design. It remains possible that any observed association is the result of perpetrators inventing or elaborating risk experiences in an attempt to excuse their behaviour.

The reliance on a retrospective design also means that the studies discussed in this chapter cannot establish the temporal relationship between the putative risk factor and sexual perpetration. This is particularly problematic for a number of the risks, because these could feasibly be a result of, rather than a cause of, sexual perpetration.

The available research fails to provide convincing evidence of the causal role of the risk factors discussed in this chapter; basic methodological improvements are needed. Chapter 9 will suggest that a catch-up longitudinal design is able to overcome these limitations. Before that, the 4th causal criterion, the identification of a causal mechanism, will be discussed in the next chapter.
7 Theories of sexual perpetration

7.1 Introduction

The criteria of association, non-spuriousness and temporal precedence are necessary but not sufficient to establish a causal relationship between a risk factor and an outcome, such as sexual perpetration (Haynes, 1992). It is also necessary to identify a logical (Haynes, 1992) or plausible mechanism (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997) to account for the relationship. Given the ease with which plausible psychological mechanisms can be generated (see chapter 4), it is usually necessary to provide evidence of a mechanism, rather than merely suggest one. Theories of sexual perpetration provide hypotheses about these mechanisms; this chapter will evaluate these theories.

As discussed in chapter 4, the theories will be assessed by evaluating the 'riskiness' of their predictions. A central problem of theory testing in non-experimental psychology is that the theories make unrisky, directional predictions, which are corroborated using null hypothesis significance tests. Even a theory that is drivel has a fair chance of clearing this low hurdle, and therefore this method does little to establish the verisimilitude of a theory. In contrast, a theory making risky predictions gets considerable 'money in the bank' (Meehl, 1990a) if there is empirical confirmation of its predictions.

A number of strategies for making risky predictions were described in chapter 4 and these will be used to evaluate the theories, these include:

- Predictions about the temporal relationship between events
- The number of quasi-independent predictions
- A theory's ability to make one or more predictions that require ad hoc adjustments from competing theories.
Theories that make predictions that are both risky and empirically corroborated will be judged most successful, but few studies make risky predictions and fewer still have empirical evidence to support these predictions. An additional limitation is the reliance on single causal explanations of sexual perpetration, which limits explanatory power. Multicausal theories of sexual perpetration are needed, but these run the risk of becoming unfalsifiable theories of everything. The challenge for theorising in this area is to develop multicausal theories that are capable of making risky predictions.

7.1.1 Limits of review

This thesis is concerned with identifying the causes of someone becoming a sexual perpetrator. A separate question is what factors are involved in the maintenance of this behaviour. As Haynes (1992) discusses, a psychopathological outcome has a number of features (e.g. onset, maintenance etc.), and different causal mechanisms may be involved in each of these. This chapter will review only those theories designed to explain why some people become sexual perpetrators; it will not review theories attempting to account for the maintenance of this behaviour.

The results of treatment studies may provide evidence for or against particular theories of sexually abusive behaviour. However, these studies will not be reviewed because the interpretation of these data is difficult for at least two reasons. The treatment studies seek to alter mechanisms involved in the maintenance of abusive behaviour. This can provide only indirect support for theories designed to explain why some people begin to abuse in the first place. In addition, treatment typically attempts to change a number of the maintaining mechanisms. If treatment is shown to be effective it is unclear which mechanism was effectively altered by the treatment.
7.1.2 Structure of review

Theories that offer a single causal explanation of sexual perpetration will be discussed first, followed by multicausal theories. Single causal theories are broadly grouped into feminist, pornography, psychoanalytic/psychodynamic, family systems, behavioural and learning, cognitive, and psychophysiological theories. Five multifactorial models will also be discussed: Finkelhor’s Four-Factor Model (Araji & Finkelhor, 1986), Marshall and Barbaree’ (1990) Integrated Theory, Hall and Hirschman’s (1992) Quadripartite Model, Ward and Siegert’s (2002) Pathways Theory and a Developmental Psychopathology model.

7.2 Feminist theories

Feminist theories of child molestation stress the importance of the socio-political system, characterised by patriarchy, in which sexual abuse takes place (Solomon, 1992). Patriarchy is a system in which men’s organisation of and control over the societal structure maintains their dominance over women as a class. Violence, including sexual violence, is one method through which this dominance is maintained, and it is integrated within the societal structure. The mechanism by which sexual offending occurs, including that directed towards children, is to be found within this patriarchal system.

This analysis implicates the entire societal structure in the development of sexual offending, but attention has tended to focus on the institution of the family (e.g. Solomon, 1992). Solomon (1992) argues that the family is instrumental in the socialisation into male and female roles, including the learning of different sexual scripts. Male’s sexual scripts emphasise dominance, taking the initiative, the attractiveness of smaller and younger sexual objects, and the misinterpretation of resistance as a cover for sexual desire. In contrast, female sexual scripts expect females to be docile and subservient sex objects. These different scripts tacitly
encourage and condone sexually abusive behaviour by males against females and the young.

The theory makes a number of predictions, particularly that a range of beliefs should be more characteristic of sexually abusive males than a non-abusive comparison group. However, similar predictions are also made by other theories, such as Cognitive Theory, which limits the riskiness of both theories. One strength of the feminist approach is the emphasis it gives to the fact that males outnumber females as perpetrators of sexual abuse. Any adequate theory of sexual perpetration must explain this phenomenon.

7.3 Pornography theories

Pornography has also been implicated in the development of sexual offending (e.g. Marshall, 1989; Marshall, 2000). The model proposed by Marshall (1989) suggests that pornography may have a role in the cause of sexual offending, but one that is neither necessary nor sufficient. Males that are placed at risk of sexual offending by developmental experiences may be most susceptible to the messages contained in pornography, such as the message that children are appropriate sexual partners. Exposure to this material in susceptible males might fuel sexual fantasising, and this might translate into overt behaviour. The model predicts that exposure to pornography will lead to an alteration in cognitions related to sex and an alteration in sexual fantasising.

At first glance, retrospective research comparing exposure to pornography in sexual perpetrators and non-perpetrators seems to question the causal role of pornography. In a review of the literature Bauserman (1996) reported that sexual offenders do not have an earlier date of first exposure to pornography than non-sex offenders, nor do they have a more unusual exposure. However, these finding do not contradict the proposed mechanism. Age of exposure and type of exposure may be the same in
both groups, but developmental experiences may have made the subsequent sexual perpetrators more susceptible to the messages contained in the material.

A large number of experimental studies have examined the relationship between exposure to pornography and changes in cognitive states that may serve as mediators for sexual perpetration (for reviews see Allen, D’Alessio, & Brezgel, 1995; Allen, Emmers, Gebhardt, & Giery, 1995; Seto, Marie, & Barbaree, 2001). However, these studies have focused on cognitions and affective states hypothesised to link exposure to pornography and rape, rather than child sexual abuse (e.g. anger towards adult females, rape myth acceptance). No experimental studies of pornography have examined the cognitive states hypothesised to link pornography and sexual activity involving children (Seto, Marie, & Barbaree, 2001).

Researchers have also examined the relationship between the availability of pornography and the prevalence rate of sexual offending (e.g. Kutchinsky, 1972; Mawhinney, 1998; Page 1989). The interpretation of the data is controversial (Mould, 1990), and in addition it is not able to address hypotheses about psychological mechanisms.

In summary, although there is a large body of research examining pornography and sexual perpetration, it has not addressed the issue of the psychological mechanism proposed to connect exposure to such material with the perpetration of child sexual abuse.

### 7.4 Psychoanalytic and psychodynamic theories

A number of theories of child molestation have been proposed that draw upon psychoanalytical or psychodynamic concepts. One early formulation stated that an offender’s Oedipal Complex, characterised by the person’s internal conflict about their mother figure, made it difficult for them to develop social and sexual
relationships with adult females (Araji & Finkelhor, 1986). Other early formulations also emphasise the inability of child molesters to form adequate relationships with adult females, but trace this back to castration fears. The main problem of theory corroboration for these theories, as with so many psychoanalytic formulations, is the loose deductive chain linking the theory to the statistical hypothesis. This chain is often so loose and requires so many difficult to substantiate assumptions that confirmation of the statistical hypothesis does little to corroborate the theory. This is, of course, a problem for all psychological theorising (Meehl, 1990a), but it seems to cause particular problems for psychoanalytic theories (Meehl 1991a; 1991b).

Although these early theoretical formulations are now largely of only historical interest, there is continued interest in Groth's (1978) psychodynamic model. The model draws upon the psychoanalytic concepts of fixation and regression to propose two developmental pathways to sexual perpetration against children. According to the model, the fixated child molester has experienced an arrestment in socio-sexual maturation, resulting from unresolved formative issues, such as sexual victimisation. Groth proposes that this leads to an exclusive sexual attraction towards children. The regressed offender is primarily attracted to adult women, but exposure to stress in coping with adult demands leads to sexual activity with children.

Because the model suggests two distinct developmental pathways to sexual perpetration, it does make a number of testable, reasonably risky predictions. Perpetrators who are regressed offenders should have a cluster of characteristics that are distinct from those of fixated offenders. Simon, Sales, Kashniak and Kahn (1992) examined this by developing an ordinal measure of perpetrator characteristics designed to range from fixated to regressed. Each characteristic can be considered a quasi-independent prediction; therefore the clustering of these characteristics would provide evidence for the theory. If the theory were correct, the scores of the child molesters would produce a binomial distribution on this measure. As Simon, Sale, Kashniak and Kahn (1992) report, the study failed to find such a distribution.
The theory also predicts an association between the experience of sexual victimisation and primary sexual interest in children. This hypothesis has been examined in a number of studies using phallometric assessment. A study by Lang and Langevin (1991) did not find higher rates of sexual victimisation among sex offenders with a phallometrically-diagnosed attraction towards children than those who had a sexual preference for adults. The results of a series of studies by Freund (Freund & Kuban, 1994; Freund, Watson, & Dickey, 1990) are also equivocal. Even if the findings were in keeping with the prediction, this would not offer strong corroboration of the theory, because other theories (e.g. behavioural) also make this prediction.

Although Groth’s theory may lack empirical support, its emphasis on multiple pathways is important. As will be argued below, there are likely to be different routes to sexually abusive behaviour, and any theory must be able to account for this. Groth’s theory also demonstrates that by proposing multiple pathways the theory can make a number of quasi-independent predictions, which can help increase its ‘riskiness’.

Other psychodynamic theories focus on the hypothesised relationship between sexual victimisation and perpetration. The identification with the aggressor hypothesis (Araji & Finkelhor, 1986) proposes that sexually victimised males who are also emotionally neglected identify with the sexual perpetrator as a result of a need for emotional attention. This may lead the victim to behave like the perpetrator.

The only testable prediction made by this theory is that sexually victimised males that subsequently perpetrate should be more likely to have experienced emotional neglect than victims that do not do so. Briggs and Hawkins (1996) compared retrospective reports of sexually victimised sex offenders and a group of sexually victimised males that were not offenders, and found that the perpetrators were less likely to report receiving physical affection as children. This is consistent with the theory, but given that only one prediction is made this does little to establish its
verisimilitude. In addition, methodological limitations of the Briggs and Hawkins (1996) study mean that there are alternative explanations of the finding. First, the retrospective nature of this study raises the possibility that the association could be a result of current status (perpetrator/non-perpetrator) affecting recall. Secondly, the perpetrator group consisted of convicted child molesters, whereas the comparison group were a non-forensic group of self-referred males who claimed to have been sexually abused and claimed not to have sexually perpetrated. The different selection procedures may account for the differences in the levels of affection experienced by the two groups.

7.5 Family systems theories

The Family Systems approach to sexual perpetration was designed to explain intrafamilial sexual abuse. According to this theory, family members engage in patterned transactions, which reciprocally influence each other in ways that are characteristic and somewhat predictable (Maddock & Larson, 1995). These actions periodically converge in ways that create systems of shared meanings that will guide the future behaviour of family members. From this perspective, all members contribute to dysfunctional behaviour, including incestuous behaviour.

According to Maddock and Larson (1995), incestuous behaviour originates and continues because of substantial boundary disturbances in the family system. A boundary distinguishes a system, such as the family, from its environment, delineating the organisational whole or autonomy of the system. Family Systems Theory argues that incestuous behaviour occurs when the family generates a barrier between itself and society. In families that are isolated from the social environment, family members meet each other’s emotional needs; these families are classified as ‘enmeshed’. The scarcity of emotional resources in these enmeshed families means that family members – regardless of age – are pressed into service to meet the needs of other members, including their sexual needs.
The theory makes one directional prediction: families in which sexual abuse occurs are likely to score higher on a measure of enmeshment than families in which there is not sexual abuse. Although clinical experience (e.g. Maddock & Lawson, 1995) supports the view of incestuous families as enmeshed, one directional prediction does little to establish a theory's verisimilitude.

7.6 Behavioural and learning theories

Early conditioning theories proposed that classical conditioning was involved in the development of sexual offending (McGuire, Carlisle, & Young, 1965; Abel & Blanchard, 1974). These models suggested that a deviant first sexual experience, such as child sexual abuse, combined with a lack of success in peer sexual relationships, resulted in the salience of the deviant experience. This experience subsequently features in masturbatory fantasies, and this pairing leads to a conditioned response in which deviant fantasies lead to sexual arousal.

The theory predicts that deviant sexual experiences should be common in the background of sexual perpetrator. As discussed in chapter 5, there is some evidence for this, but the prediction is by no means specific to the theory. Another prediction is that sexual offenders have experienced failure in peer sexual relationships. Again, this is not a prediction specific to conditioning theory.

The theory also predicts that child molesters will report deviant sexual fantasies in childhood or adolescence and these will precede the first act of sexual perpetration. As argued in chapter 4, evidence of a temporal relationship does increase the riskiness of the predictions made by a theory. High rates of deviant fantasising in adolescence have been reported by adult sex offenders (Abel, Becker, Mittelman, Cunningham-Rathner, Rouleau, & Murphy, 1987). However, in a retrospective
design it is difficult to gauge the temporal relationship between fantasising and the onset of perpetration.

Finally, the theory predicts that child molesters should exhibit sexual arousal to sexually deviant stimuli. A number of studies have found that a proportion of child molesters are sexually aroused by images of children (e.g. Freund, 1987; Marshall, Barbaree, & Butt, 1988; Marshall & Barbaree, 1988), but again these studies cannot address the temporal relationship between events, and this is needed to ensure a risky test of the prediction. It is possible that the repeated acts of sexual perpetration over time lead to a deviant arousal pattern, rather than the other way round as suggested by the theory.

This theory does make a number of testable predictions but all of these can be easily accommodated by other theories. If it were established that the temporal relationship between events was the same as that predicted by the theory, then this would increase the riskiness of the predictions. However, this research is unlikely to be carried out given the associated ethical problems.

Although operant conditioning strategies have been used in the treatment of sexual perpetration (e.g. Cautela & Kearney, 1990; McNally & Lukach, 1991), the theory has not been applied to explaining the development of this behaviour. However, researchers have developed social learning theories of sexual perpetration (e.g. Garland & Dougher, 1990; Greenberg, Bradford, & Curry, 1993). Social learning theory proposes that a person learns a behaviour by observing the modelling of it by significant others (Bandura, 1969; 1977); and this theory has been used to explain why some victims of child sexual abuse subsequently become sexual perpetrators.

Burton, Miller and Shill (2002) tested a range of predictions derived from Garland and Dougher's (1990) social learning theory of sexual offending in sexually victimised males. The study used a case-control design of 309 sexually victimised male adolescents in delinquent treatment facilities. Participants who were referred
for sexual offences or who self-reported sexually abusive behaviour (N = 216) were compared to participants who did not report sexual offending (N = 93).

The researchers predicted that sexually victimised males were more likely to be sexual offenders if their own sexual victimisation began at a young age, had a longer duration, involved force and involved penetrative acts. They also predicted that victimisation by a male would increase the risk, as would being victimised by someone such as a family member. They provided a social learning explanation of each of these predictions. For example, abuse by a family member increases risk because the victim would be more likely to attend to, learn from and later model that person's victimising behaviour. Each of the predictions, with the exception of age of onset, was supported by the data.

Although the predictions of social learning theory are only directional, the theory does have the advantage of making a number of predictions. This is impressive only if the predictions are quasi-independent (Lykken, 1991), but the ones tested by Burton, Miller and Shill (2002) do not seem to be. Many of them simply restate the prediction that as severity increases - whether defined in terms of duration, penetration or force - so too will the risk of perpetration. Even if this were not the case, most of the predictions can be reduced to the bland truism that in psychology bad things tend to go with bad things. If all the predictions take this form then, no matter how many are made, it does not increase the riskiness of the theory.

Greenberg, Bradford and Curry (1993) test one prediction of social learning theory that cannot be reduced to this truism, and so is a much more interesting test of the theory. They hypothesised that the age of a perpetrator's victim would be positively correlated with their age at their own victimisation, because it is the behaviour with a similar aged child that had been learnt. They examined this by comparing the age at victimisation among a group of paedophiles (sexual preference for prepubertal children) and hebephiles (sexual preference for pubertal young adolescents). The mean age of victimisation for the paedophiles was 8.4 years, compared to 10.9 years.
for the hebephiles (t = -3.87, df = 74, p < 0.001). Although the support for the theory is based on only one directional prediction, the case for the theory is strengthened because it is difficult to see how competing theories could account for this difference. As discussed in chapter 4, the case for a theory is strengthened if it makes a prediction that is not made by other theories, and the prediction is subsequently borne out by the data.

7.7 Cognitive theories

Cognitive theories of sexual offending are typically models of maintenance, not development. The question of how cognitive mechanisms might put someone at risk of sexually offending in the first place is largely ignored in favour of describing how these mechanisms maintain the behaviour once it has occurred. However, some of these mechanisms might be equally applicable to understanding the origins of the problem; a suggestion supported by Keenan and Ward (2000; Ward, Keenan, & Hudson, 2000). They argue that key cognitive concepts, such as intimacy deficits, lack of empathy and cognitive distortions, all reflect an underlying deficit in theory of mind, which they define as an inability to understand and attribute mental states to others.

In the field of emotional disorders cognitive theory has burgeoned because of its ability to make testable and – every bit as important – risky predictions. There are early but promising signs that the same may be true of theories of sexual offending. Certainly, in principle, it should be possible to squeeze out of the theory some reasonably risky predictions. For example, rather than make a directional prediction about the relationship between two variables, the theory is capable of stating that a third variable, the cognitive mechanism, mediates this relationship. Even tighter predictions could be made by suggesting specific relationships between particular types of childhood experiences, particular cognitive mechanisms and particular types of outcome.
A recent attempt at testing a cognitive theory of sexual offending illustrates this strategy. Simons, Wurtele and Heil (2002) used Structural Equation Modelling to examine cognitive mediators of childhood experiences and offending behaviour. As predicted by cognitive theory, offenders reporting child sexual abuse displayed less empathy for children and reported more child victims, whereas those reporting physical abuse were more likely to display less empathy for adult women and reported more adult victims.

The application of cognitive theory to the development of sexual offending does offer the opportunity to make risky predictions, but construct validity is likely to be a problem. Take the findings of Simons, Wurtele and Heil (2002) as an example. Intriguing as they are, do they indicate that a person who has abused a child has a lack of victim empathy for children or is this being used by the offender to explain or excuse their offending? Which construct is being measured: lack of victim empathy or the feigning of cognitive distortions? It is likely to prove difficult – at least in retrospective research – to disentangle these constructs. A further problem is that some of the predictions are also made by Feminist Theories, and so this limits the riskiness of both theories.

### 7.8 Attachment theory

Marshall and co-workers' application of attachment theory to sexual perpetration (e.g. Marshall, 2001; Marshall, Hudson, & Hodgkinson, 1993) is a potentially fruitful development. A central concept of attachment theory is the notion of an internal working model (Bowlby, 1969; 1973), which proposes that the child's sense of self and of attachment figures derives from interaction with those figures. These interactions lead to expectations about subsequent relationships, and this in turn influence the nature of these relationships (Bowlby, 1969; 1973). Marshall argues that child molesters experience insecure attachments in childhood, and the model of
self and other that derives from these puts the individual at risk of failing major developmental tasks, particularly in the area of intimacy. It is suggested that the resulting emotional loneliness may increase the risk of a person turning to children to find emotional and sexual fulfillment.

The general literature on child maltreatment has also suggested a link between child maltreatment, insecure attachments and subsequent problems in social interaction (Cicchetti & Toth, 1995). A number of studies have found that maltreated children are more likely to be insecurely attached than non-maltreated children (e.g. Crittenden, 1988; Egeland & Sroufe, 1981). Maltreated children also display unusual patterns of behaviour that do not fit smoothly into the traditional attachment-classification scheme (Main & Solomon, 1990). For example, Carlson, Cicchetti, Barnett and Braunwald (1989) found that more than 80% of maltreated infants displayed this unusual pattern of behaviour, termed disorganised/disorientated, compared to less than 20% of a matched non-maltreated comparison group.

The insecure attachments styles of maltreated children may predispose them to peer rejection and isolation. In their review of the literature, Mueller and Silverman (1989) identify two characteristics of maltreated children in relation to their peers. First, maltreated children are more likely than non-maltreated children to display aggression in their interactions. Second, maltreated children, and in particularly neglected children, are more likely than non-maltreated children to avoid peer interactions.

What makes attachment theory interesting as a theory of sexual perpetration is that a number of predictions flow naturally from it that would require *ad hoc* adjustments from its competitors. The source of these novel predictions is the different sexual perpetration characteristics associated with the two distinct types of insecure attachment (the disorganised/disorientated category has not been incorporated into the theory proposed by Marshall). According to Marshall, Anxious-ambivalent
attachment styles result from inconsistent, neglectful or rejecting parenting. A person with this history would desire intimacy, but is afraid of rejection. Sex offenders with this attachment style would attempt to form affectionate sexual relationships with their victims. Avoidant styles result from punitive parenting, such as physical, sexual or emotional abuse. A person with this type of history has a negative view of themselves and consequently is afraid to form close ties with others. Sexual perpetrators with this attachment style seek out impersonal sexual contacts; for example, they might molest children only once, but have several victims. In sum, the model proposed by Marshall suggests that anxious-ambivalent sexual perpetrators will be more likely than avoidant sexual perpetrators to attempt to build up relationships with their victims, abuse fewer victims, but abuse each victim over a longer period of time. If these predicted relationships are empirically confirmed this would place ‘money in the bank’ for this theory, because it is unclear how competing theories could account for the findings. Although studies have examined the relationship between attachment and coercive sexual behaviour (e.g. Craissati, McClurg, & Browne, 2002; Smallbone & Dadds, 2000) these specific predictions have not been tested as yet.

The attachment theory of sexual perpetration is in keeping with the risk factors examined in this thesis. The risk factors are all categories of child maltreatment, and as discussed above, these experiences may affect the attachment style of the maltreated children and increase the risk of subsequent perpetration.

Although attachment theory is potentially of interest, it is important to note that it is not sufficient to explain the onset of sexual perpetration. By no means all children who exhibit insecure attachment styles are at risk of sexually perpetrating.

7.9 Biological theories

Psychophysiological theories of sexual abuse are remarkably underdeveloped, and the available research is inconsistent. Rada (1981) found that testosterone levels for
rapists, child molesters and an institutional comparison group were all within the normal range. Bradford and McLean (1984) also reported that testosterone levels of sexual offenders were within normal limits. However, Berlin (1983) reported elevated testosterone and luteinizing hormone (LH) levels in 50% of a sample of sexual perpetrators, who were primarily abusers or children. This finding may be due to a sample selection bias because the subjects had all been evaluated at a specialised hormone clinic. In contrast, Gumani and Dwyer (1986) found that an outpatient group of sexual perpetrators against children had lower levels of testosterone than did a matched group of men with psychogenic erectile dysfunction. (See appendix 6 for further discussion.)

7.10 Multifactor theories

The theories already reviewed tend to assume sexual perpetration can be explained by a single mechanism. As the previous chapters examining risk factors suggests, no single risk factor is likely to be both necessary and sufficient for sexual perpetration, and by extension no single mechanism. In response to this limitation, a number of authors have generated multicausal theories.

Multifactor theories are often in danger of being unfalsifiable 'theories of everything'. These theories suggest that different combinations of risk factors operating through a series of psychological mechanisms lead to sexual perpetration, but fail to state the necessary and sufficient combination of risk factors and psychological mechanisms for the outcome. This represents the very pinnacle of unrisky theorising because any outcome can be accounted for. However, this section will argue that a carefully constructed multifactorial theory can increase the riskiness of predictions.
7.10.1 Finkelhor’s Four-Factor Model

Araji and Finkelhor’s (1986) extensive review of the theories of child molestation identified four types of theory:

- Emotional congruence: theories explaining why an adult finds sexual interaction with a child emotionally gratifying
- Sexual arousal: theories explaining why an adult is sexually aroused by children
- Blockage: theories explaining why an adult is blocked from obtaining sexual and emotional satisfaction from adults
- Disinhibition: theories explaining why a person is not inhibited against having sex with a child.

The authors propose a multicausal model of sexual offending based on the four categories. The model predicts that sexual perpetration occurs when many or all of the four factors are present. The model is summarised in figure 7.1.

Figure 7.1 – Finkelhor’s Four-Factor Model

![Diagram showing the four-factor model of child molestation]

- Sexual arousal to children
- and/or
- Emotional congruence
- and/or
- Blockage
- and/or
- Disinhibition

Child molestation
7.10.2 Marshall and Barbaree’s Integrated theory


The authors argue that these experiences lead to a series of social deficits, including:
- Insecure attachment style, resulting in an inability to form adult attachments
- Low self esteem and poor self-image
- Antisocial behaviour, including lack of empathy and cognitive distortions.

The authors propose that this constellation of deficits leads to sexual offending (see figure 7.2). Adolescents that cannot gain access to appropriate sexual outlets, because to do so requires rudimentary social skills and competences, turn to the sexual victimisation of a vulnerable victim.

Figure 7.2 – Marshall and Barbaree’s Integrated Theory

- Parental neglect and absenteeism
- Lack of attachment
- Low self esteem
- Sexual offending
- Parental coercive behaviour
- Antisocial behaviour
7.10.3 Hall and Hirschman’s Quadripartite Theory

Hall and Hirschman (1992) present a quadripartite model of sexual offending against children. They suggest that this behaviour occurs as a result of one among four distinct motivational precursors: deviant sexual arousal, cognitions justifying sexual aggression, affective dyscontrol and personality problems (figure 7.3).

The model assumes that although each of the four factors can contribute to abusive behaviour, usually one factor is prominent for each offender. Hall and Hirschman (1992) also suggest that each of the precursors will be associated with a particular offender profile. For example, those offenders characterised by affective dysfunction are likely to behave in an impulsive and unplanned manner. In contrast, offenders characterised by cognitive distortions are likely to possess good self-regulatory and planning skills, and their offending behaviour will reflect this. Those classed as having deviant sexual arousal have strong sexual preferences for children and are likely to abuse a large number of children. Those with personality problems have difficulties functioning effectively in the social world, including difficulties establishing relationships with other adults, particularly intimate relationships.

Other researchers have also suggested that those offenders with personality problems represent a distinct subgroup, but, unlike Hall and Hirschman (1992), these emphasise psychopathic personality disorder (PPD). For example, Porter, Campbell, Woodworth and Birt (2001) suggest that PPD is associated with a unique pattern of sexual behaviour (e.g. diverse patterns of sexual offending), motivation (e.g. thrill seeking), and severity levels (e.g. excessive, gratuitous violence). It is possible, therefore, that there may be a need for a further subdivision of Hall and Hirschman’s (1992) final category.

This issue aside, a key strength of the model is the specification of different routes to sexual offending; in particular it states distinct combinations of motivational precursors and offence profiles. This is important because it helps ensure that a multicausal theory is capable of making risky predictions.
7.10.4 Ward and Siegert’s Pathways Model

The most recent model of sexual offending (Ward & Siegert, 2002) has sought to combine the strengths of the three earlier multicausal models discussed above. From Finkelhor’s model Ward and Siegert (2002) take the link between psychological vulnerabilities and the offence process; from Marshall and Barbaree’s they take the description of how early developmental adversities can result in vulnerability to offending. The emphasis on typology in Hall and Hirschman’s model also appears in the Ward and Siegert (2002) model. Given the reliance of Ward and Siegert’s (2002) model on the earlier multicausal theories, it will not be explained in any more detail.

7.10.5 Developmental psychopathology

Ryan (1997b) has suggested that Developmental Psychopathology (Sroufe & Rutter, 1984) may provide a suitable framework for a multicausal theory of sexual
offending, and other researchers have also recommended the adoption of this approach (e.g. Lee, Jackson, Pattison, & Ward, 2001). A Developmental Psychopathology model is not necessarily in opposition to earlier multicausal theories; in fact it is best seen as integrating and extending them. Ryan (1997b) provides only a brief overview of a Developmental Psychopathology model of sexual offending. In the next section the model is described in greater detail.

The overarching framework for Developmental Psychopathology is the organisational perspective (Cicchetti & Cohen, 1995), which derives from the work of Waddington (1957) via Bowlby (1973). This perspective sees development as a series of challenges requiring reorganisation within and among all domains of functioning, including the biological, social, emotional, cognitive, representational and linguistic (Cicchetti & Cohen, 1995). One of the main features of this perspective is that current organisation is seen as influencing future organisation, which derives from the assumption that organisation at one stage of development is integrated into a newly emerging system, rather than being supplanted by it (Cicchetti & Cohen, 1995). The successful resolution of a challenge increases the probability of successful adaptation to subsequent challenges; conversely, maladaption at one stage increases the likelihood of future maladaption (Sroufe & Rutter, 1984).

Another core tenet of Developmental Psychopathology is its transactional view of development (Sameroff, 1995). The transactional model derives from an interactional perspective in which both environment and biology are said to contribute to an outcome (Sameroff, 1995). The transactional model also emphasises the affect of the child on the environment (Rutter, Champion, Quinton, Maughan, & Pickles, 1995). From this perspective, development consists of a reciprocal influence between, first, the child’s biological organisation and behaviour, and second, the environment. This reciprocal influence at one stage of development influences development at future stages (Sameroff, 1995).
7.10.5.1 Complex causal models
Developmental Psychopathology rejects a simple causal model in favour of more complex ones (Cicchetti & Rogosch, 1996; Kazdin, Kraemer, Kesler, Kupfer, & Offord, 1997). Rather than assume that one causal variable is both necessary and sufficient for a particular outcome, the model assumes the operation of multiple causal variables. The model also accepts that the connection between each causal variable and an outcome may involve more than one mechanism and numerous stages or steps. The operation of a causal variable may differ depending on a number of boundary conditions, such as the environmental context, developmental stage, features of the causal variables (e.g. severity, duration), and the presence or absence of other causal variables. Despite this complexity there is recognition of a need for parsimony: there is an attempt to balance the explanatory power of the model with an economy of variables and pathways (Haynes, 1992).

7.10.5.2 Risk and protective factors
The concepts of risk and protective factors are central to the explanation of the relationship between early and subsequent organisation in development (Sroufe & Rutter, 1984). Risks decrease the likelihood of a successful resolution, and therefore place that individual at a disadvantage when facing subsequent challenges. Conversely protective factors foster successful resolution, are capable of reversing earlier maladaption, and equip the individual with the capacities to face future challenges.

7.10.5.3 Developmental psychopathology and sexual perpetration
From a Developmental Psychopathology perspective, a model of sexual perpetration is likely to involve multiple causal variables and for each of these it may be necessary to delineate multiple mechanisms and multiple stages. Additionally, the effect of a risk will vary depending on environmental context, developmental status and other boundary conditions, such as severity and duration of the risk. A model of the development of offending will require the identification of risk factors that
increase the likelihood of offending and protective factors that decrease the likelihood of this outcome.

7.10.5.4 Multifactor theories and risky theoretical predictions

Although it is clear that multicausal theories are needed, they are in danger of becoming unfalsifiable theories of everything. This is a problem for a Developmental Psychopathology model just as much as it is for earlier multicausal models. It could end up, like its predecessors, stating that an unspecified number of risk factors will lead to an unspecified combination of psychological mechanisms, and these will predispose to sexual perpetration. Here too there would be a failure to state the combinations of the risks and mechanisms that are necessary and sufficient for sexual perpetration.

However, multicausal theories need not be 'theories of everything', and in fact they offer one of the best opportunities for a theory of sexual perpetration that is capable of making risky predictions.

Multicasual theories can be used to delineate a number of different pathways to sexual perpetration, each of which may be associated with distinct characteristics in terms of risk experiences and psychological mechanisms. Once a theory identifies more than one developmental route to sexual perpetration, it becomes possible to derive a number of predictions about the different routes (e.g. likely risk factors, likely indicators of psychological mechanisms, number of victims, one-off or multiple abuse of each victim). Such a theory would be capable of making risky predictions, because there would be a low prior probability of these various associations clustering together if the theory had no verisimilitude. It is also possible that such predictions, flowing naturally from the multicausal theory, would require ad hoc adjustments from its competitors.

It is beyond the scope of this thesis to generate and test such a theory. The usual method is to first meet the association criterion and then to study the more complex
problems of non-spuriousness and temporal precedence. The criterion of the causal mechanism – where the riskiness of theoretical predictions becomes important – is typically studies once a field of research has reached some level of maturity. This thesis is designed to examine the non-spuriousness and temporal precedence criterion, and the catch-up design used in the research is well placed to do that (see chapter 9). As noted in chapter 9, the catch-up design is often unable to examine the causal mechanism criterion.

The research reported in this thesis does, however, draw upon a Developmental Psychopathology framework. The research assumes that there is more than one primary route to sexual perpetration and it endeavours to examine one of these: the route involving sexual victimisation. It is hypothesised that while this experience increases the risk of sexual offending, it is not sufficient in itself. In keeping with the Developmental Psychopathology model, and the other multicausal approaches, it is assumed that other risk factors will be involved. As suggested by Marshall and Barbaree (1990), it is proposed that other types of maltreatment act as additional risk factors that in combination lead to sexual offending.

7.11 Summary

An important causal criterion is to provide evidence of the psychological mechanism connecting a risk and sexual perpetration. Theories of sexual perpetration provide hypotheses about these mechanisms and these were reviewed in this chapter. While a number of plausible mechanism have been identified, there are few – perhaps no – convincing demonstrations of these mechanisms. The unrisky predictions made by many of these theories are the main culprits for this poor state of affairs.

An additional limitation of these theories is their focus on a single causal variable, which limits explanatory power. Multicausal models address this difficulty and these were also reviewed. Although these theories often do not make risky predictions,
this need not be the case, because a multicausal theory can delineate different pathways to sexual perpetration, and the relationship between these pathways and other variables can be used to generate risky predictions. The research reported in this thesis is not testing such a theory, but many of the risk factors it examines are likely to feature in such a theory. The results of this study could feed into the next stage of causal research examining psychological mechanisms.
Summary and conclusion of literature review

There is an urgent need to identify risk factors for sexually abusive behaviour, particularly among previously sexually victimised males. The search for these risk factors is the central aim of this thesis, and this chapter summarises the research that suggests there is a need to identify these risks.

There is a considerable lack of agreement in the literature about the prevalence rate of child sexual abuse. Three major reviews of prevalence studies have found substantial variation in the rates for both males (3 to 37%) and females (6 to 62%) (Dhaliwal, Gauzas, Antonowicz, & Ross, 1996; Goldman & Padayachi, 2000; Peters, Wyatt, & Finkelhor, 1986), but even the lowest figures indicate that a substantial proportion of the population is sexually victimised during childhood.

It is more difficult to interpret the research on the developmental effect of these experiences. Research has examined a vast range of outcomes of sexual abuse, including anxiety, depression, eating disorders, sexual adjustment and social adjustment. These studies have tended to conclude that sexual abuse is pathogenic, but the methodology of these studies is less than ideal. While there is evidence of an association between child sexual abuse and a range of psychopathological outcomes, the size of the association is typically small, and few studies have addressed the criteria of non-spuriousness, temporal precedence and the causal mechanism. A definitive conclusion cannot be made as yet, but the argument that sexual abuse is pathogenic must be taken seriously.

The prevalence of childhood sexual abuse and the hypothesised impact of these experiences suggest there is a need to generate effective prevention strategies. Secondary prevention aims to reduce the likelihood of re-offending among identified abusers. The evidence of the effectiveness of this prevention strategy hinges on the
question of whether randomisation is necessary to make causal interpretations. Although the conclusions of studies that have not randomised are optimistic, the small number of randomised studies has not found treatment to be effective. Even if the evidence for effectiveness was more compelling, this prevention strategy is inherently limited because the intervention occurs after at least one – and often many more – acts of sexual perpetration. Primary prevention strategies have attempted to educate children to identify dangerous situations. There is a shortage of methodologically robust research assessing this approach, but the available evidence does raise questions about effectiveness. Both primary and secondary prevention lack a firm evidence abuse; this has led to calls to develop alternative strategies that identify and treat potential perpetrators before the onset of abusive behaviour.

A first step in this process is to identify the risk factors for sexual perpetration. This is the central aim of this thesis, and is addressed by dividing it into questions about the causal status of these risk factors and the capacity of the risk factors to predict the onset of this behaviour. Four criteria must be met to establish that a putative risk is causal: there must be an association between the risk and the outcome, that association must be non-spurious, temporal precedence must be demonstrated and a causal mechanism must be identified.

One widely quoted hypothesis is that the experience of sexual victimisation may increase the risk of perpetration. Retrospective studies have found a robust association between victimisation and perpetration; there is also some evidence that the association is non-spurious. The strength of the association does appear to be stronger for males than for female victims, which is in keeping with the observation that more females than males are sexually victimised, but that the majority of sexual perpetrators are male. This suggests that it may be useful for prevention strategies to focus on sexually victimised males. It would be useful to have stronger evidence of the causal relationship, but the evidence is probably stronger for this risk factor than any other. In addition, social workers and mental health clinicians routinely see sexually victimised males, which makes prevention a genuine possibility.
Although the experience of sexual victimisation may increase the risk of perpetration by no means every male victim of sexual abuse develops abusive behaviour; other risk factors must be involved. Effective prevention for sexually victimised males requires the identification of those putative risks that are causally related to sexual perpetration in this group. A range of additional putative risk factors has been identified, but the causal status of these is open to question. For some risks, such as emotional maltreatment, there are too few studies to assess the evidence of a causal relationship. For other risks, there are more studies, but the evidence is still limited.

One potential limitation of studies examining maltreatment risk factors is the use of dichotomous measures of maltreatment. Authors such as Cicchetti and Barnett (1991) and Wolfe and McGee (1994) have argued that research should use more sophisticated measures. One of the additional aims of this study is to compare dichotomous and more than dichotomous measures of maltreatment (e.g. severity, duration).

A more general limitation of the research discussed in the literature review is the predominance of retrospective studies. The retrospective design can establish whether there is a significant association between the risk and the outcome, but is not well placed to address the other causal criteria.

The retrospective design is also limited in its ability to test the predictive capacity of a risk factor. Biases, such as perpetrators inventing or elaborating risk experiences, could artificially inflate the predictive capacity of a risk factor, as could other biases, such as switching the temporal order of events. A variable that lacks any predictive power may appear to hold promise if its assessment is based solely on a retrospective assessment.
The following chapter will argue that the catch-up design (Robins, 1966) is well placed to address these limitations. The potential and actual limitations of the design will also be outlined. One of the limitations of it is the inability to test theories about the psychological mechanisms linking the risk factors to sexual perpetration. As the previous chapter argued, there is a need to test a multicausal theory of sexual perpetration that is capable of generating risky predictions. Given the use of the catch-up design, it is outside the scope of this thesis to test such a theory.
9 Methodology: The catch-up longitudinal design

9.1 Introduction

The faith we can place in the methodology of a study determines the faith we can place in its findings. The selection of a research strategy for this thesis must consider the range of methodologies available and identify the one that finds an appropriate balance between methodological considerations and ethical constraints. The vast majority of studies have used a retrospective design to identify risk factors for sexually abusive behaviour. While this is a useful starting point, this method seriously limits the capacity of research to make causal inferences. It would also distort the predictive capacity of a risk index designed to predict which sexually victimised males are at most risk of perpetration. Firmer conclusions about both causal relationships and predictive capacity need longitudinal research, but a longitudinal study conducted in real time is ethically unacceptable. The study reported here uses a 'catch-up' longitudinal design (Robins, 1966), a seldom used strategy, but one that successfully balances methodological and ethical considerations. This chapter will make a case for the use of this design by identifying its methodological advantages and its circumnavigation of the ethical problems facing a real-time longitudinal study. The chapter will also discuss some of the potential and actual limitations of the design.

9.2 The retrospective design

Retrospective studies have dominated research in the area of risk factors for sexually abusive behaviour. Chapter 4 described two retrospective approaches. One compares the degree of exposure to a risk in a group of known sexual perpetrators and a non-perpetrating comparison group. A second approach identifies the proportion of a
non-incarcerated, non-clinical sample reporting sexual perpetration, the proportion reporting exposure to the risk factor, and the association between the reports. In both cases the information on the risk variable is coded from the subject’s retrospective self-report.

Chapter 4 has described in detail the limitations of this retrospective method for inferring a causal relationship between a given risk factor and sexual perpetration, but it is worth briefly summarising them here.

The outcome of sexual perpetration could bias the recall of a risk factor (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997). In retrospective recall it is possible that sexual perpetrators invent or elaborate experiences of putative risks in an attempt to explain or excuse their current functioning. Hindman (1988), for example, found that 67% of sex offenders (N = 40) referred before the introduction of polygraph testing reported being sexually victimised, compared to only 29% of those referred after its introduction (N = 129). This violates the non-spuriousness criterion. A problem with all retrospective studies in this area is that any observed association between a risk and sexual perpetration may be a result of the invention rather than the actual experience of that risk.

To establish a causal relationship between a risk factor and sexual perpetration it is important to assess whether the risk preceded the onset of sexual perpetration, but this temporal precedence criterion cannot be successfully addressed in a retrospective design (Menard, 1991). For this area of research the main problem is that the perpetrators may switch the temporal order in retrospective recall in an attempt to explain or excuse their sexually abusive behaviour.

The retrospective design also struggles to examine whether risk factors can be used to predict which sexually victimised males are at risk of sexual perpetration before the onset of this behaviour. The retrospective invention of risk factors may artificially inflate the predictive capacity of a putative risk factor. The switching of
temporal order may also mean that risk factors are coded as present before the onset of perpetration when in fact they were not. This could also artificially inflate the predictive capacity of the risk.

These problems prevent firm conclusions about the causal relationship between risk factors and sexual perpetration and the predictive capacity of those risk factors. The difficulties can be overcome by using a longitudinal design, and this will be described in the next section.

9.3 The longitudinal design

A longitudinal assessment follows a group exposed to the risk and one or more non-exposed groups to establish the proportion of each that perpetrates. In the traditional longitudinal design, the study takes place in real time. An important feature of the design is it establishes that the outcome is absent at the time of exposure to the risk (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997). This permits the researcher to draw causal inferences. Temporal precedence can be established because the temporal relationship between risk and outcome is clear. Information on risk is gathered prior to the onset of perpetration, which avoids the possibility of the outcome affecting the recall of a risk. The questions of prediction can also be successfully addressed in this design. The predictive capacity of a risk factor cannot be artificially inflated by retrospective switching of temporal order nor the invention of putative risks.

There can be ethical problems associated with the use of a longitudinal design (Maughan & Rutter, 1997), and these are magnified in this particular area of research. A longitudinal design promises answers tomorrow for questions that need answers today (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997). There is an urgent need to identify the risk factors for sexual perpetration, but this urgency cannot be addressed in a design that can take decades to complete.
There is a more fundamental ethical problem. The integrity of the design requires the researcher to follow the participants without making any form of intervention, even though the reason for conducting the study is that some people, those exposed to additional risk factors, are predicted to be at heightened risk. It is not ethically permissible to assume that some participants are at increased risk of perpetration, and yet fail to intervene to prevent them abusing others.

9.4 The catch-up longitudinal design

Fortunately the catch-up longitudinal design (Robins, 1966) provides an inventive solution to this ethical problem. This technique retains much of the methodological strength of the real-time longitudinal design, but avoids the ethical problems associated with it.

Because the catch-up design is seldom used, it may be helpful to describe the differences between it, the real-time longitudinal design and the retrospective design. To do this, it is helpful to outline the four stages in the origin of research data used by Bergman, Gunnar and Magnusson (1991) in their discussion of longitudinal research:

- The reference point in time: the particular occasion or age level in a person’s life to which the data refer
- The coding point in time: the point in time when the observation is coded as data
- The collection point in time: the point in time when the researcher collects the data
- The usage point in time: the point in time when the researcher uses the data for a calculation.

The main difference between the designs is the temporal order of the coding point and the collection point of the predictor variable (see table 9.1).
Table 9.1 – Differences between the real-time longitudinal design, the catch-up longitudinal design and the retrospective design

<table>
<thead>
<tr>
<th></th>
<th>Real-time longitudinal</th>
<th>Catch-up Longitudinal</th>
<th>Retrospective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding point of the predictor</td>
<td>Before outcome</td>
<td>Before outcome</td>
<td>After outcome</td>
</tr>
<tr>
<td>Collection point of the predictor</td>
<td>Before outcome</td>
<td>After outcome</td>
<td>After outcome</td>
</tr>
</tbody>
</table>

In a ‘catch-up’ design the information on the predictor variable is collected from information coded in the past for reasons other than the research question. While the collection point occurs after the outcome, the coding point happens before it. Information on the outcome is gathered as in any other design.

The research study reported in this thesis uses a catch-up longitudinal design. In this study, information on the predictor variables was gathered from clinical and social service files. Information on sexual perpetration was collected from social service material and criminal records. To protect the integrity of the longitudinal design, the assessment of the risk factors is based entirely on data gathered before the onset of sexual perpetration. Chapter 11 provides full details of the method used in this research study.

9.4.1 Advantages of the catch-up longitudinal design

The use of the catch-up design in which information is coded prior to the outcome helps to address the causal and prediction questions that are at the centre of this thesis. As with the real-time longitudinal design, temporal order can be established and the outcome cannot affect the coding of the risk factor, which helps us draw causal inferences. As in the real-time design, the catch-up approach allows a realistic
test of the capacity of a risk index to predict which sexually victimised males are at risk of perpetration before the onset of this behaviour. This is because retrospective reporting biases such as the switching of temporal order and the invention of putative risk factors cannot operate, and therefore cannot artificially inflate the predictive capacity of the index.

The design maintains these advantages of the real-time longitudinal design, but successfully avoids the ethical problems associated with it. Risk factors can be coded on the basis of information gathered before perpetration, but, because the outcome is already known, the design avoids the ethical necessity for intervention. Also the information is coded but not collected in real time, and so the design is able to provide answers to an urgent question in years rather than decades.

**9.4.2 Limitations of the catch-up longitudinal design**

In both the real-time longitudinal design and the retrospective design, the researcher can control the coding of the predictor variable. The researcher using a catch-up design does not have this luxury:

> "The nature of the earlier-gathered information restricts the research questions that the investigator can ask... The research worker of the current moment is at once the victim and the beneficiary of the original sources." (Radke-Yarrow, Campbell, & Burton, 1970)

The majority of the potential and actual limitations of the catch-up design result from this lack of control over the coding of the predictor variable. The rest of this chapter examines the actual and potential limitations of this design.

**9.4.2.1 Random error**

The researcher is unable to ensure that information is gathered in the same way on each case. In the research study reported in this thesis, there may be variability
between social service departments and between social workers in the type and quality of information gathered on subjects. While this is a genuine problem, it is limited in that the error introduced is likely to be random, not systematic. The information on risk is collected using data coded prior to perpetration, and so it is difficult to see how this problem could lead to a systematic bias in risk scores between perpetrators and non-perpetrators. The between-subject variability may mask a causal relationship, but it is unlikely to lead to a spurious relationship between a risk and an outcome.

9.4.2.2 Measurement of risk factors
The catch-up longitudinal design uses information collected for reasons other than the research question. The contemporaneous records might contain such limited information about the risk factors that this will invalidate this approach. Certainly some of the putative additional risk factors discussed in chapter 6 could not be measured in this study, because the case files lacked sufficient information on them. The implications of this are discussed in chapter 14.

The risk factors examined in this thesis are different forms of maltreatment. Because the only other feasible alternative to the catch-up design is a retrospective study, it would be useful to compare a ‘gold-standard’ measure of maltreatment with those obtained from the two research designs. Unfortunately in the absence of a ‘gold-standard’ this is difficult.

It is possible to use a technique that makes some assessment of the validity of the retrospective and catch-up designs. This technique assumes a positive report by either a contemporaneous account or a retrospective account is a true positive. It is then possible to compare the proportion of maltreatment missed by contemporaneous records (i.e. maltreatment rated as present according to retrospective report, but absent from contemporaneous report) with that missed by retrospective reports (i.e. maltreatment rated as present according to contemporaneous reports present, but absent from retrospective report).
Wolfe, McGee, Yuen, Wilson and Carnochan (1995) provide a comprehensive assessment of the validity of the two designs using this technique. This study compared the retrospective reports of maltreatment made by adolescents (N = 160, aged 11 to 17) and the ratings made by researchers based on social work case files held on the same subjects. Table 9.2 indicates the proportion of maltreatment identified by one source but not the other. 'Contemporaneous present / retrospective absent' refers to the proportion of cases in which the contemporaneous records indicated the presence of maltreatment when the retrospective report failed to report its presence. 'Retrospective present / contemporaneous absent' refers to the proportion of cases in which the retrospective report indicated the presence of maltreatment when the contemporaneous record failed to indicate its presence.

Table 9.2 – Comparison of contemporaneous and retrospective recall of dichotomous measures of maltreatment (Wolfe, McGee, Yuen, Wilson, & Carnochan, 1995)

<table>
<thead>
<tr>
<th>Type of maltreatment</th>
<th>Contemporaneous present / Retrospective absent (%)</th>
<th>Retrospective present / Contemporaneous absent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical abuse</td>
<td>6.3</td>
<td>21.3</td>
</tr>
<tr>
<td>Family violence</td>
<td>16.3</td>
<td>6.3</td>
</tr>
<tr>
<td>Sexual abuse</td>
<td>7.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Emotional abuse</td>
<td>14.4</td>
<td>4.4</td>
</tr>
<tr>
<td>Neglect</td>
<td>30.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>

With the exception of physical abuse, the contemporaneous records identified substantially more maltreatment than did retrospective recall. Other studies have not found that contemporaneous records underreport physical abuse. Femina, Yaeger and Lewis (1990) found inconsistencies in reports of physical abuse between contemporaneous records and retrospective interview in 26 out of 69 cases. Of this 26, 18 denied or minimised being abused, when files reported that they had been. The remaining 8 reported in interview that they had been physically abused, when the records contained no evidence of this.
Interpretation of the data is difficult in the absence of a gold standard measure. However, if a positive report by either source is likely to be a true positive, we can make some estimate about the relative validity of the two approaches. Despite the maltreatment information being coded for reasons other than the research question, there is some evidence it provides a more valid measure of maltreatment than retrospective reports.

9.4.2.3 Generalisability
A catch-up longitudinal design can use cases only on which detailed contemporaneous information exists, and this may limit the generalisability of the findings. In this study a subject could be included in the sample only if a social service file existed on that person. The findings of the research study may not be generalisable to subjects who have not had substantial involvement with social services. However, this research project aims ultimately to provide guidance on risk factors for sexually abusive behaviour in sexually victimised males, and these prevention strategies could only focus on male victims who have come to the attention of professionals. While the catch-up design may limit generalisability to all male victims of sexual abuse, it does not necessarily inhibit generalisation to the target population.

9.4.2.4 Researcher bias
In a real-time longitudinal design information on risk factors is coded and collected prior to the outcome of interest. In a catch-up longitudinal design the information is coded prior to the outcome, but the researcher collects the data after the outcome. In a catch-up design, unlike in a real-time longitudinal design, it is possible that a researcher's knowledge of the outcome could affect the scoring of the risk factors. One solution would be to employ researchers blind to the study hypotheses, but this was not financially feasible in the current project. In the study described in this thesis a range of novel techniques were used to minimise the possibility of researcher bias; these are described in detail in chapter 11.
9.4.2.5 Unidentified perpetration

The most important feature of a longitudinal design is that it establishes the absence of an outcome at the time of measurement of the risk factor (Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997). If this can be established, it is clear that any information gathered on risk is free of the possible biases associated with the presence of an outcome affecting the coding of that risk. This substantially strengthens the case for a relationship between a risk and an outcome to be considered as non-spurious. The temporal precedence criterion also requires establishing the absence of an outcome when the risk factor is assessed. If we are to conclude that the risk preceded the outcome, it is necessary to establish the absence of the outcome when measuring the presence of a risk.

In this thesis it is necessary to rely on contemporaneous case material to assess whether sexual perpetration had occurred by a particular age. It could be argued that it is difficult to assume that perpetration was not present, because the researcher was not involved in the coding of this data, and was therefore not able to intensively screen for perpetration. This may mean that unidentified perpetration may have occurred, and it may influence a subject’s report of risk factors to social services. The presence of unidentified perpetration would also make it difficult to establish the temporal relationship between a risk factor and the outcome.

It should be noted that even if a real-time longitudinal design were ethically permissible, it would not be possible to conclusively screen for sexually abusive behaviour. Sexual perpetration is not readily observable and it is unlikely to be admitted in a self-report measure. Even a real-time longitudinal design cannot rule out this possible source of non-spuriousness, but an attempt to screen for this behaviour would increase the probability, compared to a retrospective design, that a relationship was non-spurious. The same point applies to the temporal relationship between the risk factor and the outcome. The possibility of unidentified perpetration occurring before exposure to a risk cannot be ruled out, but an attempt to screen for
this behaviour would increase the probability that the temporal relationship was consistent with the causal status of the risk factor.

These arguments can also be applied to the catch-up study reported in this thesis. The subjects in this study had close contact with social services, often on a weekly basis over a number of years. Any inappropriate sexual behaviour committed by subjects and either observed by or reported to social workers would be recorded in case files because of mandated reporting procedures. This provides some form of screening procedure, but it does not rule out the possibility of unidentified perpetration.

In a retrospective study – the only feasible alternative to the catch-up design – the information on risk is gathered, by definition, after the point of perpetration. The retrospective design cannot address the problem of the presence of the outcome biasing the recall of the risk factors, nor can it address the issue of temporal precedence. In contrast, the catch-up design can make some attempt at meeting these criteria.

9.4.2.6 Causal mechanism
Chapter 7 suggested that a complex causal model would be needed to adequately explain the onset of sexual perpetration. It is useful to assess the extent to which the catch-up design can test and refine such a model. Haynes (1992) distinguishes between a number of types of causal variable, including original causes, causal chains and trigger variables. The type of variable that can be measured in a catch-up design is determined by the quality of the contemporaneous material, but in this particular research study, despite the detail contained in social service material, it was only possible to examine original causes. These are defined by Haynes (1992) as the first clinically or theoretically important variable in a causal chain linking the cause to the outcome. While it is essential that these original causes are identified, this far from captures the full complexity of the development of offending. Each original cause may be connected to the outcome by a number of mechanisms, and
each of these may involve several stages of a causal chain. As Cook and Campbell (1979) have noted (p. 230), one of the main limitations of archival data is the lack of information on 'process', which makes these sources of data problematic for researchers interested in psychological constructs. Importantly this limits the ability of the study to test the fourth causal criterion: the need to specify, and preferably test, the mechanisms linking a causal variable and an outcome.

9.5 Summary

The methodological limitations of retrospective research and the ethical objections to a real-time longitudinal study are overcome by the catch-up longitudinal design. This design avoids the limitations of the retrospective approach and retains many of the methodological strengths of a standard longitudinal design, while avoiding the ethical objections to a study conducted in real time.

This chapter has assessed each of the potential limitations of the catch-up design and has argued that the potential problems do not compromise the methodological adequacy of the design. The catch-up design may increase the random error in measurement, but this would attenuate any relationship between a risk and an outcome, rather than increase the likelihood of a spurious relationship. The available evidence, in fact, suggests that detailed contemporaneous records gathered for reasons other than the research question may provide a more accurate picture of maltreatment experiences than retrospective self-reports. The subjects used in a catch-up design are restricted to those on whom contemporaneous case material exists, and, while this may limit generalisability, it may not limit it to the target population of the research.

The absence of sexual perpetration cannot be conclusively demonstrated using a catch-up design, but this would also prove difficult in a study conducted in real time. Although a conclusive demonstration was not possible, the design does decrease the
probability that an observed association between a risk and an outcome can be accounted for by the presence of the outcome affecting the measurement of the risk. The same point applies to the temporal precedence criterion, which also relies on a demonstration of the absence of an outcome.

A genuine limitation of the catch-up design is that it cannot assess and refine theories about the mechanisms connecting a risk and an outcome. This limitation will be discussed in greater detail in chapter 14.

One of the aims of this thesis is to examine the causal status of putative risk factors for sexual perpetration. The catch-up longitudinal design presents a methodological advance for research seeking to do this. Another aim is to establish whether it is possible to identify those sexually victimised males who are at risk of sexually perpetrating before the onset of this behaviour. This too requires a catch-up longitudinal design, because a realistic test of predictive capacity requires the researcher to use data gathered before the outcome occurs. The next chapter will explore the methodological issues related to this predictive question.
10 Methodology:
Actuarial prediction of sexual perpetration

10.1 Introduction

The study described in this thesis capitalises on risk factor information gathered before the onset of sexual perpetration. The use of this material allows an exploration of whether it is possible to predict which sexually victimised males are at risk of sexually perpetrating before the onset of this behaviour.

A range of methodological, financial and ethical considerations determines the usefulness of a risk index that predicts an outcome such as sexual perpetration. This chapter will describe these considerations and use them to derive an analysis strategy for the indexes developed as part of this thesis.

10.2 Actuarial prediction of sexual recidivism

There are currently no actuarial instruments designed to predict the onset of sexual perpetration, but this is not surprising. The majority of research has focussed on the prevention of sexual recidivism; only recently has attention turned to the prevention of abusive behaviour before its onset (Ryan, 1997a). However, instruments for predicting which sexual perpetrators are at risk of re-offending do exist, and a number of studies have examined their predictive validity (e.g. Grubin, 1998; Hanson, 1997; Hanson & Thornton, 1999; Quinsey, Harris, Rice, & Cormier, 1998). Although these instruments are designed to predict who will re-abuse rather than who will abuse, it is worth examining the accuracy of these instruments. One reason for doing so is that they can provide an idea of the level of predictive accuracy that
can be achieved for a similar class of behaviour. The behaviour is not only similar in that it is a form of sexual perpetration, but it is also likely to be temporally distant from the point at which the prediction is made and determined by an INUS condition in which a large number of causal factors may be effectively random. Also the research reported in this thesis and the research on predicting sexual recidivism both face the problem of predicting an outcome that may be difficult to measure with any accuracy given its secretive nature.

Table 10.1 summarises the predictive validity of four of the most commonly used risk-assessment instruments for sexual recidivism. In the studies reported in the table, the researchers assessed predictive accuracy by using the risk instrument to calculate a risk score for each sex offender at time 1 and then establishing which sex offenders have re-offended at time 2. The measure of predictive accuracy used in these studies is the area under the ROC-curve (McFall & Treat, 1999; Rice & Harris, 1995), which is described in detail in chapter 11. In brief a ROC-curve is obtained by plotting sensitivity (the hit rate) against 1 minus specificity (the false alarm rate) for each possible cutting point on an instrument. The area under the ROC-curve represents overall discriminatory capacity. The area value can range from 0.5 (prediction no better than chance) to 1.0 (perfect prediction). In general, the area under the ROC-curve can be interpreted as the probability that a randomly selected person with the outcome (e.g. a sexual recidivist) will have a higher score on the instrument than a randomly selected person without the outcome (e.g. a non-recidivist) (Rice & Harris, 1995). As indicated in table 10.1 the various instruments perform at a level significantly better than chance (lower boundary of 95% CI greater than 0.5), with areas under the ROC-curves ranging from 0.62 to 0.71. It is possible, therefore, to predict which sex offenders are at risk of re-offending at a level significantly better than chance. However, given the low base rate of recidivism in these studies, the instruments may be less accurate than predicting that no one will reoffend. The area under the curve also indicates that there will be some degree of misclassification no matter which cutting point is used.
Table 10.1 – Predictive validity of sexual recidivism risk assessment instruments

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Sample</th>
<th>Area under ROC-curve (95% CI)</th>
<th>Pearson’s point biserial (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid Risk Assessment for Sex Offence Recidivism (RRASOR; Hanson, 1997)</td>
<td>Male sex offenders (N = 1208) (Hanson &amp; Thornton, 1999)</td>
<td>0.68 (0.65 to 0.72)</td>
<td>0.28 (0.23 to 0.33)</td>
</tr>
<tr>
<td>Sex Offender Risk Appraisal Guide (SORAG; Quinsey et al., 1998)</td>
<td>Male sex offenders (N=142) (Quinsey et al. (1998))</td>
<td>0.62</td>
<td>-</td>
</tr>
<tr>
<td>Static-99 (Hanson &amp; Thornton, 1999)</td>
<td>Male sex offenders (N = 1208) (Hanson &amp; Thornton, 1999)</td>
<td>0.71 (0.68 to 0.74)</td>
<td>0.33 (0.28 to 0.38)</td>
</tr>
<tr>
<td>Structured Anchored Clinical Judgement (SACJ; Grubin, 1998)</td>
<td>Male sex offenders (N = 1208) (Hanson &amp; Thornton, 1999)</td>
<td>0.67 (0.63 to 0.71)</td>
<td>0.23 (0.18 to 0.28)</td>
</tr>
</tbody>
</table>

10.3 Factors determining clinical applicability

Whether a particular actuarial instrument should be used to predict an outcome in a particular setting depends on a range of questions that will be discussed in turn.

10.3.1 Base rate of sexual perpetration

The base rate of the outcome in the particular population to which the risk instrument is to be applied can influence the utility of the instrument (Meehl & Rosen, 1955; Rice & Harris, 1995). For the risk instrument discussed in this thesis the base rate is the proportion of sexually victimised males that subsequently perpetrate. As an example of the influence of the base rate, assume that a risk instrument can successfully identify 70% of those who subsequently perpetrate (sensitivity) and 70% of those who do not go on to perpetrate (specificity). If the base rate were 10%, this would lead to the conclusion that the test had no predictive value, because 90% accuracy could have been achieved by predicting that no one
would perpetrate. (This assumes that equal weight is given to the cost of a false alarm (false positive) and a miss (false negative), a point that will be addressed shortly.) If the base rate were 50%, then the risk instrument providing 70% accuracy would be a valuable improvement on the prediction that no one would perpetrate or that a randomly selected one half would perpetrate (50% accuracy in both cases).

It is important to have an idea, even if it is an approximate one, of the base rate of sexual perpetration among sexually victimised males. The proportion of perpetrators discussed in the current study cannot be used as an estimate, because that rate is artificially inflated. The sample consisted of sexually victimised males who had been referred to Great Ormond Street Hospital as either victims of sexual abuse or as perpetrators. To identify the base rate it would be necessary to follow the group referred as victims (who had not perpetrated at the time of referral), and identify the proportion that subsequently perpetrated. Skuse et al. (in preparation), using a sample that has some overlap with the one discussed in this thesis, did just that. They identified 224 sexually victimised males who had not sexually perpetrated at the point of referral to Great Ormond Street Hospital. According to hospital, social service or criminal records, 11% of these subsequently perpetrated.

A number of other studies also provide some information on the degree of continuity between victimisation and perpetration. The catch-up longitudinal study by Widom and Ames (1994) discussed in chapters 5 and 6 found that 1 out 30 (3.3%) sexually victimised males subsequently obtained a conviction for a sexual offence (calculation made by present author). However, this is likely to be an underestimate because the outcome data is based solely on conviction data. Watkins and Bentovim (1992) describe three studies and these found rates of 13% (4/31) (Friedrich, Beilke, & Urquiza, 1988), 50% (3/6) (Sansonnett-Hayden, Haley, Marriage, & Fine, 1987) and 66% (2/3) (Chasnoff, Burns, Schnoll, Burns, Chisum, & Kylespore, 1986). Considerable caution is needed when interpreting the results of the two studies finding the highest base rate because both had a very small sample sizes. On the
basis of the three studies, Watkins and Bentovim calculate a pooled base rate of 22% (9/40).

The base rates reported for sexual perpetration among sexually victimised males vary considerably (range 3% to 66%). However, this may be due to the very high base rates reported by two studies with small sample sizes. If these two studies are excluded, the base rates ranges from 3% to 13%, with the largest study (Skuse et al., in preparation) reporting a figure of 11%.

The variations in base rates reported in these studies suggest that the efficacy of the risk instrument should be examined for a number of base rates. Also it is likely that different assessment and treatment centres will have somewhat different base rates. For these reasons the analysis section will assess the risk instrument using three different base rates (3%, 10% and 17%).

10.3.2 Utility decisions

Utility decisions refer to the relative weight placed on the four possible outcomes of classifying people as 'at risk' or 'not at risk' on the basis of a cutting point on a risk instrument. These four outcomes are summarised in table 10.2. The costs and benefits referred to in this table may be those for a person or those for society.

Table 10.2 – Utility of four possible classification outcomes

<table>
<thead>
<tr>
<th>Utility</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of a false alarm</td>
<td>The cost of a male victim being labelled 'at risk for perpetration' when he has not perpetrated and will not perpetrate in the future.</td>
</tr>
<tr>
<td>(false positive)</td>
<td></td>
</tr>
<tr>
<td>Cost of a miss</td>
<td>The cost of a male victim being labelled 'not at risk for perpetration' when he subsequently perpetrates.</td>
</tr>
<tr>
<td>(false negative)</td>
<td></td>
</tr>
<tr>
<td>Benefit of a hit</td>
<td>The benefit of a male victim who subsequently perpetrates being correctly identified as 'at risk for perpetration'.</td>
</tr>
<tr>
<td>(true positive)</td>
<td></td>
</tr>
<tr>
<td>Benefit of a correct rejection</td>
<td>The benefit of a male victim who does not perpetrate being correctly identified as 'not at risk for perpetration'.</td>
</tr>
<tr>
<td>(true negative)</td>
<td></td>
</tr>
</tbody>
</table>
Each of these costs and benefits will now be considered in turn.

10.3.2.1 Cost of a false alarm (false positive)

A false alarm (false positive) occurs when a person who is not and will not sexually perpetrate is labelled as at risk of perpetration. The main cost is to the person being falsely labelled in this way, but there may be additional costs. There are, for example, the numerous negative consequences for the person of knowing that they have been labelled as a potential perpetrator. As Watkins and Bentovim (1992) discuss, this label may lead to a self-fulfilling prophecy, increasing the probability of subsequent abusive behaviour. These consequences would be exacerbated for that person if others were aware of the label. It is clear that the cost of a false alarm (false positive) will be dependent on with whom the information is shared. The way in which the label is applied may also be important. There is a difference between begin labelled a ‘child at very high risk of sexual perpetration’ and a ‘child requiring additional treatment’.

If the cost of a false alarm was considered to be high then one solution would be to label all sexually victimised males as not at risk of perpetration. Here the false alarm rate would be reduced to zero and the base rate of perpetration would be the misclassification rate. For example, if 10% of sexually victimised males subsequently perpetrated, then the misclassification rate would also be 10%. In this situation, a risk instrument would need to increase the proportion correctly classified (by correctly labelling subsequent perpetrators) without introducing any false alarms (false positives).

However, labelling every sexually victimised male as not at risk of subsequent sexually abusive behaviour assumes that there is no cost of a miss (false negative), and there may well be.
10.3.2.2 Cost of a miss (false negative)

A miss (false negative) occurs if a sexually victimised male is labelled as not at risk of sexual perpetration when in fact he subsequently perpetrates. Here the costs apply to the child or children sexually victimised by this perpetrator. One position would be to argue that the cost of a miss (false negative) is so substantial that all sexually victimised males should be assumed to be at risk and all should receive treatment. Although this would lead to a very high false alarm rate (100% - base rate; e.g. 90% if base rate is 10%), it would ensure that all subsequent perpetrators were identified. Under these conditions a risk instrument would need to increase the number of correct classifications without introducing any misses (false negatives).

10.3.2.3 Benefits of correct rejections (true negatives) and hits (true positives)

It is also worth examining the benefits of correct classification. The benefit of a subsequent perpetrator being correctly identified is that treatment can be given for this problem before it begins. This raises the possibility of a treatment preventing the behaviour from occurring.

The benefit of a correct rejection (true negative) is that unnecessary treatment is not given to a person, and the treatment given can then focus on the other problems that may be experienced by a sexually victimised male.

10.3.2.4 Balancing utilities

There is an inevitable trade-off between these utilities (McFall & Treat, 1999). For example, if we attach a substantial cost to false alarms (false positives) and we choose a cut-off point to reflect this concern, we will inevitably increase the number of misses (false negatives) (Meehl & Rosen, 1955). The question of how to balance the utilities can be extremely difficult to answer (Swets, 1992). A particular difficulty is the tension between the costs and benefits to the person of being labelled as 'at risk' and the costs and benefits to society. Rather than present a pseudo-definitive answer to this difficult balancing act, the approach taken in this thesis will be to assess the risk instrument using two decision analyses in which
different weights are assumed. A decision analysis identifies the optimal cutting point for a risk index when the utilities are given different weightings. The first analysis assumes that it is twice as important to obtain a true positive than the other utilities; the second analysis assumes that true negatives are twice as important as the other utilities. The next chapter describes how a decision analysis is conducted.

### 10.3.3 Treatment factors

A number of treatment factors also affect the usefulness of a risk index, including the efficacy of treatment, the proportion of the target population given treatment and dose-response relationships in treatment.

#### 10.3.3.1 Effectiveness of treatment

The effectiveness of treatment has a marked influence on the usefulness of a risk index. To illustrate this, assume that a treatment was entirely ineffective. If the treatment did not work, nothing would be gained by the identification of those at highest risk. The proportion that subsequently perpetrated would be exactly the same whether a risk index was used to select the proportion for treatment or that proportion was picked at random. As the treatment becomes more effective then any advantage for the identification of those at highest risk will be increased along with it.

As yet there are no treatment programs designed to decrease the likelihood of a non-perpetrating sexually victimised male from sexually abusing others. Therefore, to develop a risk index before the treatment may seem questionable. We have no idea whether an effective treatment could be developed, nor the extent of its efficacy. For exploratory purposes the discriminatory capacity of the risk instrument will be assessed under different assumptions about treatment efficacy.

This degree of efficacy can be quantified in terms of the proportion reduction in the base rate. For example, if 10% of sexually victimised males perpetrate without
treatment, but this is reduced to 5% with a particular treatment then that treatment has an efficacy of 50%. In the absence of information on treatment efficacy, the risk index will be examined assuming a wide range of treatment efficacies (25%, 50% and 75%).

10.3.3.2 The proportion of the target population selected for treatment
If a high value is placed on the reduction of misses (false negatives) and there are sufficient finances to treat all sexually victimised males, then the risk instrument is redundant. Treatment would be given to all males identified as sexually victimised, and so there would be no reason to identify those at highest risk. However, financial restrictions may mean that not all of an at-risk population can be treated. Here there may be a role for a risk index to help select for treatment those at highest risk.

10.3.3.3 Dose-response relationships in treatment
Dose-response relationships in treatment efficacy may complicate decisions about the delivery of treatment. For example, the extent to which treatment is effective may be related to the duration and intensity of the treatment program. A decision would need to be made about whether clinicians should treat fewer people with a more effective but costly treatment or treat more people with a less effective but cheaper treatment. Again, in these circumstances a risk instrument may be of use.

10.3.3.4 Assessing treatment factors
A further set of analyses, described in the next chapter, will be carried out to examine the influence of these treatment factors on the performance of the risk index. Given that a treatment aimed at reducing the probability of perpetration has not been developed, these analyses should not be seen as an evaluation of the risk index, but rather an example of how the risk index could be evaluated.

In circumstances in which only a proportion of a clinical population could be treated because of financial restrictions, the analyses can identify the extent to which the use of the risk index reduces perpetration compared to random selection of that
proportion. The analyses also allow an examination of how the improvement varies as a function of the proportion selected for treatment, the base rate of perpetration and the efficacy of treatment. We can also compare across different treatment efficacies to explore decision-making when there is a dose-response relationship in treatment. For example, the analyses can be used to answer questions such as whether 30% of the population should be treated using a more effective (but more costly) treatment or whether 60% of the population should be treated using a less effective (but cheaper) treatment.

10.4 Summary

Whether the risk instrument developed as part of this thesis would be of use in a clinical or social work setting is dependent on a wide range of factors. These include assumptions about the base rate of sexual perpetration among sexually victimised males, the balance given to the costs and benefits of correct and incorrect classification, the effectiveness of treatment and financial considerations. Given that these are likely to vary across situations, the predictive capacity of the risk instrument will be examined under a range of conditions. The performance of the index will be examined under three different base rates. A decision analysis will also be used to identify the optimal cutting point when true positives are given twice the weight of other utilities; a second decision analysis will assume that true negatives are twice as important as the other utilities. A further set of analyses will examine the influence of treatment factors on the performance of the index. These factors include the extent to which treatment reduces the probability of perpetration, the proportion of the target population who are offered treatment and dose-response relationships in treatment.

These analyses, however, are exploratory. The balance of utilities in the two decision analyses cannot be considered definitive, because it is far from clear how
the utilities should balanced. Also the technique used to assess treatment factors make assumptions about treatment efficacy, but as yet no treatment exists.
11 Method

11.1 Introduction

The research reported in this thesis uses a catch-up longitudinal design (Robins, 1966). This design was selected because of its ability to address the two central questions of this thesis: cause and prediction.

11.2 Aim and hypotheses

The aim of the research reported in this thesis is to identify risk factors for sexual perpetration in previously sexually victimised males. More specifically, the study aimed to examine the causal status of putative risk factors. It also aimed to establish whether a risk index, consisting of a set of *a priori* risk factors, could successfully predict which sexually victimised males were at risk of perpetration before the onset of this behaviour.

11.2.1 Hypotheses

Hypothesis 1: Key maltreatment risk factors (sexual victimisation by a female, experiencing physical abuse, witnessing intrafamilial violence, neglect - failure to provide, neglect - lack of supervision, rejection by carers, number of separations from primary carers) will discriminate at the bivariate level between those sexually victimised males who subsequently perpetrate from those who do not do so.

Hypothesis 2: A risk index summating key maltreatment risk factors (sexual victimisation by a female, experiencing physical abuse, witnessing intrafamilial violence, neglect - failure to provide, neglect - lack of supervision, rejection by
carers, number of separations from primary carers) will successfully discriminate between those sexually victimised males who go on to perpetrate sexually from those who do not do so.

11.3 Design

A description of the catch-up longitudinal design and the rationale for using it are given in chapter 9. In brief, a ‘catch-up’ longitudinal design capitalises on information gathered in the past for reasons other than the research question and it uses this information as the data source for the predictor variables (risk factors). Information on the outcome (sexual perpetration) is gathered as in any other design. In contrast to a longitudinal study conducted in real time, in a catch-up design the researcher collects information on the predictor variable after the outcome has occurred.

The catch-up longitudinal method was used to examine hypothesised risk factors for sexually abusive behaviour in sexually victimised males. Information on hypothesised risk factors was collected from clinical and social services case files held on sexually victimised males referred to Great Ormond Street Hospital between 1980 and 1992. Information on sexual perpetration was collected from social service material and criminal records. To protect the integrity of the longitudinal design, and to ensure that the questions of cause and prediction could be successfully addressed, the assessment of the risk factors is based entirely on data written before the onset of sexual perpetration.

The reliance on information gathered before the onset of perpetration is a crucial feature of the design (see figure 11.1).
For a subject who developed sexually abusive behaviour, the assessment of the experience of key risk factors was based entirely on information in social service or clinical files that was *written before the onset of perpetration*. For example, if a subject began perpetrating in May 1987, only information written before this date would be used in the scoring of a risk factor. Information written after this date would be excluded from the assessment, even if it referred to events that clearly occurred before the onset of perpetration.

This strategy protects the methodological integrity of the catch-up longitudinal design, and is necessary if the thesis is to successfully examine the questions of cause and prediction. The non-spuriousness criterion of causality is always a challenge for non-experimental studies, but one important threat for this area of research is the possibility that perpetrators invent or elaborate risk factors in an attempt to explain or excuse their behaviour. If information on risk factors is
documented before the onset of perpetration then this threat is eliminated. It also protects against subtler forms of bias, such as a social worker taking more interest in and collecting more information on the family background of someone they know to be a perpetrator. The use of information collected before the onset of perpetration also helps establish a clear temporal order between the risk factor and sexual perpetration; and therefore the design can meet the temporal precedence criterion of causality. For the same reasons the use of information gathered before the onset of perpetration provides a realistic and stringent test of the capacity of a risk index to predict which males are at risk of perpetration. If perpetrators invent experiences of risk factors then the use of information gathered after perpetration could artificially inflate the predictive capacity of a risk index. They may also switch the temporal order of putative risk factors and onset date, which would also suggest that a risk had predictive value when in fact it did not.

For those subjects who were not perpetrators, information on the risk factors was based on information written before the subject’s 16th birthday. If a subject began perpetrating after this date, the risk score would be based on information written before the 16th birthday. This helped to ensure that comparable amounts of information were available on the two groups. If both the clinical and social service files ended before the subject turned 16 or the onset of perpetration, the end of the case material would be the end-point for scoring.

11.4 Ethical approval

Ethical approval for the study was given by the Great Ormond Street Hospital for Children NHS Trust / Institute of Child Health Research Ethics Committee (see appendix 2). The Association of Directors of Social Services gave its backing to the study and permission was also given by the Director of Social Services or the Child Protection Co-ordinator of each of the social service departments involved. The Office of National Statistics gave approval for the use of medical records for
research purposes (see appendix 3). The Crime Committee of the Association of
Chief Police Officers gave its support to the research, and the Chief Constables of
the constabularies involved gave permission to obtain caution and conviction data
(see appendix 4).

11.5 Sample

11.5.1 Inclusion-exclusion criteria

The sample was generated from records held at the Department of Psychological
Medicine, Great Ormond Street Hospital for Children, NHS Trust. The inclusion-
exclusion criteria for the sample are given below, and, where necessary, the reasons
for a criterion are explained.

11.5.1.1 Stage one
Males under the age of 18 referred between 1980 and 1992 to the Child Sexual
Abuse Team at the Department of Psychological Medicine, Great Ormond Street
Hospital were considered for inclusion in the study. In addition, male children
mentioned in the hospital records as victims or perpetrators of child sexual abuse
were also included.

The start and end-dates of this time frame were selected for specific reasons. The
Department of Psychological Medicine first received referrals for childhood sexual
abuse in 1980. Children referred during or after 1992 were included in a hypothesis
generating study that informed the selection of risk factors in the research reported
in this thesis. These subjects were excluded to ensure an independent assessment of
the risk factors.

11.5.1.2 Stage two
The subject had to be aged over 18 on 01.05.99. The age criteria were imposed to
increase the homogeneity of the sample.
11.5.1.3 Stage three
The Great Ormond Street Hospital file had to contain sufficient information about the subject to enable the social service file and criminal record to be traced. Sufficient information was defined as the file recording at least the first name, last name and date of birth of either the subject or a sibling of the subject.

11.5.1.4 Stage four
The social service department believed to be holding a subject’s social service file had to agree to take part in the study. Subjects in areas that declined to take part were excluded.

11.5.1.5 Stage five
Subjects whose social service file could not be located were excluded from the sample. This and the previous criterion were imposed to ensure that subjects in the final sample had comparable amounts of risk factor information available on them.

11.5.1.6 Stage six
A number of subjects were excluded after the researcher visited the social service department to examine the file. The possible reasons for exclusion at this stage were as follows:

- The files contained insufficient information to classify the subject as a victim of sexual abuse. These subjects were excluded whether they were perpetrators (non-victimised perpetrators) or not (non-victimised non-perpetrators). The definition of sexual victimisation used in this study is given in section 11.6.3.
- Both the hospital and social services did not have contact with the subject until after the onset of perpetration, and therefore there was no information on the risk factors that predated this onset.
- Although the social service department reported that a file existed, when the researcher went to examine the file it could not be located. These departments
had computerised records indicating that the file existed, but the file could not be physically located.

- Miscellaneous reasons (including subject no longer living in England and subject turned out to be female).

11.5.1.7 Stage seven

In a number of cases more than one subject mentioned in a file met the above criteria. Here one subject from each file was selected at random to ensure that the assumption of statistical independence was not violated.

Figure 11.2 summarises the inclusion-exclusion criteria, and reports the number of subjects excluded at each stage.

11.5.2 Comparison of larger samples with final sample

Ideally, key characteristics of the final sample (stage 7) would be compared with the same characteristics of one or more of the larger samples. However, this was not possible because of a lack of detailed information about the excluded subjects. The majority of the subjects were excluded because of the absence of or inability to access a social service file. Without these files the only information on these subjects was their Great Ormond Street file, which often contained only sketchy information. The methodological implications of this are discussed in chapter 14.

11.5.3 Description of final sample

This section summarises the descriptive characteristics of the final sample (N = 104). Information is reported separately for those males who subsequently perpetrated (victim-perpetrators; N = 21) and for those who did not do so (victim-only; N = 83).
Stage 1
Starting sample
N = 601

Stage 2
Aged > 18 (01.05.99)
N = 449
Aged < 18 (01.05.99)
N = 152

Stage 3
Sufficient information to trace
N = 427
Insufficient information to trace
N = 22

Stage 4
Social service department agreed to take part
N = 408
Social service department refused to take part
N = 19

Stage 5
Social service department reported the file exists
N = 292
Social service department could not locate a file
N = 116

Stage 6
File kept
N = 139

File excluded
N = 153
Died = 1
Left country = 1
Female = 2
SSD lost file = 44
Not a victim = 61
No predating information = 44

Stage 7
Selected at random
N = 104
Not selected at random
N = 35
11.5.3.1 Demographic factors

**Age (on 01.05.99)**

The age of the subjects (on 01.05.99) was calculated for the two groups. The victim-perpetrator group (age = 23.29; sd = 4.28; range = 18.11 to 32.19) was approximately 1 ½ years older than the victim-only group (age = 21.67; sd = 2.84; range = 18.07 to 30.76), and this difference approached statistical significance (t = -1.65, df = 24.62, p = 0.11). However, the two groups were no different in terms of their age at referral to Great Ormond Street Hospital (victim-only mean age = 10.68; sd = 3.0; range = 4.58 to 16.50) (victim-perpetrator mean age = 10.74; sd = 3.69; range = 5.0 to 19.33) (t = -0.08, df = 95, p = 0.94). This combination of findings indicates that the victim-perpetrators were referred to Great Ormond Street Hospital on average 1½ years before the victim-only group.

**Ethnicity**

The clinical and social service files provided details of ethnicity for only 18 of the 104 subjects. However, ethnicity is routinely recorded in criminal records, and so was available for the 50 subjects who had received a caution or a conviction. Data on ethnicity was therefore available for 68 of the 104 (65.38%) subjects, including 50 (60.24%) of the victim-only group and 18 (85.71%) of the victim-perpetrators.

84.0% (N = 42) of the victim-only group were Caucasian, 2.0% (N = 1) was Asian, 6.0% (N = 3) were African-Caribbean, and 8.0% (N = 4) were of mixed ethnicity. Of the victim-perpetrator group, 94.4% (N = 17) were white and 5.3% (N = 1) were African-Caribbean. Ethnicity was re-coded as a dichotomous variable (Caucasian / Not Caucasian) because of the small cell sizes. The difference in proportions was not statistically significant ($X^2$ with continuity correction = 0.51, df = 1, p = 0.47)

**Socio-economic status**

Information on socio-economic status was available for only 29 of the 104 subjects (27.9%), including 25 (30.1%) subjects in the victim-only group and 5 (23.8%) in the victim-perpetrator group. For this reason, the SES status of the two groups was
not calculated. Previous studies using subjects referred to the Child Sexual Abuse Team at the Department of Psychological Medicine, Great Ormond Street Hospital indicate that referrals to this Team tend to come from the lower SES groups (e.g. Skuse et al., 1998).

Abode at time of referral
Over half of both groups were living with their family at the time of referral (victim-only = 65.1%; victim-perpetrator = 52.4%). 18.1% of the victim only group were living in a Children’s Home, and the remainder (16.9%) were living with a foster family. 23.8% of the victim-perpetrators were living in a Children’s Home and the same proportion were living with a foster family. The small cell sizes meant that the data were collapsed for statistical analysis (Living with family / Not living with family). The difference between the two groups was not statistically significant ($X^2$ with continuity correction = 0.67, df = 1, p = 0.41).

Geographical location at time of referral
Similar proportions of the two groups (victim-only = 37.3%; victim-perpetrators = 38.1%) were living in Greater London at the time of referral. 50.6% of the victim-only group were living in the Home Counties and 12.0% were living in the provinces. This compares to 33.3% of the victim-perpetrators living in the Home Counties and 28.6% living in the provinces. The differences in geographical location at referral approached statistical significance ($X^2$ with continuity correction = 4.04, df = 2, p = 0.13).

Summary
Table 11.1 summarises the comparison of the victim-only and victim-perpetrator group in terms of the demographic factors. As indicated in table 11.1, there were no statistically significant differences between the two groups on any measure. However, age (on 01.05.99) and geographical location had p values of less than 0.20. These two variables were controlled for in the multivariate analysis.
### Table 11.1 – Demographic information on victim-only and victim-perpetrator group

<table>
<thead>
<tr>
<th>Demographic factor</th>
<th>Victim-only</th>
<th>Victim-perpetrator</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (on 01.05.99)</td>
<td>21.67 yrs (sd = 2.84)</td>
<td>23.29 yrs (sd = 4.28)</td>
<td>( t = -1.65, \text{df} = 24.62, p = 0.11 )</td>
</tr>
<tr>
<td></td>
<td>(range = 18.07 to 30.76)</td>
<td>(range = 18.11 to 32.19)</td>
<td></td>
</tr>
<tr>
<td>Age at referral</td>
<td>10.68 yrs (sd = 3.0)</td>
<td>10.74 yrs (sd = 3.69)</td>
<td>( t = -0.08, \text{df} = 95, p = 0.94 )</td>
</tr>
<tr>
<td></td>
<td>(range = 4.58 to 16.50)</td>
<td>(range = 5.0 to 19.33)</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Caucasian = 84%</td>
<td>Caucasian = 94.4%</td>
<td>( \chi^2 = 0.51, \text{df} = 1, p = 0.47 )</td>
</tr>
<tr>
<td>Abode</td>
<td>With family = 65.1%</td>
<td>With family = 52.4%</td>
<td>( \chi^2 = 0.67, \text{df} = 1, p = 0.41 )</td>
</tr>
<tr>
<td>Geographical location</td>
<td>Greater London = 37.3%</td>
<td>Greater London = 38.1%</td>
<td>( \chi^2 = 4.04, \text{df} = 2, p = 0.13 )</td>
</tr>
<tr>
<td></td>
<td>Home Counties = 50.6%</td>
<td>Home Counties = 33.3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Provinces = 12.0%</td>
<td>Provinces = 28.6%</td>
<td></td>
</tr>
</tbody>
</table>

### 11.5.3.2 Professional involvement

#### Source of referral

The majority of the victim-only group (79.3%) were referred to Great Ormond Street Hospital by social services. Of the remainder, 2.4% were referred by a *Guardian ad litem*, 12.2% by medical/mental health professionals and 6.1% by parents or carefigures (\( N = 82; 1 \text{ missing case} \)). A similarly high proportion of the victim-perpetrator group (71.4%) were referred by social services, 9.5% were referred by medical/mental health professionals and 19.0% were referred by parents or carefigures. Because of the small cell sizes the data were collapsed (referred by social services / referred by other source). The difference between the two groups was not statistically significant (\( \chi^2 \) with continuity correction = 0.28, df = 1, \( p = 0.63 \)).
Social service involvement

Similar proportions of both groups were on care orders at the time of referral (victim only = 43.4%) (victim-perpetrators = 33.3%) ($X^2$ with continuity correction = 0.34, df = 1, p = 0.56), and approximately half of each group were on the Child Protection Register (victim-only = 50.6%) (victim-perpetrator = 57.1%) ($X^2$ with continuity correction = 0.08, df = 1, p = 0.77). The victim-only group were known to social services on average 1 year earlier than the victim-perpetrators (victim-only = 6.24, sd = 4.68, range = 0 to 15.83) (victim-perpetrator = 7.55, sd = 4.28, range = 0 to 14.08); however, the difference was not statistically significant ($t = -1.17$, df = 102, p = 0.25).

Legal status

Over one quarter of both groups had been made Wards of Court (victim-only = 27.7%) (victim-perpetrator = 28.6%) ($X^2 = 0$, df = 1, p = 1.0).

Educational statementing

A substantial minority of both groups were educationally statemented at some point during their school career (victim-only = 16.9%; victim-perpetrator = 23.8%). The difference between the two groups was not statistically significant ($X^2$ with continuity correction = 0.18, df = 1, p = 0.68).

Contact with Great Ormond Street Hospital

Although all of the subjects were referred to GOSH, only a proportion was assessed at the hospital. Approximately half of both groups (victim-only = 53.0%; victim-perpetrators = 47.6%) were seen at Great Ormond Street Hospital ($X^2$ with continuity correction = 0.04, df = 1, p = 0.84). The mean age of the victim-only group (N = 78) at the time of referral to the hospital was 10.68 (sd = 3.00; range = 4.58 to 16.50), which is similar to the age of the victim-perpetrators (N=19) who were aged 10.74 (sd = 3.70; range = 5.00 to 19.33) ($t = -0.08$, df = 95, p = 0.94).
Treatment

Half (50.6%) of the victim-only group had received some form of psychological treatment in connection with the experience of child sexual abuse. This compares to 38.1% of the victim-perpetrators who had received some form of psychological treatment for their sexual victimisation before the onset of their sexually abusive behaviour. This difference is not statistically significant ($\chi^2$ with continuity correction = 0.61, df = 1, $p = 0.44$).

Summary

Table 11.2 summarises the comparison of the victim-only and victim-perpetrator group for professional involvement. Across a number of measures the two groups appear to be broadly comparable.

Table 11.2 – Professional involvement for victim-only and victim-perpetrator group

<table>
<thead>
<tr>
<th>Professional involvement</th>
<th>Victim-only</th>
<th>Victim-perpetrator</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source of referral</td>
<td>Social services = 79.3%</td>
<td>Social services = 71.4%</td>
<td>$\chi^2 = 0.28$, df = 1, $p = 0.63$</td>
</tr>
<tr>
<td>Care Order</td>
<td>Care order = 43.4%</td>
<td>Care order = 33.3%</td>
<td>$\chi^2 = 0.34$, df = 1, $p = 0.56$</td>
</tr>
<tr>
<td>Child Protection Register</td>
<td>Registered = 50.6%</td>
<td>Registered = 57.1%</td>
<td>$\chi^2 = 0.08$, df = 1, $p = 0.77$</td>
</tr>
<tr>
<td>Ward of Court</td>
<td>Ward = 27.7%</td>
<td>Ward = 28.6%</td>
<td>$\chi^2 = 0$, df = 1, $p = 1.0$</td>
</tr>
<tr>
<td>Educationally statemented</td>
<td>Statemented = 16.9%</td>
<td>Statemented = 23.8%</td>
<td>$\chi^2 = 0.18$, df = 1, $p = 0.68$</td>
</tr>
<tr>
<td>Assessed at Great Ormond Street Hospital</td>
<td>Assessed = 53.0%</td>
<td>Assessed = 47.6%</td>
<td>$\chi^2 = 0.04$, df = 1, $p = 0.84$</td>
</tr>
<tr>
<td>Psychological treatment</td>
<td>Treatment = 50.6%</td>
<td>Treatment = 38.1%</td>
<td>$\chi^2 = 0.61$, df = 1, $p = 0.44$</td>
</tr>
</tbody>
</table>

11.6 Instruments and measures

The sheer volume of case material presented a problem for this study. To further complicate matters, the information written by social workers and clinicians was not
gathered with this particular research study in mind, and so information on risk factors was not recorded in a systematic or easily identifiable way. A Risk Factor Manual and a computerised aid for the extraction of case material were developed to facilitate the reduction of qualitative information to quantitative risk scores. The following sections describe the procedure used to develop the Risk Factor Manual, the structure of that manual and the computer program designed to be used alongside it. The complete Risk Factor Manual can be found in appendix 5.

11.6.1 Development of the Risk Factor Manual


An iterative process of piloting and redrafting was used to refine the manual. Piloting consisted of extracting and scoring information from Great Ormond Street Hospital files held on cases referred to the Department of Psychological Medicine, but who were not included in the sample used in this thesis. The cases used in the piloting exercise included sexually victimised females referred between 1980 and 1992, and male and female victims referred after 1992. The manual was generated and refined independently of cases that were used to establish reliability, because this ensured that the reliability was generalisable beyond the cases on which piloting took place. Each of the researchers involved in the extraction and scoring of case material (including the author) would independently extract information from the pilot file and then score it. (Further details of the procedure used to extract and score
The researchers would then meet to compare the information they had chose to extract, the scores they had assigned to each of the risk factors, and to discuss any discrepancies and problems. Consensus decisions were made about changes to the manual, which was then updated. This procedure was repeated until no new substantial difficulties emerged.

The final draft of the Risk Factor Manual (see appendix 5) consists of an introduction section discussing general guidelines for extracting and coding case material, and a separate chapter for each of the risk factors. The structure of each of these chapters is summarised in the following sections.

11.6.1.1 Definitions
The first section part of each chapter provides definitions of the risk factor and other key terms needed for scoring that factor (e.g. the child sexual abuse chapter also includes a definition of the genital area and penetration).

11.6.1.2 Perpetrator criteria
The criteria that a person must meet before he or she can be considered a perpetrator of that risk factor.

11.6.1.3 Correct factor
The definitions of the factors are supported by this section, which provide guidance on the allocation of maltreating behaviours that caused difficulties during piloting.

11.6.1.4 Potential difficulties
Piloting also identified a number of additional difficulties for extracting information about a particular risk factor or scoring that factor; this section was designed to clarify these difficulties (e.g. Witnessing Intrafamilial Violence provided a distinction between intrafamilial violence and physical chastisement of siblings).
11.6.1.5 Dichotomous measurement
This section summarised the criteria that must be met before the risk factor could be scored as present.

11.6.1.6 Severity levels
For a number of the risk factors the severity of that factor was also measured. The severity levels described the behaviours that would be scored at a particular level.

11.6.1.7 Duration
This section summarised the procedure for scoring the duration of the subject's contact with the perpetrators of the risk factor. A full description of the procedure for scoring duration is given in the introductory section of the manual.

11.6.2 The Computerised Extraction Program
The Risk Factor Manual evolved alongside the Computerised Extraction Program, and was designed to facilitate the collection, storage and retrieval of material from the case files. The program runs on a Windows 95 ACCESS format; a copy of the program can be obtained from the author. The following sections gives a description of how the program operates.

11.6.2.1 Cover sheet
Basic demographic information on the subject (e.g. name, date of birth) and the case file (e.g. start date and end-date of file) is recorded on a computerised cover sheet.

11.6.2.2 Quote screen
The researcher extracting information reads a file until he or she reaches a quote meeting the manual's criteria for extraction. This quote is then typed into the text box of the Quote Screen. The risk referred to in this quote is recorded using a series of electronic tick boxes. The researcher also records the date of the document from
which the quote is extracted. This procedure is repeated until the researcher has read the entire file.

11.6.2.3 Review screen

While reading the file the researcher, by double clicking on a risk factor, can access all the extracted quotes referring to that factor. This allowed the researcher to establish what information has been extracted and what is still needed.

11.6.2.4 Care table

The researcher uses the quotes extracted about the subject’s care history to generate a care table. Each change in the group of carers that lasted for one month or more is recorded in the care table. For each group of carers, the researcher records the names of each carer, the relationship of the subject to each carer (e.g. step father, foster mother), the date the care set-up began, the date it ended and the reason it ended. The researcher also records the relevant quotes from the case file that are used in building that part of the subject’s care history. The introduction section to the Risk Factor Manual (appendix 5) provides further details of the procedure for constructing the care table.

11.6.2.5 Example

As an example of how the program was used in this study, consider the following hypothetical quote:

'Jonathan reported that since his mother’s boyfriend moved in early last year, the boyfriend has often punched both him and his mother. Jonathan reported that on one occasion the boyfriend gave him a black eye.'

(Extract of a letter from the head teacher of subject’s school to social services, dated 24.09.87)

Once the researcher had read this sentence he or she would decide whether this quote needed extracting on the basis of the guidelines given in the manual. The
sentence contains information about the subject’s experience of physical abuse, witnessing intrafamilial physical abuse and care history. Therefore the manual states that this quote should be extracted unless the researcher had already extracted previous quotes that contained this information. The researcher would also need to consider whether the quote might change the scoring of any of the risk factor variables. For example, this quote contains information about bruises to the subject’s face, which would lead to a higher physical abuse severity score than a report of being punched.

If the quote meets the criteria for extraction it would be entered in the on-screen text box. The electronic tick boxes for ‘experiencing physical abuse’, ‘witnessing intrafamilial physical abuse’ and ‘care history’ would be ticked to indicate that this quote contains information of relevance to these risk factors. The date of the letter would also be recorded. This is essential to establish whether the risk factors referred to in this quote preceded the onset of perpetration.

11.6.3 Definition of child sexual abuse

The definition of child sexual abuse used in this study required three criteria to be met: 1) the subject was under the age of 16 years at the time of the abuse, 2) the abuse involved sexual contact between the subject and the perpetrator (see definition of sexual contact given below), and 3) either the perpetrator was more than two years older than the subject or the perpetrator used physical coercion (see definition of physical coercion given below).

11.6.3.1 Sexual Contact

The criterion for sexual contact is met if the act involves physical contact between the subject and the perpetrator and that contact involves attempts, successful or otherwise, to perform sexual acts with the victim. Sexual acts include non-genital contact (e.g. the perpetrator ‘French kisses’ the subject), non-penetrative genital contact (e.g. masturbation of the perpetrators by the subject) and penetrative sex.
11.6.3.2 Physical coercion

Physical coercion is defined as physical behaviour designed to ensure compliance in a sexual act against the subject's wishes. This must include contact either between the perpetrator and the subject or an object and the subject, provided that the perpetrator has deliberately and maliciously caused this contact, but it does not necessarily involve marks being left on the subject. For example, physical coercion would be assumed if the perpetrator had grabbed the subject and pushed him into the room where the sexual assault took place.

Further details of the definition of child sexual abuse can be found in the Risk Factor Manual (appendix 5).

11.6.4 Definition of sexual perpetration

Evidence of perpetration was taken from the clinical and social service files, and also police caution and conviction records held on the subjects. Sexual perpetration was restricted to perpetration directed against children, and therefore excluded rape.

The criteria used to classify a subject as a perpetrator are described in detail in the Risk Factor Manual (appendix 5). The criteria are summarised in brief here. Note that reference to physical coercion in these criteria uses the same definition as that given above for child sexual abuse.

The subject was coded as a sexual perpetrator if one or more of the following criteria were met:

11.6.4.1 Criterion one

The subject had a caution or conviction for one or more of the following offences:

- Rape of male/female under 14
- Rape of male/female under 16
- Unlawful intercourse with male/female under 14
- Unlawful intercourse with male/female under 16
- Indecent assault on male/female under 14
- Indecent assault on male/female under 16
- Gross indecency with a child

An offence of incest was classed as sexual perpetration if it was clear that the subject was two or more years older than the victimised family member.

11.6.4.2 Criterion two

Based on social service and clinical records, a subject aged over 12 years at the time of the event was considered a perpetrator if:
- The victim was aged under 16 at the time of the event
- The abuse involved physical contact of a sexual nature between the subject and the victim (non-genital, genital or penetrative)
- Either the subject was two years or more older than the victim or the subject employed physical coercion.

11.6.4.3 Criterion three

On the basis of social service and clinical case material a subject aged 6 years or more, but less than 12 years, was considered a perpetrator in three possible ways:

1) Genital contact
- At least one of the following applies:
  - The subject had direct contact with victim’s unclothed genitals
  - The victim was forced to touch the subject’s unclothed genitals
- For either situation at least one of the following applies:
  - The subject was two years or more older than the victim and the event occurred more than twice
  - The subject used physical coercion
2) Penetration
- At least one of the following applies:
  - The victim was penetrated (vaginal, anal or oral) by the subject
  - The subject was penetrated by the victim (anal or oral)
- For either situation either the subject was two or more years older than the victim or the subject used physical coercion

3) General statement
- Either a clinician or a social worker states that the subject sexually abused another child.

11.6.4.4 Determining the onset of perpetration

The integrity of the longitudinal design required the researcher to establish the date on which a person began sexually abusing others. For the causal question, the onset date is necessary to ensure that the risk factor occurred prior to this date (temporal precedence), and to ensure that a perpetrator did not elaborate or invent risk factor after onset in an attempt to excuse their abusive behaviour (non-spuriousness). The prediction question also requires an accurate onset date, because the use of retrospective reports could artificially inflate the predictive capacity of a risk index.

The case material reported the onset of perpetration in a variety of ways. These included reports of the subject’s or victim’s age (e.g. ‘Jane reports that Jonathan began abusing her when she was seven’); the length of time since its onset (e.g. ‘Jane reports that Jonathan began abusing her about two years ago); a specific time or date (e.g. ‘Jane said it first happened in the summer holidays before she started her new school.’). Details of how this information was used to establish a date of onset are given in detail in the Risk Factor Manual (appendix 5). As these examples indicate, the onset of perpetration could be reported with varying degrees of accuracy. In brief, when faced with vague information the general approach was to take the earliest date of onset. This approach was designed to minimise the
possibility of the risk factor score using information that occurred after the onset of perpetration.

**11.6.5 Measurement of the risk factors**

A number of the potential additional risk factors discussed in chapter 6 could not be used in the study, because the clinical and social service files lacked sufficient detail to score these factors. The following risk factors were excluded: social rejection, intelligence, non-sexual antisocial behaviour and psychopathology. The implications of this are discussed in chapter 14. Pervasive sexual atmosphere was also excluded because early on in the extraction phase it became clear that there was no variance on this measure (all subjects had been exposed to this risk factor).

Definitions of the risk factors used in the study are summarised in table 11.3; full details are given in the Risk Factor Manual (appendix 5).

Each of the risk factors was measured in a number of ways (e.g. dichotomous, severity, duration). The severity measures are based on the work of Barnett, Manly and Cicchetti (1993). These measures use a six-point scale ranging from 0 (maltreatment factor not present), through 1 (low severity) to 5 (high severity). The authors define severity as the relative seriousness of the act with regard to the potential negative psychological impact that a caregiver's act may have on the child's socio-emotional development (Barnett, Manly, & Cicchetti, 1993). Dichotomous measures are based on a recoding of the severity measure (0 = absent; severity level 1 to 5 = present). Duration was calculated on the basis of the length of contact between the subject and the perpetrator or perpetrators. Note that it is designed as a measure of contact with the perpetrator rather than a measure of the duration of maltreatment. This proxy measure was necessary because the files often did not contain sufficient information to score the duration of maltreatment. The quantitative risk factor measures derived from the case files are described in detail in
the Risk Factor Manual (appendix 5). A brief summary of the measures is given in table 11.4.

**Table 11.3 – Definition of risk factors**

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiencing physical abuse</td>
<td>Physical abuse is defined as physical behaviour that does, or has the potential to do, physical harm to the subject. This must include contact either between the perpetrator and the subject or an object and the subject, provided that the perpetrator has caused this contact.</td>
</tr>
<tr>
<td>Witnessing intrafamilial physical abuse</td>
<td>Witnessing is defined as the subject being either visually or aurally exposed to, but neither a recipient nor a perpetrator of, an event of intrafamilial physical abuse.</td>
</tr>
<tr>
<td>Neglect (failure to provide)</td>
<td>Neglect (failure to provide) is scored when a caregiver fails to exercise a minimum degree of care in meeting the child’s physical needs. Neglect (failure to provide) covers 5 domains: supplying the child with adequate food; ensuring that the child has clothing that is sanitary, appropriate for the weather and permits the child freedom of movement; providing adequate shelter; ensuring adequate medical, dental, and mental health care; and ensuring the child’s adequate hygiene.</td>
</tr>
<tr>
<td>Neglect (lack of supervision)</td>
<td>Neglect (lack of supervision) is scored when a main caregiver does not take adequate precautions to ensure a child’s safety in and out of the home, given the child’s particular emotional and developmental needs. There are four broad elements that caregivers may violate to jeopardise children’s physical safety: supervision (failing to take steps to ensure that the child is engaging in safe activities); environment (failing to ensure that the child is playing in a safe area); substitute care (failing to provide for adequate substitute care in the caregiver’s absence or mental or physical incapacity); and developmental needs (failing to recognise the developmental needs of the child in providing adequate supervision to ensure the child’s safety).</td>
</tr>
<tr>
<td>Rejection by carers</td>
<td>This risk factor is often underlined by the carer’s dislike of or hostility towards the child. This may be manifest behaviourally in several ways: blaming (including references to blaming the subject, ‘scapegoating’, ‘targeting’ and negative misattribution); the subject is perceived as deserving harsh discipline and punishment (expression of hostility towards the subject); degradation (including references to degrading the subject or criticising the subject); rejection (including references to rejection, hostility towards the subject and threats to remove the subject from the home or kill the subject); and locking up or containing the child.</td>
</tr>
<tr>
<td>Separations</td>
<td>A separation occurs when the subject is physically separated from one or more primary carers for a period of one month or more. A primary carer is a person who has ultimate responsibility for the subject or who shares this responsibility with one other person, and has cared for the subject for one year or more.</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Child sexual abuse</strong></td>
<td></td>
</tr>
<tr>
<td>Severity</td>
<td>1 = non-genital contact; 2 = genital contact; 3 = one form of penetration (e.g. anal sex only); 4 = two forms of penetration (e.g. oral and anal sex); 5 = penetration plus physical force</td>
</tr>
<tr>
<td>Duration of contact with perpetrator/s</td>
<td>Duration of contact with perpetrator or perpetrators, taking account of any gaps in contact.</td>
</tr>
<tr>
<td>Sexual abuse by a female</td>
<td>A dichotomous measure of whether the subject was sexually victimised by one or more females.</td>
</tr>
<tr>
<td><strong>Experiencing physical abuse</strong></td>
<td></td>
</tr>
<tr>
<td>Dichotomous</td>
<td>Dichotomous measure based on a recoding of the severity level measure (0 absent; 1 to 5 severity = present).</td>
</tr>
<tr>
<td>Severity</td>
<td>0 = not physically abused; 1 = minor marks below the neck; 2 = Non-minor marks below the neck; 3 = Marks to the neck or above, serious bruising, minor lacerations, minor burns; 4 = Serious injuries (e.g. choked, smothered, second degree burns, hospitalised &lt; 24 hours); 5 = &gt; 24 hours hospitalisation or permanent damage</td>
</tr>
<tr>
<td>Duration of contact with perpetrator/s</td>
<td>Duration of contact with perpetrator or perpetrators, taking account of any gaps in contact.</td>
</tr>
<tr>
<td><strong>Witnessing intrafamilial physical abuse</strong></td>
<td></td>
</tr>
<tr>
<td>Dichotomous</td>
<td>Dichotomous measure based on a recoding of the severity level measure (0 absent; 1 to 5 severity = present).</td>
</tr>
<tr>
<td>Severity</td>
<td>Scored as the experiencing physical abuse measure of severity.</td>
</tr>
<tr>
<td>Duration of contact with perpetrator/s</td>
<td>Duration of contact with perpetrator or perpetrators, taking account of any gaps in contact.</td>
</tr>
<tr>
<td><strong>Neglect (failure to provide)</strong></td>
<td></td>
</tr>
<tr>
<td>Dichotomous</td>
<td>Dichotomous measure based on a recoding of the severity level measure (0 absent; 1 to 5 severity = present).</td>
</tr>
<tr>
<td>Severity</td>
<td>0 = not neglected (failure to provide); 1 = e.g. subject is washed by carer infrequently; 2 = e.g. carer sometimes fails to provide adequate meals for subject; 3 = e.g. carer does not treat rat or vermin infestation of house; 4 = e.g. faeces present in living areas of house; 5 = e.g. subject has non-organic failure to thrive</td>
</tr>
<tr>
<td>Duration of contact with perpetrator/s</td>
<td>Duration of contact with perpetrator or perpetrators, taking account of any gaps in contact.</td>
</tr>
</tbody>
</table>

(Continued)
Table 11.4 – Summary of risk factor measures (continued)

<table>
<thead>
<tr>
<th>Neglect (lack of supervision)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dichotomous measure based on a recoding of the severity level measure (0 absent; 1 to 5 severity = present).</td>
</tr>
</tbody>
</table>

Severity
0 = no neglect (lack of supervision); 1 = e.g. carer leaves the eight year old subject unsupervised for three hours; 2 = e.g. carer leaves one year old subject in the care of a nine year old; 3 = e.g. carer leaves subject unsupervised for a whole day; 4 = e.g. carer allows a six year old subject to play unsupervised near a major road; 5 = e.g. carer makes a nine year old subject leave home with no alternative living arrangements
These are examples only; they are not the full contents of each level. Full details of each level are given in the Risk Factor Manual (appendix 5).

Duration of contact with perpetrator/s
Duration of contact with perpetrator or perpetrators, taking account of any gaps in contact.

Rejection by carers
Present / absent rating, rated as present if the carer exhibits one or more of a list of reject ing behaviours, including: scapegoating, hostility, degradation, threats to kill the subject, expression of dislike.

Duration of contact with perpetrator/s
Duration of contact with perpetrator or perpetrators, taking account of any gaps in contact.

Separations
One or more separations
Coded as present if the subject experienced one or more separations before either age 16, end of case material or onset of perpetration, whichever took place first.

Number of separations
The total number of separations from primary carers before either age 16, end of case material, onset of perpetration, whichever took place first. Simultaneous separation from two primary carers counted as two separations.

11.7 Reliability

The reliability of data extraction and coding by each member of the research team was assessed on a subgroup of subjects from the study sample. Reliability was assessed at two time points: before the extraction of data from the files and before the coding of the files. Full details of the procedure for establishing reliability are given in chapter 12.
11.8 Structure of the study

The research study involved a number of interrelated stages and this section describes these. A summary of the structure of the study is given in figure 11.10.

11.8.1 Extraction of identifying information from clinical files

Clinical files held at the Department of Psychological Medicine, Great Ormond Street Hospital, were examined to identify the sample and to enable the social service file and criminal records of the sample to be traced. The following information was extracted from each file:

- First, middle and last names of the subject, including all changes in name and variations in spelling
- Date of birth, including all recorded variations
- Most recent addresses of subject, family members, General Practitioner and the social services department involved in the case.

11.8.2 Search of criminal records

Complete criminal records, including caution and conviction data, held by the police forces of England, Wales and Scotland were obtained for all the subjects.

11.8.2.1 Search for conviction records

Convictions are given if a finding of guilt is made by either the Magistrates Court or the Crown Court. A conviction remains on a person’s criminal record for life.

The Metropolitan Police, using the Police National Computer, conducted the search for criminal records. The members of staff who conducted the search decided whether the identity of a particular subject matched with that given on a criminal record. Ideally, the researcher would have made this decision, but this was not
possible because it could mean the police disclosing criminal records that did not refer to one of the subjects.

11.8.2.2 Search for caution records

Cautions are given by the police, but they do not constitute a finding of guilt. They are, however, recorded on a person’s criminal record. They remain there for a length of time that varies according to the police force giving the caution, the age of the subject at the time of the caution and the date on which the caution was received. Until April 1996 in the Metropolitan Police Force, a caution remained active for 3 years or until the person turned 18, whichever took place last. Since this time, the caution has remained active for 5 years or until the subject turned 18, whichever took place last.

The police forces outside the London Metropolitan district have always used a five-year expiry date. It may be possible to obtain information on caution records given more than five years ago, because of variation in the regularity with which expired cautions are removed from the system. Police forces using computerised criminal records are more likely than those using a card system to remove expired cautions after five years.

The Police National Computer is designed to hold all caution records received in Britain. However, at the time the caution searches were conducted, the database did not record all relevant cautions. To obtain a comprehensive list of cautions it was necessary to search the databases of the local police forces of the regions in which the subject has lived over the last five years.

The database held by the Office of National Statistics (ONS) supplied the details of where the subject had lived. The Office of National Statistics uses NHS records to identify the geographical area in which a person is registered with a GP. Each subject was searched in the areas in which he had been living over the last five years, according to this database. If an ONS trace was not made for a subject,
caution records were searched in police forces on the basis of information contained in the clinical files, ONS information for any siblings, and the social service department in which the subject’s file had been located.

As described above, in two situations the caution may remain active for more than five years. This would occur if the subject received a caution at a young age in the Metropolitan Police Force area. It would also occur if the subject received a caution in a local force that did not routinely extract cautions from their system once they expired. Because some subjects may have a caution lasting for more than five years in the Metropolitan area, the entire sample was searched there. For logistical reasons, it was not possible to do this for all of the local forces in which cautions may remain for more than five years.

The police forces covering the regions in which subjects had lived were contacted and asked to take part in the study. The initial approach was made to the Chief Constable of each police force. Of the 31 forces approached, 30 (97%) agreed to take part in the study. The one force that declined to take part did so because of data protection issues. The staff at the local police forces conducted all the caution searches. As with the search of the criminal records, the police staff conducting the search made the decision about whether the identity of a subject matched with the details of a particular criminal record.

11.8.3 Tracing of social service files

Subjects who had sufficient information to permit tracing (first name, last name, date of birth, or sibling with this information) were assigned to a social service department. This assignment was made either on the basis of the social service address given in the file (N = 257; 54%) or, if this information was not available, on the basis of the subject’s address (N = 219; 46%). In cases in which a social service address had not been given the likely social service was contacted and asked if they covered the address.
The social services departments to which an assignment had been made were contacted and invited to take part in the study. 53 social service departments were contacted, including 29 London Borough departments and 24 English departments outside London. 50 of the 53 departments agreed to take part (94%), which excluded 19 of the 476 subjects (4%). Of the three departments that declined to take part, 2 were London Borough Departments and one was outside London. One department refused to take part because of the sensitivity of the cases held by them; two refused because of data protection issues.

For those departments agreeing to take part in the study, an information sheet was generated on each subject that gave details of the subject’s name (including changes and variations), date of birth (including variations) and addresses (subject, family member, GP, social services). This was sent to the assigned social service department to help trace the files.

If a file was not located in the first search, the clinical file would be re-examined for additional information, which would lead to either a second search in the same area using the additional information or a search in a different area. In some cases, the computerised system used by a social service department indicated that the social service file had been held, but had since been moved to another department. In these cases the new department would be contacted and invited to take part in the study.

No limit was placed on the number of times a subject’s file could be searched for in different social service areas; the search continued until no new possibilities were found.

11.8.4 Extraction and scoring of case material

The procedure for extracting and scoring case material involved a division of tasks between researchers and the double coding of risk factor information; this was
essential to minimise the possibility of researcher bias. Before describing the extraction and scoring of material, it is useful to describe this blinding strategy.

11.8.4.1 Blindness

The task of translating qualitative information into a quantitative risk score consisted of three stages: extraction of information from the file, scoring sexual perpetration status and scoring risk factors. For each subject, a different researcher would complete each of the three tasks. If a researcher had extracted the information from a subject’s file, he or she was prohibited from scoring either the subject’s perpetration status or the risk score. If a researcher had scored a subject’s risk score, he or she was prohibited from scoring his perpetration status. Finally, if a research had scored a subject’s perpetration status, he or she was prohibited from scoring the subject’s risk scores.

This division of tasks was necessary to minimise the possibility of researcher bias. Knowledge of a person’s perpetration status may bias the scoring of a risk factor. For example, if a subject was a perpetrator, and it was unclear whether a particular risk factor was present or absent, there may be a tendency to score it as present. In the same way, knowledge of a person’s risk factor scores could bias the coding of perpetration status. If it was unclear whether a subject met the criteria for perpetration, a high or low risk score may bias that decision. If a researcher had extracted information about a subject from a case file, then s/he knew the subject’s perpetration status as well as his experience of the risk factors. Therefore, this research could not be involved in coding that file.

An additional strategy was used to minimise the possibility of researcher bias. One of the reasons the process of coding was divided into two stages, carried out by different researchers, was to ensure that knowledge of perpetration status did not influence the scoring of risk factors. However, one feature of the coding procedure could inadvertently indicate perpetration status. The integrity of the catch-up design required that only information gathered before perpetration could be used in scoring
the risk factors. Therefore, when a researcher was coding the risk factors s/he would be given this 'end-date'. Because the subjects typically began perpetrating before the age of 16, this end-date would be a date sometime before the subject's 16th birthday. In contrast, when scoring the victim-only group there was no need to impose such a cut-off, and the risk factor could use information written up to a subject's 16th birthday. If a risk factor scorer were told to code a file up to the subject's 16th birthday, he or she would be able to guess that the subject was unlikely to be a perpetrator. Whereas, if the researcher were told to code to a date other than the 16th birthday, the researcher would be able to guess that the subject was likely to be a perpetrator. These inferences would not always be accurate. For example, if the file of a subject who had not perpetrated ended before the subject's 16th birthday, this would be used as the end-point. If a perpetrator did not begin to perpetrate until after the age of 16, the end-point would be taken to be 16 years. However, the guess was more likely to be accurate than not, and this could introduce a bias.

The strategy used to avoid this possibility involved the double coding of each file. In the first condition the researcher would score the risk factors using an ‘end-date’ of 16 years. An 'end-date' means that all information written before this point could be used in calculating a risk score, but no information written after that point could be used, even if it referred to events that occurred before that date. This would be the genuine end-date for non-perpetrators whose files ended after their 16th birthday and for perpetrators who began perpetrating after the age of 16. In the second condition the researcher would use an end-date between 6 and 16 years (excluding 16). This second date may be the genuine end-date of the file. For example, if the subject began perpetrating before the age of 16, this end-date would be the date of onset. In addition, if the file of a non-perpetrator ended before the age of 16 years, the end-date of the file would be this date. If 16 was the genuine end-date, the perpetrator coder would use a computer program to generate a random date between the subject's 6 and 16th birthday (excluding 16).
Therefore, for each subject the risk factor scorer had to calculate two sets of risk scores: one in which the end-date was assumed to be 16 and one in which it was assumed to be a date between the subject’s 6 and 16th birthday (excluding 16). One of these dates would be a genuine end-date, and one would be a dummy condition. The researcher scoring the risk factors would be unaware of which of the two dates was the genuine end-date, and so could not guess the subject’s perpetration status. The perpetrator coder would keep a record of the genuine end-date, and the scores coded to this date were the ones used in the analysis.

11.8.4.2 Extraction of information
The extraction of information from clinical files took place in the Behavioural Sciences Unit, Institute of Child Health. The extraction of information from social services files took place in the social service office at which the file had been located. The files were trawled by one of five researchers, including the author. All were trained in using the Risk Factor Manual and the Computerised Extraction Program, and had demonstrated reliability in extraction (see chapter 12).

The allocation of files was based on the geographical accessibility of files for a researcher. Social service departments that were not easily accessible for any of the researchers were allocated at random.

The procedure for extracting information from the social service files is described in the section above on the Computerised Extraction Program.

11.8.4.3 Scoring of information
The scoring of risk factors and sexual perpetration status were allocated to the different researchers using the procedure outlined in section 11.8.4.1 above. The scorer would examine the quotes that could be used and then by consulting the Risk Factor Manual would decide on the score to be assigned to each variable. Full details of the scoring guidelines are given in the Risk Factor Manual (appendix 5).
11.9 Statistical analysis

The procedure for data entry, data cleaning and statistical analysis of the main dataset are described in this section. The statistical analysis procedure for the reliability data is described in chapter 12.

11.9.1 Data entry

For each subject the score using the genuine end-date was entered into an SPSS (Statistics Package for Social Sciences) datasheet. The score using the dummy end-date played no part in the analysis.

11.9.2 Data cleaning

Means and standard deviations were calculated to check that the values of the variables were plausible. In addition, ranges were checked to ensure the values were within the possible range.

The data for 39 randomly selected cases were re-entered to examine the accuracy of data entry. For each of the 39 cases, 155 data points were checked, making a total of 6045 checked data points. 27 data entry mistakes were made, which means that there were 0.45% data entry errors. However, in all cases the mistakes were minor and would not affect the conclusions drawn from the analyses (e.g. a decimal year of 7.82 entered as 7.83); therefore the data entry was assumed to be accurate.

11.9.3 Assessing normality and homogeneity of variance

For all parametric tests, normality and variance were checked before conducting the analysis. The researcher eyeballed histograms of the data in SPSS (Statistics
Package for Social Sciences) to assess normality. Formal measures of kurtosis and skew were not used because these employ an extremely liberal criterion of normal. Distributions that deviate substantially from normal are still assumed to be normal according to these two statistics. Equality of variances was assessed using Levene’s test of equality of variance.

11.9.4 Bivariate analysis of descriptive variables

Categorical descriptive variables were analysed using chi square tests with Yates’s Continuity Corrections. If more than 20% of cells had expected frequencies of less than 5, the levels of the categorical descriptive variable were combined and the new variable reanalysed.

Ordinal measures were analysed using the Mann-Whitney U test. Variables that met parametric assumptions were analysed using $t$ tests. If the assumption of homogeneity of variance was not met, then the $t$ test for unequal variance is reported.

11.9.5 Bivariate analysis of risk factors

Odds ratios and 95% confidence intervals were used as the measure of the bivariate association between the dichotomously measured risk factors and sexual perpetration. Calculations were made by running a logistic regression in SPSS (Statistics Package for Social Sciences). An odds ratio represents the ratio between the odds of an outcome (sexual perpetration) in a group exposed to a risk and the odds of an outcome in a group not exposed to that risk. Odds of greater than one indicate an increased likelihood of the outcome in the exposed group; an odds ratio of one indicates no association.
As with the descriptive variables, ordinal measures were analysed using the Mann-Whitney U test. Variables that met parametric assumptions were analysed using \( t \) tests; the \( t \) test for unequal variance is reported if the assumption of homogeneity of variance is not met.

11.9.6 Multivariate analysis of risk factors

A series of logistic regressions were run to examine the independent contribution of risk factors to the prediction of sexual perpetration status. Adjusted odds ratios and 95% confidence intervals were calculated for each of the variables entered into the regression. These can be interpreted in the same way as unadjusted odds ratios, but represent the association between the variable and sexual perpetration, when other variables are controlled for.

11.9.7 Construction of the Risk Indexes

Beyond the variable-by-variable comparison of the differences between the victim-only and victim-perpetrator groups a series of aggregate risk indexes were generated. It was assumed that sets of related variables could be used to compute a risk index that would have greater discriminatory power than any one of those variables alone.

Four different risk indexes were constructed, including two based on \textit{a priori} predictions and two using a \textit{post hoc} approach.

11.9.7.1 Construction of the Seven-Factor Dichotomous Risk Index

The factors included in the Seven-Factor Dichotomous Risk Index were selected on the basis of the literature review of potential risk factors discussed in chapter 6. The decision about which variables to include was made before the data were scored or
analysed. The Seven-Factor Dichotomous Risk Index consisted of the following factors, each measured as present or absent:

- Sexual victimisation by a female
- Experience of physical abuse
- Witnessing intrafamilial physical abuse
- Neglect (failure to provide)
- Neglect (lack of supervision)
- Rejection by carers
- One or more separations before age 16, first perpetration or end of file (whichever took place first).

Each subject scored 1 for each risk factor that was present; these scores were then summed to give an overall risk score. Scores could range from 0 to 7. A high score indicated that the boy was at high risk of sexual perpetration; a low score indicated that the boy was at low risk.

### 11.9.7.2 Construction of the Seven-Factor Severity Risk Index

A further seven-factor index was generated to permit a comparison between a risk index based purely on dichotomous measures of maltreatment and one based on more than dichotomous measures. The Seven-Factor Severity Risk Index consisted of the same factors used in the Seven-Factor Dichotomous Risk Index, but, when such data existed, a more than dichotomous level of measurement was used:

- Sexual victimisation by a female (rated present / absent)
- Experience of physical abuse (six-point severity scale)
- Witnessing intrafamilial physical abuse (six-point severity scale)
- Neglect (failure to provide) (six-point severity scale)
- Neglect (lack of supervision) (six-point severity scale)
- Rejection by carers (present / absent)
- Number of separations before age 16 / first perpetration / end of file (whichever took place first) (number).
For some of the measures it would be possible to use either severity or duration measures, because both use a more than dichotomous level of measurement. A decision was made not to use the duration measures because of their doubtful validity; the duration variables measure the duration of contact with the perpetrator, which may not be the same as the duration of maltreatment.

Because the seven variables were not all measured using the same levels, the were first transformed to z scores and these scores were then summed. A high score indicated that the subject was at high risk of sexual perpetration; a low score indicated that the subject was at low risk.

The Seven-Factor Severity Risk Index was also generated on the basis of a priori predictions. The decision about which variables to include and how they were to be measured was made before the data were scored or analysed.

11.9.7.3 Construction of shortened forms of the risk indexes

A further set of risk indexes were calculated post hoc to examine whether it was possible to obtain an increased degree of discrimination between the two groups.

One method that might improve discrimination is to generate a risk index based solely on those risk factors that were significant at the bivariate level. Therefore a shortened form of the Seven-Factor Dichotomous Risk Index (Four-Factor Dichotomous Risk Index) was generated by using only those risk factors that were significant or near significant at the bivariate level. Each subject scored 1 if a risk factors was present and these were then summed to give an overall risk score. A higher score indicated an increased probability of perpetration according to the risk index.

The shortened severity risk index was generated in a similar way. This index consisted of those factors that were significant at the bivariate level. As with the seven-factor severity index, the variables were transformed to z scores. A high score
indicated that the subject was at high risk of sexual perpetration; a low score indicated that the subject was at low risk.

11.9.8 Quantifying predictive validity

A number of statistical techniques were used to assess the discriminative capacity of the risk indexes. The difference between the victim-only group and the victim-perpetrator group was analysed using a $t$ test. An effect size was also calculated, which expresses the difference between the means of the two groups in terms of the pooled standard deviation:

$$\frac{\text{mean of victim-perpetrator group} - \text{mean of victim-only group}}{(\text{sd of victim-perpetrator group} + \text{sd of victim-only group}) / 2}$$

Other measures used in the study included sensitivity, specificity, positive predictive value and negative predictive value. The main method used to assess discriminative capacity of the risk index was the ROC-curve. As Rice and Harris (1995) point out, there are a large number of methods for quantifying predictive validity, including percent correctly classified, false positives, false negatives, sensitivity, specificity, the correlation coefficient and relative improvement over chance. However, these methods are limited because the statistic they produce is dependent on the cut-off selected, and some are also influenced by the base rate in the sample (McFall & Treat, 1999). The measures only provide information on the predictive accuracy of the instrument at one particular cutting point. Given that the optimal cutting point for an instrument will vary according to the use to which it is put in a particular setting (Rice & Harris, 1995), it would be useful to have a measure of predictive accuracy that provided a value of accuracy over all cutting points. If samples with different base rates are used to compare the predictive validity of different instruments, then it is necessary to use a measure that is independent of base rates. If the measure is dependent on the base rate then a comparison cannot be made.
The ROC-curve has been recommended as the measure of choice for quantifying predictive accuracy because it is independent of both the base rate and the cutting point (McFall & Treat, 1999; Rice & Harris, 1995). This measure provides a visual description of the overall discriminatory capacity of the risk index by plotting the sensitivity score against 1-specificity for each cutting point. If the index predicted at a level no better than chance then the plot would be a straight line with 50% of the boxed area under the curve, as depicted in Figure 11.3. If an index is to predict at a level better than chance, the ROC-curve must have more than 50% of the area in the box below it, as depicted in figure 11.4.

**Figure 11.3** – ROC-curve with 50% of the area under the line

![ROC-curve with 50% of the area under the line](image)

**Figure 11.4** – ROC-curve indicating prediction better than chance

![ROC-curve indicating prediction better than chance](image)
ROC-curves and 95% confidence intervals for the amount of area under the curve were calculated using the ROCKIT program (Metz, 1998).

11.9.8.1 Further analysis of the Seven-Factor Dichotomous Risk Index

If a risk index achieves a respectable area under the curve score, this does not guarantee that the risk index is of use. As discussed in chapter 10, the usefulness of a risk index is dependant on a number of considerations including:

- The base rate of the outcome (sexual perpetration) in a particular population
- Decisions about the costs and benefits of correct and incorrect classification
- Treatment factors (e.g. different efficacy of treatment and dose-response relationships between the amount and the cost of treatment).

The results section argues that of the four risk indexes discussed in this thesis, the Seven-Factor Dichotomous Risk Index offered the most promise. This index was therefore selected for further analysis, which examined how variations in these factors affected its performance.

The performance of the risk index was examined under three different base rates of sexual perpetration among sexually victimised males: 3%, 10% and 17%. For each of these three base rates, two sets of analyses were then carried out.

The first set used two decision analyses to examine the performance of the risk index under two assumptions about the weighting of utilities. In the first analyses it was assumed that true positives were twice as important as other utilities; in the second, true negative were given twice the weighting of the other utilities. To conduct a decision analysis a $U_{\text{Overall}}$ score is calculated for each cutting point on the risk index (McFall & Treat, 1999):
Figure 11.5 – Equation for a decision analysis

\[ U_{\text{Overall}} = (Pr)(HR)(U_h) + (Pr)(1-HR)(U_m) + (1-Pr)(FAR)(U_{fa}) + (1-Pr)(1-FAR)(U_{cr}) \]

Where:
- \( Pr \) = The probability (base rate) of the outcome (e.g. the proportion of sexually victimised males who become sexual perpetrators in a specific population)
- \( HR \) = Hit rate (sensitivity) for the particular cutting point
- \( FAR \) = False alarm rate (1 – specificity) for the particular cutting point
- \( U_h \) = Utility (benefit) of a hit (true positive)
- \( U_m \) = Utility (cost) of a miss (false negative)
- \( U_{fa} \) = Utility (cost) of a false alarm (false positive)
- \( U_{cr} \) = Utility (benefit) of a correct rejection (true negative)

The \( U_h, U_m, U_{fa} \) and \( U_{cr} \) values reflect the different weightings given to the four utilities. For example, if true positives are twice as important as the other utilities then \( U_h \) is given twice the value as \( U_m, U_{fa} \) and \( U_{cr} \). The cutting point with the highest \( U_{\text{Overall}} \) score is the optimal cutting point for that particular weighting of the utilities.

The second set of analyses examined the influence of treatment factors on the performance of the risk index. Each of the analysis steps described below was repeated for each cutting point on the risk index. The analyses were also repeated assuming the three different treatment efficacies (25%, 50% and 75%).

The first step in this analysis was to identify the proportion of sexually victimised males referred to a service that would be treated given a particular cutting point. For example, if 30% of sexually victimised males scored above a cutting point then for that cutting point 30% would be treated. If a decision was made to treat 30% of a population of sexually victimised males referred to a particular service, then this cutting point would be the appropriate one to use.
The proportion of a referred population who would be treated given a particular cutting point was calculated as follows:

**Figure 11.6 – Equation for proportion treated**

\[
Tr = (Pr)(HR)+(1-Pr)(FAR)
\]

Where:
- \(Tr\) = the proportion treated
- \(Pr\) = The probability (base rate) of the outcome (e.g. the proportion of sexually victimised males who become sexual perpetrators in a specific population)
- \(HR\) = Hit rate (sensitivity) for the particular cutting point
- \(FAR\) = False alarm rate (1 – specificity) for the particular cutting point

The second step of the analysis estimated the proportion of sexually victimised males who would go on to perpetrate if random selection was used for treatment. The estimation also took account of both the base rate of perpetration and efficacy of the treatment. For example, if financial restrictions meant that only 30% of a given population could be treated, and the 30% were selected at random, this stage of the analysis indicated the proportion of the total population that would go on to perpetrate given this base rate and treatment efficacy. This was calculated as follows:

**Figure 11.7 – Equation for proportion perpetrating if treatment population selected at random**

\[
Ra = Pr - (Tr)(Pr)(Ef)
\]

Where:
- \(Ra\) = The proportion of the population of sexually victimised males referred to a service who would go on to perpetrate if the proportion for treatment was selected at random
- \(Pr\) = The probability (base rate) of the outcome (e.g. the proportion of sexually victimised males who become sexual perpetrators in a specific population)
- \(Tr\) = the proportion treated
- \(Ef\) = The efficacy of treatment (set at either .25, .50 or .75; see above for discussion)
In the third step of the analysis an estimation was made of the proportion of sexually victimised males who would go on to perpetrate if selection for treatment used the relevant cutting point on the risk index, rather than selecting at random. For example, if 30% of the population were to be treated, the cutting point above which 30% of the population scored would be identified and people scoring above this would be selected for treatment. This stage of the analysis indicated the proportion of the total population that would go on to perpetrate if this 30% were treated, and again it takes account of the base rate of perpetration and treatment efficacy. This is calculated as follows:

**Figure 11.8** – Equation for proportion perpetrating if risk index used to select treatment population

\[
R_I = Pr - (HR)(Pr)(Ef)
\]

Where:

- \(R_I\) = The proportion of the population of sexually victimised males referred to a service who would go on to perpetrate if the proportion treated were selected using the relevant cutting point of the risk index
- \(Pr\) = The probability (base rate) of the outcome (e.g. the proportion of sexually victimised males who become sexual perpetrators in a specific population)
- \(HR\) = Hit rate (sensitivity) for the particular cutting point
- \(Ef\) = The efficacy of treatment (set at either .25, .50 or .75; see above for discussion)

The final stage of the analysis estimates the reduction in perpetration when the risk index is used to select who is to be treated compared to random selection. This value is expressed as a proportion of the total population of sexually victimised males referred to a particular service. For example, a figure of 1% indicates that for every 100 sexually victimised males referred to a service, the use of the risk index to select the proportion for treatment would lead to 1 less person perpetrating compared to random selection. This value is calculated as follows:
**Figure 11.9** – Equation for improvement of selection using risk index over selection at random

\[
\text{Improvement} = Ra - RI
\]

Where:
- \(Ra\) = The proportion of the population of sexually victimised males referred to a service who would go on to perpetrate if the proportion for treatment was selected at random
- \(RI\) = The proportion of the population of sexually victimised males referred to a service who would go on to perpetrate if the proportion treated were selected using the relevant cutting point of the risk index

As described in chapter 10, these analyses can be used to examine how the improvement varies as a function of the proportion selected for treatment, the base rate of perpetration and the efficacy of treatment. By comparing across different treatment efficacies, we can use the analyses to explore decision-making when there is a dose-response relationship in treatment. For ease of interpretation, each of the equations in figures 11.6 to 11.9 is translated into a percentage when reported in the results section.

**11.10 Summary**

An important methodological feature of the research reported in this thesis is the use of a catch-up longitudinal design. In this design information on risk factors is taken from information gathered in the past for reasons other than the research question. Only information recorded before the onset of perpetration was used in scoring the risk factors. This design was selected because of its ability to address the two central questions of this thesis: *cause* and *prediction*.

The starting sample for the study consisted of sexually victimised males referred to Great Ormond Street Hospital between 1980 and 1992 (\(N = 601\)). A range of inclusion-exclusion criteria were imposed on the sample, and the final sample used in the analysis consisted of 104 subjects. The study used contemporaneous social
service and clinical files to code the risk factors. Information on sexual perpetration was coded from case files and police caution and conviction records. The structure of the study is summarised in figure 11.10.

Two analysis strategies were used to examine the two cause and prediction questions. The causal status of the risk factors was assessed using bivariate and multivariate analyses (logistic regression). The predictive capacity of risk indexes derived from the risk factors was assessed using a variety of techniques including the calculation of ROC-curves.
Figure 11.10 – Structure of the study

1. Generate Manual and Extraction Program
2. Identification of the samples
3. Pilot Manual and Extraction Program
4. Extraction of identifying information
5. Establish extraction reliability
6. Location of social service files
7. Establish scoring reliability
8. Trawling of social service and clinical files
9. Search of Register of Births in England and Wales
10. Scoring of social service and clinical files
11. Office of National Statistics search
12. Search of caution and convictions records
13. Analysis
12 Reliability

12.1 Introduction

The catch-up longitudinal design offers a number of distinct methodological advantages for research in this area, but, like any other design, it has limitations. The main problem is that data are collected for reasons other than the research study. For example, in this study neither the social service nor clinical records were designed to be used for research purposes. In these voluminous records the relevant information is not collected in one place, it is not recorded in an easily identifiable format and information that would help clarify a particular score is not always present. These difficulties raise substantial problems for inter-rater reliability. The catch-up design is of use only if it can be established that the case material -- with all its limitations -- can be translated into reliable quantitative scores.

This chapter discusses the issues involved in establishing reliability, the methods used to do this and the results of the reliability analyses.

12.2 Sources of potential unreliability

Problems with reliability could occur at two points in the study. At the extraction stage different researchers may extract different amounts or types of information from the case files, which would affect the quantitative scores derived from this information. At the scoring stage different researchers may assign different scores, even when they were scoring the same information. Inter-rater reliability, therefore, was assessed in two stages. The first stage assessed the reliability with which information was extracted from the case files (extraction reliability). This assessment was completed before the researchers began extracting information from the case files. After the information had been extracted from the files of all the
subjects, the second stage of reliability took place. This stage assessed the reliability for the scoring of case material (scoring reliability).

An additional reason for conducting reliability at two time points is that the data collection was carried out over a long period of time – 12-months. The demonstration of reliability before the extraction procedure began does not ensure reliability 12-months later, when the scoring stage began.

The long period of time over which case material was collected has an additional implication for the reliability analyses discussed in this chapter. Over this time there were some changes to the way in which some risk factors were measured. The extraction reliability for these variables is not based on the same measures as that used in the analysis, and, therefore, these variables lack extraction reliability. This applies to three variables: sexual abuse by a female, the dichotomous measure of rejection by carers and the duration measure of rejection by carers. Scoring reliability is available for these measures.

12.3 Method

The following sections outline the method and results of the extraction reliability, and the same for scoring reliability. The statistical analysis of inter-rater reliability was common to both stages and will be described first.

12.3.1 Statistical analysis of inter-rater reliability

The reliability of dichotomous variables were analysed using Cohen’s kappa. For a dichotomous variable to be considered reliable, a kappa of at least 0.5 was needed.

An adapted form of Kendall’s tau-b was used to assess reliability of ordinal level measures. Kendall’s tau-b is a measure of association rather than a measure of
reliability. The statistic, if used for reliability purposes, is flawed because it fails to adjust for systematic differences in the scores of the researches. For example, if one researcher consistently scores exactly twice the value scored by the other researcher, this would lead to a perfect association despite there being no reliability. For variables meeting parametric assumptions, this difficulty is dealt with by the calculation of an intraclass correlation rather than a measure of association, such as Pearson $r$. However, no such comparable measure exists for ordinal-level data.

One method of approximating an intraclass correlation for interval-level data is to take researcher A’s scores in the database and copy them below researcher B’s scores, and then copy researcher B’s scores under researcher A’s scores. Pearson’s $r$ is then calculated on these data. Intuitively this makes sense. If researcher B were scoring exactly twice the value scored by researcher A, Pearson’s $r$ would be 1.00 if this procedure were not carried out. However, if this procedure were carried out, half the time researcher A would be scoring half as much as researcher B, and the other half of the time he or she would be scoring twice as much as researcher B. The same applies to researcher B. The best fitting line will be the diagonal, and Pearson’s $r$ will reflect how much the two sets of scores deviate from this diagonal. This same technique was applied to the ordinal level data, but was analysed using Kendall’s tau-b rather than Pearson’s $r$.

In the reliability analysis, no continuous measures met the parametric assumption of normality. Non-normality was caused by a high proportion of subjects scoring 0 on the continuous variables. This means that transformation techniques could not be used to normalise the data – no matter how the data is transformed, there would still be a pile up of scores on the transformed 0 value. All continuous measures were therefore analysed using the adapted form of Kendall’s tau-b described above. It should be noted that this technique artificially doubles the sample size, and, therefore, the reliability statistic, rather than the p value should be the focus of attention. A value of 0.70 was required for acceptable reliability.
12.4 Reliability of extracting case material

Although the aim of the first stage of reliability testing was to assess the inter-rater reliability with which information was extracted from case files, it is difficult to envisage how this could be assessed independently of the reliability of scoring. This is because to test whether two researchers extract the same type of information it is necessary to score that information and to compare the reliability of these two scores. Therefore the assessment of extraction reliability is inevitably an assessment of both extraction and scoring.

The stage of extraction reliability was also designed to assess whether the division of tasks designed to minimise bias (see chapter 11) was feasible. This division was necessary to minimise the operation of bias in the catch-up design, but posed particular problems for the scorer, because he or she had to score data without the benefit of background information gained from reading the file in its entirety. The concern was that it might be difficult for a researcher to score a file without also having read that file. It was possible that important pieces of contextual information, not contained in the quotes, but provided by reading the file, would not be available to the researchers scoring the case material. The manual provided advice to the researcher extracting information about when the scorer would need additional contextual information, but it was still necessary to establish that information could be scored in the absence of this background information. This was assessed by comparing two scorings: one made by a researcher who had also extracted the information from the file (‘extract and score task’) and one made by an individual who had not (‘score-only task’). The ‘score-only task’ was made on the basis of material extracted by a third researcher (‘extract-only task’). If acceptable reliability ratings could be achieved using the division it gave some indication that accurate scoring could be made without background information. It should be emphasised that it provided a particularly stringent test of reliability, because the two scorers were working from different knowledge bases.
Reliability was assessed using the files of 18 subjects selected at random from the sample. Initially, three researchers, including the author, were involved in the extraction and scoring of files. Each one of the three researchers was assigned at random to the 'extract and score task' for six files. Assignment to extract-only and score-only conditions ensured that each researcher would act in these roles to the other researchers in the same proportion of times as each of the other researchers.

At a later date two additional researchers were introduced. It was not possible to ensure that each of the five researchers stood in exactly the same number of relationships to each other in terms of the three tasks. Half of the original 18 files were selected at random and the original researcher’s extract and score task was used. The two additional researchers were then randomly assigned to the two remaining conditions (extract-only and score-only). For the remaining nine files, the two new researchers were randomly assigned to the score and extract condition. The researcher not randomly assigned to the 'extract and score' condition, was assigned to a score-only condition, and scored the material extracted by one of the three original researchers. The results of the extraction stage of reliability are summarised in table 12.1.

Of the 36 reliability estimates, 28 (78%) reached acceptable levels of reliability (kappa > 0.5; Kendall’s tau-b > 0.7). The dichotomous measures and the more than dichotomous measures differed in their reliability. All of the dichotomous measures reached acceptable reliability levels. Indeed, 10 of the 14 kappa statistics (71%) had values above 0.8, and 12 (86%) had values above 0.7. The reliability problems were confined to the more than dichotomous measures. Despite the difficulties with these measures, they were retained because it allowed comparison of the relative predictive capacity of dichotomous measures and more than dichotomous measures of maltreatment. However, the reduced reliability of the severity measures could deflate the predictive capacity of a risk index based on these measures, because of the increased random error in measurement.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Three researchers</th>
<th>Five researchers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sexual perpetration</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Child sexual abuse</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>1.00</td>
<td>0.82</td>
</tr>
<tr>
<td>- Severity (tau-b)</td>
<td>0.65</td>
<td>0.69</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.67</td>
<td>0.85*</td>
</tr>
<tr>
<td><strong>Experiencing physical abuse</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>0.67*</td>
<td>0.89</td>
</tr>
<tr>
<td>- Severity (tau-b)</td>
<td>0.60</td>
<td>0.76</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.65</td>
<td>0.83</td>
</tr>
<tr>
<td><strong>Witnessing intrafamilial physical abuse</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>0.77</td>
<td>0.77</td>
</tr>
<tr>
<td>- Severity (tau-b)</td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.64</td>
<td>0.72</td>
</tr>
<tr>
<td><strong>Neglect (failure to provide)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>0.57</td>
<td>0.82</td>
</tr>
<tr>
<td>- Severity (tau-b)</td>
<td>0.60</td>
<td>0.84</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.55</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Neglect (lack of supervision)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>0.82</td>
<td>1.00</td>
</tr>
<tr>
<td>- Severity (tau-b)</td>
<td>0.82</td>
<td>0.97</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.75</td>
<td>0.99</td>
</tr>
<tr>
<td><strong>Separations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>- Number (tau-b)</td>
<td>0.91</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Note: All variables significant at $p < 0.001$, unless otherwise stated. * = $p < 0.01$
12.5 Reliability of scoring the case material

Although five researchers were involved in the extraction of the case material, only four were involved in the scoring. In this stage of the reliability procedure 20 cases with both a trawled clinical and social service file were selected at random. Each one of the four researchers scored each of the 20 files.

Reliability estimates were calculated separately for each possible pairing of researchers (i.e. researcher A and B, A and C, A and D, B and C, B and D, C and D). The same statistical analyses and reliability criteria were used in the scoring reliability as in the extraction reliability. The results of this stage of reliability are summarised in table 12.2.

Of the 126 reliability estimates, 113 (90%) reached acceptable levels of reliability (kappa > 0.5; Kendall’s tau-b > 0.7). In this second stage of reliability the difficulties with reliability were less clearly confined to the more than dichotomous measures. 93% (50 out of 54) of the dichotomous measures reached acceptable levels of reliability, with 52% (28 out of 54) having kappa values above 0.8, and 76% (41 out of 54) having values of above 0.7. Of the 72 more than dichotomous measures, 63 (88%) reached acceptable levels of reliability.
Table 12.2 – Scoring reliability for each pairing of four researchers

<table>
<thead>
<tr>
<th>Variable</th>
<th>Researcher pairing</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A-B</td>
<td>A-C</td>
<td>A-D</td>
<td>B-C</td>
<td>B-D</td>
<td>C-D</td>
</tr>
<tr>
<td>Sexual perpetration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>0.89</td>
<td>0.88</td>
<td>1.00</td>
<td>0.76</td>
<td>0.89</td>
<td>0.88</td>
</tr>
<tr>
<td>Child sexual abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>0.44*</td>
<td>0.77</td>
<td>0.44*</td>
<td>0.77</td>
<td>1.00</td>
<td>0.77</td>
</tr>
<tr>
<td>- Severity (tau-b)</td>
<td>0.87</td>
<td>0.91</td>
<td>0.78</td>
<td>0.96</td>
<td>0.89</td>
<td>0.86</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.71</td>
<td>0.90</td>
<td>0.75</td>
<td>0.81</td>
<td>0.95</td>
<td>0.85</td>
</tr>
<tr>
<td>- Abused by a female (k)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Experiencing physical abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>1.00</td>
<td>0.78</td>
<td>0.52</td>
<td>0.78</td>
<td>0.52</td>
<td>0.37</td>
</tr>
<tr>
<td>- Severity (tau-b)</td>
<td>0.81</td>
<td>0.81</td>
<td>0.74</td>
<td>0.75</td>
<td>0.69</td>
<td>0.58</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.80</td>
<td>0.81</td>
<td>0.78</td>
<td>0.98</td>
<td>0.79</td>
<td>0.80</td>
</tr>
<tr>
<td>Witnessing intrafamilial physical abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>0.69</td>
<td>0.76</td>
<td>0.86</td>
<td>0.48</td>
<td>0.58</td>
<td>0.88</td>
</tr>
<tr>
<td>- Severity (tau-b)</td>
<td>0.66</td>
<td>0.72</td>
<td>0.68</td>
<td>0.59</td>
<td>0.75</td>
<td>0.73</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.68</td>
<td>0.67</td>
<td>0.66</td>
<td>0.58</td>
<td>0.56</td>
<td>0.79</td>
</tr>
<tr>
<td>Neglect (failure to provide)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>0.90</td>
<td>0.80</td>
<td>0.70</td>
<td>0.90</td>
<td>0.80</td>
<td>0.90</td>
</tr>
<tr>
<td>- Severity (tau-b)</td>
<td>0.79</td>
<td>0.77</td>
<td>0.69</td>
<td>0.85</td>
<td>0.75</td>
<td>0.87</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.80</td>
<td>0.78</td>
<td>0.67</td>
<td>0.84</td>
<td>0.72</td>
<td>0.87</td>
</tr>
<tr>
<td>Neglect (lack of supervision)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>0.90</td>
<td>0.79</td>
<td>0.90</td>
<td>0.89</td>
<td>1.00</td>
<td>0.89</td>
</tr>
<tr>
<td>- Severity (tau-b)</td>
<td>0.89</td>
<td>0.79</td>
<td>0.84</td>
<td>0.82</td>
<td>0.92</td>
<td>0.85</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.94</td>
<td>0.83</td>
<td>0.85</td>
<td>0.87</td>
<td>0.89</td>
<td>0.78</td>
</tr>
<tr>
<td>Rejection by carers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>0.79</td>
<td>0.59</td>
<td>0.78</td>
</tr>
<tr>
<td>- Duration (tau-b)</td>
<td>0.75</td>
<td>0.66</td>
<td>0.84</td>
<td>0.81</td>
<td>0.68</td>
<td>0.74</td>
</tr>
<tr>
<td>Separations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dichotomous (k)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>- Number (tau-b)</td>
<td>0.90</td>
<td>0.94</td>
<td>0.84</td>
<td>0.93</td>
<td>0.91</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Note: All variables significant at p < 0.001, unless otherwise stated. * = p < 0.05
12.6 Summary

One of the challenges of using the catch-up longitudinal design is the difficulty of ensuring reliability when the data were originally gathered for reasons other than the research study. A number of features of the case material used in this study could endanger reliability, these included the often voluminous case files, the absence of information that would help clarify a particular scoring, and information not being collected in one place nor recorded in an easily identifiable format.

As expected, reliability was something of a problem, with between 10 and 22% of the reliability estimates falling below the predetermined levels of acceptability (Kappa > 0.5; Kendall’s tau-b > 0.7). However, there is some evidence that the dichotomous measures were more reliable than the other measures. Summarising both sets of analyses, 94% (64 out of 68) of the dichotomous measures met acceptable reliability levels, compared with 82% (77 out of 94) of the other variables. This has implications for the comparison of the relative predictive capacity different risk indexes. The increase in random error due to reliability problems will reduce the predictive capacity of a risk index using more than dichotomous levels of measurement.
13 Results

13.1 Introduction

The results in this section, unless stated otherwise, are based on the 104 cases used in the longitudinal analysis. 21 of the 104 met criteria for sexual perpetration (victim-perpetrators); the remaining 83 were classified as victim-only.

13.2 Description of social service and clinical files

A series of calculations examined the comparability of the amount of case material held on the victim-only and victim-perpetrator groups. If the amount of information affected a person’s risk index score, then any systematic differences between the two groups may account for any differences in risk factor scores.

To make the meaning of these calculations clear it is helpful to summarise the types of information that could be used in scoring the risk factors. For the victim-only group, risk factor scores were based on information that was recorded either prior to the age of 16 or the end-point of the case material, whichever took place first. This includes information about events that occurred before the start of a file, but were recorded retrospectively in that file (e.g. subject’s file began in 1974; in a document dated 1975, the mother reports that the boyfriend living with her and the subject in 1972 used to beat the subject). To reflect the inclusion of information referring to events that occurred before the start of the file, the length of time from the subject’s date of birth to the end cut-off point (age 16 or end of case material) was calculated for each subject. Information from both social service and clinical files was used in calculating a risk score, and therefore the end-point of the case material was taken to be either the end-point of the clinical file or the end-point of the social service file, whichever occurred last.
This variable was calculated in the same way for the victim-perpetrator group, with one additional criterion. For this group, the risk scores were based on information that was recorded either prior to the age of 16, the end-point of the case material or the onset of perpetration, whichever took place first. If the subject began perpetrating before the age of 16 or before the end of the case material, the date of perpetration onset was taken as the cut-off point.

Table 13.1 summarises the mean decimal years between the birth of the subject and the end of usable information. As indicated in the table, there is a significant difference between the amounts of information available on the two groups ($U = 343.00$, $z = -4.71$, $p < 0.001$); the victim-perpetrator group have significantly less usable information (mean = 13.0, sd = 2.14) than the victim-only group (mean = 14.97, sd = 2.13). This significant difference is unlikely to act as a confound because the difference works against the hypotheses of the study. The victim-perpetrator group have less time to build up a risk score; therefore if this group still have a significantly higher risk score than the victim-only group then this makes the findings all the more compelling.

<table>
<thead>
<tr>
<th></th>
<th>Victim-only (decimal years)</th>
<th>Victim-perpetrator (decimal years)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>14.97 (sd = 2.13) (range = 6.33 to 16.0)</td>
<td>13.0 (sd = 2.14) (range = 8.50 to 16.0)</td>
<td>$U = 343.0, z = -4.71, p &lt; 0.001$</td>
</tr>
</tbody>
</table>

Although information about events that occurred before the start of the file could be used in coding a subject's risk score, it is possible that the quality of this information differed from information recorded contemporaneously. Therefore differences between the two groups in terms of the length of time the case files were open may also bias the two groups' risk scores. A further variable was calculated to examine this possibility. The same method of calculating the end-point was used, but in this
calculation the start-point was taken to be the date the clinical or social service file began, whichever took place first.

Table 13.2 summarises the mean decimal years between the start and end of usable information for the two groups using this second calculation. As indicated in the table, both groups are comparable on this variable.

Table 13.2 – Mean decimal years between start date and end of usable information

<table>
<thead>
<tr>
<th>Victim-only (decimal years)</th>
<th>Victim-perpetrator (decimal years)</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.94 (sd = 4.75) (range = 0.25 to 16.0)</td>
<td>7.45 (sd = 3.81) (range = 1.92 to 13.83)</td>
<td>t = 0.44, df = 102, p = 0.66</td>
</tr>
</tbody>
</table>

As an additional check on the comparability of case material, it would have been useful to assess the thickness of the case files held on each subject. However, the documents in both the clinical and social service files were often not ordered by date. It would be necessary to order the documents in this way to compare the two groups in terms of ‘usable’ information (i.e. information written before the cut-off point). For example, the file of a victim-perpetrator may be 50cm in length, but 45cm of it may have been written after the onset of perpetration, and therefore only 5cm of the file was ‘usable’. For pragmatic reasons, such as the length of time it would take to re-order the files, these calculations were not made.

13.2.1 Summary

The first stage of the analysis assessed the comparability of the case material held on the two groups. Although the victim-only group had more usable information when the start-date of usable information was assumed to be the date of birth, this difference works against the study hypothesis, because it makes it more difficult for
the victim-perpetrator group to have higher risk scores than the victim-only group. For this reason, the difference was not controlled for in the multivariate analyses.

13.3 Description of perpetration

All of the victim-perpetrators (N = 21) were coded as perpetrators on the basis of information contained in social service or clinical files. 2 also had a caution for a sexual offence against a child.

Over half of the perpetrators (52.4%) had committed penetrative acts against their victims. Of the remainder, perpetration involved genital contact in 19.0% of cases and non-genital contact in 28.6% of cases.

The mean age of the perpetrators at the earliest recorded date of perpetration was 13.10 (sd = 2.26, range 8.50 to 16.83). For 11 of the 21 perpetrators (52.38%) the evidence of abusive behaviour related to a single incident and the duration of the perpetrators' abusive careers needs to be interpreted alongside this fact. When these 11 subjects are included in the calculation, the mean duration is 1.13 years (sd = 2.31, range = 0 to 7.67). The duration of the abusive careers of those who abused on more than one occasion is 2.63 years (sd = 2.99, range = 0.25 to 7.67).

33.3% of the perpetrators abused more than one victim; the maximum number of victims abused by one perpetrator was eight. The gender of the perpetrators' victims was known for 20 out of 21 cases. 35% of the perpetrators had abused males only, 30% had abused females only, and 35% had abused both males and females. The relationship to the victim was also known for 20 out of 21 cases. 40.0% abused a family member (close or extended family), 50.0% abused a non-family member and 10.0% had abused both inside and outside the family.
All of the perpetrators had abused children substantially younger than themselves (more than two years); in addition, three of the perpetrators had abused adults (aged 16 or older). The mean age of the youngest victim abused by a perpetrator was 6.91 yrs (sd = 3.40, range = 1.50 to 14.0) and the mean age of the oldest victim was 9.71 yrs (sd = 5.57, range = 1.50 to 26.00).

### 13.3.1 Summary

The 21 perpetrators were a heterogeneous group. The abusive behaviour of the perpetrators ranged from non-genital contact to penetration, from single instances of abuse to an abusive career lasting over 7 years, and from abusing a single victim to abusing eight victims. The victim characteristics further illustrate the diversity of abusive behaviour. Some perpetrators abused male victims, some abused females, and others abused both. Half of the perpetrators abused outside of the family, 40% abused within the family, and 10% abused both inside and outside the family. Table 13.3 summarises the perpetration of the victim-perpetrator group.

### 13.4 Bivariate analysis of risk factors

This section summarises the results of the bivariate analysis of the risk factors. In the following analyses, duration of abuse (decimal years) is analysed using the U test rather than the t test because of non-normal distributions. The non-normal distributions were caused by a pile-up of scores on 0, and therefore transformation would not make the distributions any more normal.
Table 13.3 – Summary of the perpetration characteristics of the victim-perpetrator group

<table>
<thead>
<tr>
<th>Perpetration characteristic</th>
<th>Proportion / Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity of act</td>
<td></td>
</tr>
<tr>
<td>Penetration</td>
<td>52.4%</td>
</tr>
<tr>
<td>Genital</td>
<td>19.0%</td>
</tr>
<tr>
<td>Non-genital contact</td>
<td>28.6%</td>
</tr>
<tr>
<td>Duration of perpetration</td>
<td>1.13 yrs (sd = 2.31; range 0 to 7.67)</td>
</tr>
<tr>
<td>Age at onset</td>
<td>13.10 yrs (sd = 2.26, range = 8.40 to 16.83)</td>
</tr>
<tr>
<td>Abused more than one victim</td>
<td>33.3%</td>
</tr>
<tr>
<td>Gender of victim</td>
<td></td>
</tr>
<tr>
<td>Males only</td>
<td>35%</td>
</tr>
<tr>
<td>Females only</td>
<td>30%</td>
</tr>
<tr>
<td>Both</td>
<td>35%</td>
</tr>
<tr>
<td>Relationship to victim</td>
<td></td>
</tr>
<tr>
<td>Family member only</td>
<td>40%</td>
</tr>
<tr>
<td>Non-family member only</td>
<td>50%</td>
</tr>
<tr>
<td>Both</td>
<td>10%</td>
</tr>
<tr>
<td>Age of victim</td>
<td></td>
</tr>
<tr>
<td>Mean age of youngest victim</td>
<td>6.91 yrs (sd = 3.40, range = 1.50 to 14.0)</td>
</tr>
<tr>
<td>Mean age of oldest victim</td>
<td>9.71 yrs (sd = 5.57, range = 1.50 to 26.0)</td>
</tr>
</tbody>
</table>

* N = 20; 1 missing case

13.4.1 Child sexual abuse

Child sexual abuse was defined as sexual acts involving physical contact between the subject and perpetrator, in which the perpetrator was either two or more years older than the subject or the perpetrator used physical force. The severity level of sexual victimisation was scored on a five-point scale (1 = Non-genital contact; 2 = Non-penetrative genital contact; 3 = One form of penetration; 4 = More than one
form of penetration; 5 = Penetration plus physical force, prostitution or defecation and urination). The mean severity level did not differ significantly between the two groups (victim-only mean score = 2.81, sd = 1.14) (victim-perpetrator mean score = 3.00, sd = 1.41) (U = 806.00, z = -0.55, p = 0.58). Duration of contact with the perpetrator was calculated on the basis of the length of contact between the subject and the perpetrators. This method was used for all the maltreatment variables discussed in the following section (experiencing physical abuse, witnessing intrafamilial physical abuse, neglect – failure to provide, neglect – lack of supervision, rejection by carers). The mean duration (decimal years) of the sexual victimisation did not differ between the two groups (victim-only = 5.52, sd = 4.55, range 0 to 15.92) (victim-perpetrator = 6.28, sd = 4.72, range = 0 to 16.00) (U = 801.50, z = -0.57, p = 0.57). All of the sample was abused by males, but a proportion was also abused by females, and this differed between the two groups. Twice the proportion of victim-perpetrators (38.1%) reported being sexually abused by a female than the victim-only group (16.9%) (OR = 3.03; 95% CI = 1.06 to 8.68). These findings are summarised in table 13.4.

Table 13.4 – Summary of child sexual abuse characteristics of the victim-only and victim-perpetrator groups

<table>
<thead>
<tr>
<th>Child sexual abuse characteristic</th>
<th>Victim-only</th>
<th>Victim-perpetrator</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity (mean &amp; sd)</td>
<td>2.81 (sd = 1.41)</td>
<td>3.00 (sd = 1.41)</td>
<td>U = 806.00, z = -0.55, p = 0.58</td>
</tr>
<tr>
<td>Duration (mean &amp; sd)</td>
<td>5.52 (sd = 4.55)</td>
<td>6.28 (sd = 4.72)</td>
<td>U = 801.50, z = -0.57, p = 0.57</td>
</tr>
<tr>
<td>Female perpetrator (%)</td>
<td>16.9</td>
<td>38.1</td>
<td>OR = 3.03; 95% CI = 1.06 to 8.68</td>
</tr>
</tbody>
</table>

13.4.2 Physical abuse

Physical abuse was defined as physical behaviour that has the potential to do physical harm to the subject. Similar proportions of both groups had experienced
physical abuse (victim-only = 62.7%; victim-perpetrator = 72.2%) (OR = 1.91; 95% CI = 0.64 to 5.72). The severity of physical abuse was coded on a six-point scale (0 = No physical abuse; 1 = Minor marks below the neck; 2 = Non-minor marks below the neck; 3 = Marks or restraint to neck or above / serious bruising, minor lacerations, minor burns; 4 = Serious injuries but not hospitalised / <24 hrs hospitalisation / asphyxiation). A higher proportion of the victim-perpetrator group (47.62%) than the victim-only group (30.12%) had experienced serious physical abuse (scoring 3 or above), but the mean difference in severity was not statistically significant (victim-only mean = 1.41, sd = 1.41) (victim-perpetrator mean = 1.95, sd = 1.47) (Mann-Whitney U = 690.00, z = -1.52, p = 0.13). The mean duration of physical abuse did not differ significantly between the two groups (victim-only = 5.21, sd = 5.60, range = 0 to 16.00) (victim-perpetrator = 5.28, sd = 5.88, range = 0 to 16.00) (U = 804.00, z = -0.56, p = 0.58). Table 13.5 summarises the physical abuse characteristics of the victim-only group and the victim-perpetrators.

**Table 13.5 – Summary of physical abuse characteristics of the victim-only and victim-perpetrator groups**

<table>
<thead>
<tr>
<th>Experiencing physical abuse characteristic</th>
<th>Victim-only</th>
<th>Victim-perpetrator</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present (%)</td>
<td>62.7</td>
<td>72.7</td>
<td>OR = 1.91; 95% CI = 0.64 to 5.72</td>
</tr>
<tr>
<td>Severity (mean &amp; sd)</td>
<td>1.41 (sd = 1.41)</td>
<td>1.95 (sd = 1.47)</td>
<td>U = 690.00, z = -1.52, p = 0.13</td>
</tr>
<tr>
<td>Duration (mean &amp; sd)</td>
<td>5.21 (sd = 5.60)</td>
<td>5.28 (sd = 5.88)</td>
<td>U = 804.00, z = -0.56, p = 0.58</td>
</tr>
</tbody>
</table>

**13.4.3 Witnessing intrafamilial physical abuse**

Witnessing intrafamilial physical abuse was defined as the subject being visually or aurally exposed to an event of intrafamilial abuse, in which either the victim or perpetrator was a family member. In cases in which there was evidence of violence in the home, but no clear statement about whether the subject had witnessed the
violence, this was inferred using a number of criteria, which are given in full in the Risk Factor Manual (appendix 5). 81.0% of the perpetrators had witnessed intrafamilial physical abuse, compared to 57.8% of the victim-only group. Although this difference was not statistically significant (OR = 3.01; 95% CI = 0.96 to 10.02), the victim-perpetrators had witnessed more severe abuse than the victim-only group (victim-only mean = 1.17, sd = 1.35) (victim-perpetrator mean = 2.24, sd = 1.61) (Mann-Whitney U = 532.00, z = -2.86, p = 0.004). The scale for measuring severity of witnessing intrafamilial violence was the same as that used to measure physical abuse. The mean duration of witnessing intrafamilial violence did not differ significantly between the two groups (victim-only = 5.24, df = 5.90, range = 0 to 16.00) (victim-perpetrator = 6.10, sd = 4.91, range = 0 to 14.25) (U = 744.50, z = -1.06, p = 0.29). These findings are summarised in table 13.6.

Table 13.6 – Summary of witnessing intrafamilial physical abuse characteristics of the victim-only and victim-perpetrator groups

<table>
<thead>
<tr>
<th>Witnessing intrafamilial physical abuse characteristic</th>
<th>Victim-only</th>
<th>Victim-perpetrator</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present (%)</td>
<td>57.8</td>
<td>81.0</td>
<td>OR = 3.01; 95% CI = 0.96 to 10.02</td>
</tr>
<tr>
<td>Severity (mean &amp; sd)</td>
<td>1.17 (sd = 1.35)</td>
<td>2.24 (sd = 1.61)</td>
<td>U = 532.00, z = -2.86, p = 0.004</td>
</tr>
<tr>
<td>Duration (mean &amp; sd)</td>
<td>5.24 (sd = 5.90)</td>
<td>6.10 (sd = 4.91)</td>
<td>U = 744.50, z = -1.06, p = 0.29</td>
</tr>
</tbody>
</table>

13.4.4 Neglect (failure to provide)

Neglect (failure to provide) was defined as a carer failing to provide a minimum degree of care in meeting the child’s physical needs in one or more of five domains (food, clothing, shelter, health care and hygiene). The victim-perpetrators (71.4%) were significantly more likely than the victim-only group (42.2%) to have experienced neglect (failure to provide) (OR = 3.43; 95% CI = 1.21 to 9.72) and
their experiences were also more likely to be severe (victim-only mean = 0.76, sd = 1.14) (victim-perpetrator mean = 1.29, sd = 1.15) (U = 605.50, z = -2.37, p = 0.02). Severity was scored on a six-point scale (0 = no neglect; 1 = mild to 5 = severe); for each level of severity examples were given from the five domains. For example, level 1 of the food domain was scored if the subject missed occasional meals due to parental neglect; level 5 required a diagnosis of non-organic failure-to-thrive. The duration of this type of maltreatment also differed between the two groups (victim-perpetrator = 7.78, sd = 5.94, range = 0 to 16.00) (victim-only = 4.61, sd = 6.17, range = 0 to 16.00) (U = 625.00, z = -2.15, p = 0.03). Table 13.7 summarises the neglect (failure to provide) characteristics of the victim-only group and the victim-perpetrators.

Table 13.7 – Summary of neglect (failure to provide) characteristics of the victim-only and victim-perpetrator groups

<table>
<thead>
<tr>
<th>Neglect (failure to provide) characteristic</th>
<th>Victim-only</th>
<th>Victim-perpetrator</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present (%)</td>
<td>42.2</td>
<td>71.4</td>
<td>OR = 3.43; 95% CI = 1.21 to 9.72</td>
</tr>
<tr>
<td>Severity (mean &amp; sd)</td>
<td>0.76 (sd = 1.14)</td>
<td>1.29 (sd = 1.15)</td>
<td>U = 605.50, z = -2.37, p = 0.02</td>
</tr>
<tr>
<td>Duration (mean &amp; sd)</td>
<td>4.61 (sd = 6.17)</td>
<td>7.78 (sd = 5.94)</td>
<td>U = 625.00, z = -2.15, p = 0.03</td>
</tr>
</tbody>
</table>

13.4.5 Neglect (Lack of supervision)

Neglect (lack of supervision) is defined as a failure by a carer to take adequate precautions to ensure the subject’s safety in and out of the home, given the child’s developmental needs. Two thirds of the victim-perpetrator group (66.7%) had experienced neglect (lack of supervision) compared to 39.8% of the victim-only group (OR = 3.03, 95% CI = 1.11 to 8.30). The victim-perpetrator group had also experienced more severe neglect (failure to provide) than the victim-only group (victim-only mean = 0.84, sd = 1.29) (victim-perpetrator mean = 1.52, sd = 1.69)
Severity was scored on a six-point scale (0 = no neglect, 1 = mild to 5 = severe). The level at which a subject scored was dependent on the subject’s developmental status at the time of the event, the length of time the subject was left unsupervised and the amount of immediate sources of danger in the environment. The duration of this type of maltreatment did not differ between the two groups (victim-perpetrator = 6.28, sd = 5.26, range = 0 to 14.25) (victim-only = 4.38, sd = 6.11, range = 0 to 16.00) (U = 696.00, z = -1.56, p = 0.12). Table 13.8 summarises the neglect (lack of supervision) characteristics of the victim-only group and the victim-perpetrators.

**Table 13.8 – Summary of neglect (lack of supervision) characteristics of the victim-only and victim-perpetrator groups**

<table>
<thead>
<tr>
<th>Neglect (lack of supervision) characteristic</th>
<th>Victim-only</th>
<th>Victim-perpetrator</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present (%)</td>
<td>39.8</td>
<td>66.7</td>
<td>OR = 3.03, 95% CI = 1.11 to 8.30</td>
</tr>
<tr>
<td>Severity (mean &amp; sd)</td>
<td>0.84 (sd = 1.29)</td>
<td>1.52 (sd = 1.69)</td>
<td>U = 636.00, z = -2.10, p = 0.04</td>
</tr>
<tr>
<td>Duration (mean &amp; sd)</td>
<td>4.38 (sd = 6.11)</td>
<td>6.28 (sd = 5.26)</td>
<td>U = 696.00, z = -1.56, p = 0.12</td>
</tr>
</tbody>
</table>

**13.4.6 Rejection by carers**

Rejection by carers is defined as expressions of dislike or hostility towards the subject by the carer. It include behaviours such as blaming, perceiving the subject as deserving of harsh discipline and punishment, hostility, degradation, rejection, locking up or containing the child. 57.1% of the victim-perpetrators had experienced rejection by one or more carers compared to 42.2% of the victim-only group. This difference between the two groups was not statistically significant (OR = 1.83, 95% CI = 0.69 to 4.81). The duration of this type of maltreatment did not differ significantly between the two groups (victim-perpetrator = 5.73, sd = 5.59, range = 0
to 14.42) (victim-only = 4.44, sd = 5.81, range = 0 to 16.00) (U = 758.50, z = -1.00, p = 0.32). These findings are summarised in table 13.9.

Table 13.9 – Summary of rejection by carer characteristics of the victim-only and victim-perpetrator groups

<table>
<thead>
<tr>
<th>Rejection by carer characteristic</th>
<th>Victim-only</th>
<th>Victim-perpetrator</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present (%)</td>
<td>42.2</td>
<td>57.1</td>
<td>OR = 1.83, 95% CI = 0.69 - 4.81</td>
</tr>
<tr>
<td>Duration (mean &amp; sd)</td>
<td>4.44 (sd = 5.81)</td>
<td>5.73 (sd = 5.59)</td>
<td>U = 758.50, z = -1.00, p = 0.3</td>
</tr>
</tbody>
</table>

13.4.7 Separations from carers

The majority of both groups had experienced one or more separations by the age of 16 or first perpetration (victim-perpetrator = 95.2%) (victim-only = 90.4%) (OR = 2.13; 95% CI = 0.25 to 17.98). A separation was defined as a physical separation lasting one month or more from a carer who had consistently cared for a subject for one year or more. The difference between the two groups in terms of the mean number of separations before the age of 16 or first perpetration was also examined. The victim-perpetrator group had experienced a mean of 3.24 separations (sd = 2.41; range = 0 to 8), and this was similar to the mean of the victim-only group (mean = 3.29, sd = 2.22; range = 0 to 9) (t = 0.09, df = 102, p = 0.93). Table 13.10 summarises the care history characteristics of the two groups.

Table 13.10 – Summary of care history characteristics of the victim-only and victim-perpetrator groups

<table>
<thead>
<tr>
<th>Care history characteristic</th>
<th>Victim-only</th>
<th>Victim-perpetrator</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 or more separations (%)</td>
<td>90.4</td>
<td>95.2</td>
<td>OR = 2.13; 95% CI = 0.25 to 17.98</td>
</tr>
<tr>
<td>Number of separations</td>
<td>3.29 (sd = 2.22)</td>
<td>3.24 (sd = 2.41)</td>
<td>t = 0.09, df = 102, p = 0.93</td>
</tr>
</tbody>
</table>
13.4.8 Summary

At the bivariate level the victim-perpetrators scored higher than the victim-only group on a number of risk factor measures. These included being sexually victimised by a female; the severity of witnessing intrafamilial physical abuse; the dichotomous measure, severity and duration of neglect (failure to provide); and the dichotomous and severity measures of neglect (lack of supervision).

13.5 Multivariate analysis of risk factors

Multivariate analysis of the risk factors consisted of a series of logistic regressions. In these regressions the dichotomous outcome variable was sexual perpetration status (perpetrator / non-perpetrator). Although the two groups did not differ significantly in terms of any of the descriptive variables, age (on 01.05.99) and geographical location at time of referral did approach significance (Age: p = 0.11; Geographical location at time of referral: p = 0.13). To assess whether any difference between the two groups on the risk factor measures were due to these differences, the two variables were entered in the regression. Geographical location is a categorical variable, and was therefore re-coded as a dummy variable. (See appendix 7 for a discussion of multicolinearity.)

In the first regression, the dichotomous measures of the risk factors were examined. Of the seven dichotomous risk factors, three were significant at the bivariate level (sexual victimisation by a female, neglect – failure to provide, neglect – lack of supervision), and witnessing intrafamilial physical abuse fell just short of statistical significance. These four variables were entered into the regression. All control variables and risk factors were entered in one step using forced entry.

Table 13.11 summarises the results of this first regression and table 13.12 provides details of the adjusted odds ratios and 95% confidence intervals for each risk factor.
Table 13.11 – Logistic regression examining four dichotomous risk factors

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.20</td>
<td>0.09</td>
<td>5.18</td>
<td>0.02</td>
</tr>
<tr>
<td>Geographical location</td>
<td>-</td>
<td>-</td>
<td>3.31</td>
<td>0.19</td>
</tr>
<tr>
<td>Geographical location (1)</td>
<td>-0.74</td>
<td>0.76</td>
<td>0.95</td>
<td>0.33</td>
</tr>
<tr>
<td>Geographical location (2)</td>
<td>-1.40</td>
<td>0.78</td>
<td>3.24</td>
<td>0.07</td>
</tr>
<tr>
<td>Sexually abused by a female</td>
<td>0.75</td>
<td>0.62</td>
<td>1.46</td>
<td>0.23</td>
</tr>
<tr>
<td>Witnessing intrafamilial abuse</td>
<td>0.84</td>
<td>0.67</td>
<td>1.57</td>
<td>0.21</td>
</tr>
<tr>
<td>Neglect - failure to provide</td>
<td>1.32</td>
<td>0.61</td>
<td>4.61</td>
<td>0.03</td>
</tr>
<tr>
<td>Neglect - lack of supervision</td>
<td>0.63</td>
<td>0.59</td>
<td>1.14</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Table 13.12 – Adjusted odds ratios and 95% CI from logistic regression examining dichotomous risk factors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Adjusted odds ratio</th>
<th>Adjusted 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.22</td>
<td>1.03 - 1.44</td>
</tr>
<tr>
<td>Geographical location (1)</td>
<td>0.48</td>
<td>1.11 - 2.11</td>
</tr>
<tr>
<td>Geographical location (2)</td>
<td>0.25</td>
<td>0.05 - 1.13</td>
</tr>
<tr>
<td>Sexually abused by a female</td>
<td>2.12</td>
<td>0.63 - 7.19</td>
</tr>
<tr>
<td>Witnessing intrafamilial abuse</td>
<td>2.31</td>
<td>0.62 - 8.60</td>
</tr>
<tr>
<td>Neglect - failure to provide</td>
<td>3.73</td>
<td>1.12 - 12.42</td>
</tr>
<tr>
<td>Neglect - lack of supervision</td>
<td>1.88</td>
<td>0.59 - 5.99</td>
</tr>
</tbody>
</table>

As table 13.12 indicates, of the four dichotomous risk factors only neglect (failure to provide) remained a significant predictor of sexual perpetration status in the logistic regression (Adjusted OR = 3.73; 95% CI 1.12 to 12.42).

The second logistic regression examined the risk factor measures that used a more than dichotomous level of measurement (i.e. severity and duration). Four of these measures were statistically significant (p < 0.05): severity of witnessing intrafamilial physical abuse, severity of neglect (failure to provide), duration of neglect (failure to provide), and severity of neglect (lack of supervision). Two further variables approached statistical significance: severity of experiencing physical abuse (p = 0.13) and duration of neglect (lack of supervision) (p = 0.12).
The dichotomous measure of sexual abuse by a female was also entered into the regression. This is because it is important to establish whether the severity and duration measures still predict sexual perpetration when this variable is controlled for. The other risk factors that were significant at the dichotomous level also had significant severity levels, and these severity levels encompassed the dichotomous measure. It was therefore unnecessary to enter these into the regression. The two control variables (age and geographical location at referral) were also entered in the regression.

Table 13.13 summarises the results of this second regression and table 13.14 gives details of the adjusted odds ratios and 95% confidence intervals for each risk factor.

**Table 13.13 – Logistic regression examining severity and duration measures of the risk factors**

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.18</td>
<td>0.09</td>
<td>3.96</td>
<td>0.05</td>
</tr>
<tr>
<td>Geographical location</td>
<td>-</td>
<td>-</td>
<td>3.38</td>
<td>0.18</td>
</tr>
<tr>
<td>Geographical location (1)</td>
<td>-1.01</td>
<td>0.78</td>
<td>1.68</td>
<td>0.20</td>
</tr>
<tr>
<td>Geographical location (2)</td>
<td>-1.48</td>
<td>0.81</td>
<td>3.36</td>
<td>0.07</td>
</tr>
<tr>
<td>Sexually abused by a female</td>
<td>0.93</td>
<td>0.67</td>
<td>1.94</td>
<td>0.16</td>
</tr>
<tr>
<td>Experiencing physical abuse (severity)</td>
<td>0.05</td>
<td>0.24</td>
<td>0.04</td>
<td>0.85</td>
</tr>
<tr>
<td>Witnessing intrafamilial abuse (severity)</td>
<td>0.43</td>
<td>0.21</td>
<td>4.20</td>
<td>0.04</td>
</tr>
<tr>
<td>Neglect - failure to provide (severity)</td>
<td>-0.08</td>
<td>0.33</td>
<td>0.06</td>
<td>0.81</td>
</tr>
<tr>
<td>Neglect - failure to provide (duration)</td>
<td>0.10</td>
<td>0.07</td>
<td>2.12</td>
<td>0.15</td>
</tr>
<tr>
<td>Neglect - lack of supervision (severity)</td>
<td>0.17</td>
<td>0.28</td>
<td>0.36</td>
<td>0.55</td>
</tr>
<tr>
<td>Neglect - lack of supervision (duration)</td>
<td>-0.03</td>
<td>0.07</td>
<td>0.21</td>
<td>0.65</td>
</tr>
</tbody>
</table>
Table 13.14 – Adjusted odds ratios and 95% CI from logistic regression examining severity and duration measures of risk factors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Adjusted odds ratio</th>
<th>Adjusted 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.20</td>
<td>1.00 - 1.42</td>
</tr>
<tr>
<td>Geographical location (1)</td>
<td>0.36</td>
<td>0.08 - 1.68</td>
</tr>
<tr>
<td>Geographical location (2)</td>
<td>0.23</td>
<td>0.05 - 1.11</td>
</tr>
<tr>
<td>Sexually abused by a female</td>
<td>2.54</td>
<td>0.68 - 9.43</td>
</tr>
<tr>
<td>Experiencing physical abuse (severity)</td>
<td>1.05</td>
<td>0.66 - 1.66</td>
</tr>
<tr>
<td>Witnessing intrafamilial abuse (severity)</td>
<td>1.53</td>
<td>1.02 - 2.30</td>
</tr>
<tr>
<td>Neglect - failure to provide (severity)</td>
<td>0.92</td>
<td>0.48 - 1.77</td>
</tr>
<tr>
<td>Neglect - failure to provide (duration)</td>
<td>1.10</td>
<td>0.97 - 1.26</td>
</tr>
<tr>
<td>Neglect - lack of supervision (severity)</td>
<td>1.18</td>
<td>0.68 - 2.05</td>
</tr>
<tr>
<td>Neglect - lack of supervision (duration)</td>
<td>0.97</td>
<td>0.84 - 1.12</td>
</tr>
</tbody>
</table>

As table 13.14 indicates, the only risk factor measure to remain a significant predictor in the second regression was the severity of witnessing intrafamilial violence (Adjusted OR = 1.53; 95% CI = 1.02 to 2.30).

13.5.1 Summary

Once potential confounds and other risk factors were controlled for, few of the risk factors remained significant predictors of sexual perpetration. In the first logistic regression, only the dichotomous measure of neglect (failure to provide) remained a significant predictor. However, neither the severity nor duration measures of this factor were significant predictors in the second logistic regression. The opposite conclusion applies to witnessing intrafamilial violence. A dichotomous measure of witnessing intrafamilial violence failed to predict the outcome in the first regression, but a severity measure of this risk factor was the only significant predictor in the second logistic regression.
13.6 Analysis of the seven-factor risk indexes

A further set of analyses examined the predictive capacity of the risk factors.

13.6.1 Analysis of the Seven-Factor Dichotomous Risk Index

The construction of the Seven-Factor Dichotomous Risk Index is described in detail in the analysis section of chapter 11. In brief, on the basis of *a priori* predictions the following seven dichotomous variables were combined into a risk index:

- Sexual victimisation by a female
- Experience of physical abuse
- Witnessing intrafamilial physical abuse
- Neglect (failure to provide)
- Neglect (lack of supervision)
- Rejection by carers
- One or more separations before age 16 / first perpetration / end of file (whichever took place first).

Scores could range from 0 to 7, with higher scores indicating an increased risk of perpetration.

A significant difference was found between the mean risk score for the victim-only and victim-perpetrator groups (t = -3.39, df = 102, p = 0.001); the mean victim-perpetrator score (mean = 4.86, sd = 1.50) was significantly higher than the mean victim-only score (mean = 3.52, sd = 1.65) (Mean difference = -1.34; 95% CI of mean difference = -2.12 to -0.55). The effect size for this difference was 0.85 (difference in means of the two groups / mean standard deviation of the two groups).

Figure 13.1 gives the distribution of the risk index scores for the two groups.
Figure 13.1 – Distribution of Seven-Factor Dichotomous Risk Index scores for victim-only and victim-perpetrator group

The discriminatory capacity of the risk index was assessed by calculating sensitivity, specificity, positive predictive values and negative predictive values. These are summarised in table 13.15.

Table 13.15 – Sensitivity, specificity, positive predictive values and negative predictive value of the Seven-Factor Dichotomous Risk Index

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive predictive value (%)</th>
<th>Negative predictive value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
<td>20.19</td>
<td>Undefined</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>3.61</td>
<td>20.79</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>95.24</td>
<td>10.84</td>
<td>21.28</td>
<td>90.00</td>
</tr>
<tr>
<td>3</td>
<td>95.24</td>
<td>28.92</td>
<td>25.32</td>
<td>96.00</td>
</tr>
<tr>
<td>4</td>
<td>76.19</td>
<td>48.19</td>
<td>27.12</td>
<td>88.89</td>
</tr>
<tr>
<td>5</td>
<td>71.43</td>
<td>69.88</td>
<td>37.50</td>
<td>90.63</td>
</tr>
<tr>
<td>6</td>
<td>42.86</td>
<td>90.36</td>
<td>53.94</td>
<td>86.21</td>
</tr>
<tr>
<td>7</td>
<td>14.29</td>
<td>98.80</td>
<td>75.00</td>
<td>82.00</td>
</tr>
</tbody>
</table>
To further assess the discriminatory capacity of the risk index, a ROC-curve was calculated. This gives a visual description of the discriminatory capacity of the risk index by plotting sensitivity against 1-specificity, which is summarised in the area under the curve statistic. If the index is no better than chance at predicting perpetration then the area under the curve will be 0.5. The greater the area under the curve, the better the discrimination of the index. As figure 13.2 indicates, 0.73 of the area is beneath the curve for this risk index (95% CI = 0.61 to 0.84).

**Figure 13.2 – ROC-curve for Seven-Factor Dichotomous Risk Index**

![ROC-curve for Seven-Factor Dichotomous Risk Index](image)

**13.6.2 Analysis of the Seven-Factor Severity Risk Index**

The Seven-Factor Severity Risk Index consisted of the following variables:

- Sexual victimisation by a female (rated present / absent)
- Experience of physical abuse (six-point severity scale)
- Witnessing intrafamilial physical abuse (six-point severity scale)
- Neglect (failure to provide) (six-point severity scale)
- Neglect (lack of supervision) (six-point severity scale)
- Rejection by carers (present / absent)
- Number of separations before age 16 / first perpetration / end of file (whichever took place first) (number).
The subject’s score on each of the seven variables were transformed to *z* scores and these seven values were then summed. A high score indicated that the boy was at high risk of sexual perpetration; a low score indicated that the boy was at low risk.

A significant difference was found between the mean risk scores for the victim-only and victim-perpetrator groups (*t* = -3.10, *df* = 102, *p* = 0.003); the mean victim-perpetrator score (mean = 2.32, sd = 4.29) was significantly higher than the mean victim-only score (mean = -0.54, sd = 3.65) (Mean difference = -2.86; 95% CI of mean difference = -4.69 to -1.03). The effect size for this difference was 0.72 (difference in means of the two groups / mean standard deviation of the two groups).

The distribution of the risk index scores for the two groups is given in figure 13.3.

**Figure 13.3** – Distribution of Seven-Factor Severity Risk Index scores for victim-only and victim-perpetrator group
The discriminatory capacity of the risk index was again assessed by calculating sensitivity, specificity, positive predictive values, negative predictive values and a ROC-curve. These calculations are summarised in table 13.16 and figure 13.4. The area under the ROC-curve was 0.71 (95% CI = 0.56 to 0.82).

**Table 13.16** – Sensitivity, specificity, positive predictive value and negative predictive value of the Seven-Factor Severity Risk Index

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive predictive value (%)</th>
<th>Negative predictive value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6 (-6.5 to &lt; -5.5)</td>
<td>100</td>
<td>0</td>
<td>20.19</td>
<td>Undefined</td>
</tr>
<tr>
<td>-5 (-5.5 to &lt; -4.5)</td>
<td>100</td>
<td>8.43</td>
<td>21.65</td>
<td>100</td>
</tr>
<tr>
<td>-4 (-4.5 to &lt; -3.5)</td>
<td>95.24</td>
<td>15.66</td>
<td>22.22</td>
<td>92.86</td>
</tr>
<tr>
<td>-3 (-3.5 to &lt; -2.5)</td>
<td>90.45</td>
<td>24.01</td>
<td>23.17</td>
<td>90.91</td>
</tr>
<tr>
<td>-2 (-2.5 to &lt; -1.5)</td>
<td>90.45</td>
<td>35.53</td>
<td>25.33</td>
<td>93.10</td>
</tr>
<tr>
<td>-1 (-1.5 to &lt; -0.5)</td>
<td>71.43</td>
<td>45.78</td>
<td>25.00</td>
<td>86.36</td>
</tr>
<tr>
<td>0 (-0.5 to &lt; 0.5)</td>
<td>71.43</td>
<td>54.22</td>
<td>25.00</td>
<td>86.36</td>
</tr>
<tr>
<td>1 (0.5 to &lt; 1.5)</td>
<td>71.43</td>
<td>61.45</td>
<td>31.91</td>
<td>89.47</td>
</tr>
<tr>
<td>2 (1.5 to &lt; 2.5)</td>
<td>61.90</td>
<td>68.67</td>
<td>33.33</td>
<td>87.69</td>
</tr>
<tr>
<td>3 (2.5 to &lt; 3.5)</td>
<td>47.62</td>
<td>80.72</td>
<td>38.46</td>
<td>85.90</td>
</tr>
<tr>
<td>4 (3.5 to &lt; 4.5)</td>
<td>38.10</td>
<td>87.95</td>
<td>44.44</td>
<td>84.88</td>
</tr>
<tr>
<td>5 (4.5 to &lt; 5.5)</td>
<td>33.33</td>
<td>90.36</td>
<td>46.67</td>
<td>84.27</td>
</tr>
<tr>
<td>6 (5.5 to &lt; 6.5)</td>
<td>28.57</td>
<td>93.98</td>
<td>54.55</td>
<td>83.87</td>
</tr>
<tr>
<td>7 (6.5 to &lt; 7.5)</td>
<td>19.05</td>
<td>97.59</td>
<td>66.67</td>
<td>82.65</td>
</tr>
<tr>
<td>8 (7.5 to &lt; 8.5)</td>
<td>19.05</td>
<td>97.59</td>
<td>66.67</td>
<td>82.65</td>
</tr>
<tr>
<td>9 (8.5 to &lt; 9.5)</td>
<td>4.76</td>
<td>97.59</td>
<td>33.33</td>
<td>80.20</td>
</tr>
</tbody>
</table>
Figure 13.4 – ROC-curve for Seven-Factor Severity Risk Index

13.6.3 Summary

The area under the ROC-curve was significantly greater than 0.5 for the Seven-Factor Dichotomous Risk Index and Seven-Factor Severity Risk Index. This suggested that both indexes are able to predict sexual perpetration status at a level significantly better than change. The area value for the two indexes was similar (Dichotomous $A(z) = 0.73$; Severity $A(z) = 0.71$). This indicates that the introduction of a more than dichotomous level of measurement does not improve discriminatory capacity.

13.7 Analysis of shortened forms of risk indexes

The risk indexes described above were developed on the basis of a priori predictions. Two additional indexes were calculated post hoc to examine whether it was possible to obtain an increased degree of discrimination between the two groups. Because these risk indexes were developed post hoc, there is a clear need to
replicate these findings with independent samples of sexually victimised and sexually abusive males.

One method that may improve discrimination is to generate a risk index based solely on those risk factors that were significant at the bivariate level. Of the seven dichotomous risk factors, three were significant at the bivariate level (sexual victimisation by a female, neglect – failure to provide, and neglect – lack of supervision); one other factor, witnessing intrafamilial physical abuse, fell just short of statistical significance. Of the seven factors in the severity risk index, the same three were statistically significant and witnessing intrafamilial physical abuse was also statistically significant. These four risk factors were used in the construction of a shortened form of the dichotomous risk index (Four-Factor Dichotomous Risk Index) and the shortened form of the severity risk index (Four-Factor Severity Risk Index).

13.7.1 Analysis of the Four-Factor Dichotomous Risk Index

As described above, this Four-Factor Dichotomous Risk Index consisted of the following variables, all measured at the dichotomous level:

- Sexual victimisation by a female
- Witnessing intrafamilial physical abuse
- Neglect (failure to provide)
- Neglect (lack of supervision).

As with the full seven risk factor index, each subject scored 1 if a risk factors was present and these were then summed to give an overall risk score.

The mean score of the victim-perpetrator group (mean = 2.57; sd = 1.03) was significantly higher than the mean score of the victim-only group (mean = 1.57; sd = 1.14) (t = -3.68, df = 102, p < 0.001) (Mean difference = -1.01; 95% CI of mean
difference = -1.55 to -0.46). The effect size for this was difference was 0.92 (difference in means of the two groups / mean standard deviation of the two groups).

Figure 13.5 gives the distribution of the risk index scores for the two groups.

**Figure 13.5** – Distribution of Four-Factor Dichotomous Risk Index scores for victim-only and victim-perpetrator group

![Distribution of Four-Factor Dichotomous Risk Index scores for victim-only and victim-perpetrator group](image)

Table 13.17 and figure 13.6 summarise the measures of discriminatory capacity. The area under the ROC-curve is 0.75 (95% CI = 0.62 to 0.85).

**Table 13.17** – Sensitivity, specificity, positive predictive value and negative predictive value of the Four-Factor Dichotomous Risk Index

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive predictive value (%)</th>
<th>Negative predictive value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
<td>20.19</td>
<td>Undefined</td>
</tr>
<tr>
<td>1</td>
<td>95.24</td>
<td>20.48</td>
<td>23.26</td>
<td>94.44</td>
</tr>
<tr>
<td>2</td>
<td>85.71</td>
<td>49.40</td>
<td>30.00</td>
<td>93.18</td>
</tr>
<tr>
<td>3</td>
<td>61.90</td>
<td>78.31</td>
<td>41.94</td>
<td>89.04</td>
</tr>
<tr>
<td>4</td>
<td>23.53</td>
<td>96.39</td>
<td>57.14</td>
<td>82.47</td>
</tr>
</tbody>
</table>
13.7.2 Analysis of the Four-Factor Severity Risk Index

This risk index consisted of the following variables:

- Sexual victimisation by a female (rated present / absent)
- Witnessing intrafamilial physical abuse (six-point severity scale)
- Neglect (failure to provide) (six-point severity scale)
- Neglect (lack of supervision) (six-point severity scale)

The subject’s score on each of the variables were transformed to z scores and these values were then summed. A significant difference was found between the mean risk score for the victim-only and victim-perpetrator groups ($t = -3.68$, df = 102, $p < 0.001$); the mean victim-perpetrator score (mean = 1.77, sd = 2.92) was significantly higher than the mean victim-only score (mean = -0.43, sd = 2.32) (Mean difference = -2.20; 95% CI of mean difference = -3.39 to -1.02). The effect size for this difference was 0.84 (difference in means of the two groups / mean standard deviation of the two groups).

Figure 13.7 gives the distribution of the risk index scores for the two groups.
Figure 13.7 – Distribution of Four-Factor Severity Risk Index scores for victim-only and victim-perpetrator group

Table 13.18 and figure 13.8 summarise the measures of discriminatory capacity for this index. The area under the ROC-curve is 0.72 (95% CI = 0.59 to 0.83).

Table 13.18 – Sensitivity, specificity, positive predictive value and negative predictive value of the Four-Factor Severity Risk Index

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive predictive value (%)</th>
<th>Negative predictive value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3 (-3.5 to &lt; -2.5)</td>
<td>100</td>
<td>0</td>
<td>20.19</td>
<td>Undefined</td>
</tr>
<tr>
<td>-2 (-2.5 to &lt; -1.5)</td>
<td>95.24</td>
<td>20.48</td>
<td>23.26</td>
<td>94.44</td>
</tr>
<tr>
<td>-1 (-1.5 to &lt; -0.5)</td>
<td>85.71</td>
<td>24.94</td>
<td>25.00</td>
<td>90.63</td>
</tr>
<tr>
<td>0 (-0.5 to &lt; 0.5)</td>
<td>71.43</td>
<td>56.63</td>
<td>29.41</td>
<td>88.68</td>
</tr>
<tr>
<td>1 (0.5 to &lt; 1.5)</td>
<td>71.43</td>
<td>74.70</td>
<td>41.67</td>
<td>91.18</td>
</tr>
<tr>
<td>2 (1.5 to &lt; 2.5)</td>
<td>57.14</td>
<td>80.72</td>
<td>42.86</td>
<td>88.16</td>
</tr>
<tr>
<td>3 (2.5 to &lt; 3.5)</td>
<td>42.86</td>
<td>86.75</td>
<td>45.00</td>
<td>85.71</td>
</tr>
<tr>
<td>4 (3.5 to &lt; 4.5)</td>
<td>23.81</td>
<td>91.57</td>
<td>41.67</td>
<td>82.61</td>
</tr>
<tr>
<td>5 (4.5 to &lt; 5.5)</td>
<td>14.29</td>
<td>93.98</td>
<td>37.50</td>
<td>81.25</td>
</tr>
<tr>
<td>6 (5.5 to &lt; 6.5)</td>
<td>9.52</td>
<td>97.59</td>
<td>50.00</td>
<td>81.00</td>
</tr>
<tr>
<td>7 (6.5 to &lt; 7.5)</td>
<td>4.76</td>
<td>97.59</td>
<td>50.00</td>
<td>80.20</td>
</tr>
<tr>
<td>8 (7.5 to &lt; 8.5)</td>
<td>4.76</td>
<td>100.00</td>
<td>100.00</td>
<td>80.20</td>
</tr>
</tbody>
</table>
13.7.3 Summary

A shortened form of both the dichotomous and severity risk indexes did not improve their discriminatory capacity. For the dichotomous risk index, the shortened form did have a marginally greater area under the curve (0.75 compared to 0.73). The same applies to the severity risk index; here the shortened form had 0.01 more area under the curve (0.72 compared to 0.71).

As a result of these analyses, the Seven-Factor Dichotomous Risk Index was selected for further assessment of its predictive capacity. The shortened forms of the risk indexes were not selected because their contents were based on a post hoc analysis, unlike the contents of the seven-factor indexes, which were selected a priori. In addition, although the shortened forms did have a slightly greater area under the curve than their seven-factor counterparts, this difference was marginal, and would almost certainly be lost in cross-validation (Meehl & Rosen, 1955). The Seven-Factor Dichotomous Risk Index was selected instead of the Seven-Factor
Severity Risk Index, because its area under the ROC-curve is comparable to that of the severity risk index and it has the advantage of being easier to code.

13.8 Further analysis of the Seven-Factor Dichotomous Risk Index

The extent to which a risk index is useful is dependant on the base rate of sexual perpetration, the balance of utilities and a range of treatment factors. To further explore the performance of the Seven-Factor Dichotomous Risk Index, it was examined under each of these conditions. First, with a base rate of 3%, two utility analyses were conducted and the performance of the index was examined for three different treatment efficacies. In the first utility analysis true positives were given twice the weight of the other utilities; in the second, true negatives were given twice the weight of other utilities. The treatment efficacies examined were 25%, 50% and 75%. (If a treatment has an efficacy of 25%, this means that the proportion of victims who will subsequently perpetrate will be reduced by 25%. For example, if 10% of victims will subsequently perpetrate without treatment, then providing treatment with an efficacy of 25% will reduce this proportion to 7.5%). The utility analyses and the treatment efficacy analyses were repeated with a base rate of 10% and again with a base rate of 17%. A full description of the calculations used in these analyses is given in chapter 11.

13.8.1 Analysis assuming a base rate of 3%

In the first set of analyses it was assumed that without treatment 3% of sexually victimised males would subsequently perpetrate. The formula given by McFall and Treat (1999) was used to calculate a decision analysis, in which it was assumed that the benefit of identifying a true positive was twice that of the other utilities (benefit of identifying a true negative, cost of a false positive, cost of a false negative). This analysis provided information on the usefulness of the Seven-Factor Dichotomous Risk Index in circumstances in which it was more important to correctly identify
someone who would subsequently perpetrate than all other considerations. The results of this analysis are given in table 13.19. The highest $U_{\text{Overall}}$ score indicates the cutting point that best achieves this aim. As indicated in table 13.19, the cutting points of 0 and 1 share the highest $U_{\text{Overall}}$ score. The results of this analysis indicate that in a situation in which a base rate of 3% is expected, and it is considered important to correctly identify sexual perpetrators, the risk index is not of use. In these circumstances all sexually victimised males would be identified as potential sexual perpetrators and all would be treated. This of course assumes that it is possible to treat all sexually victimised males, which may not be the case.

Table 13.19 – Decision analysis: True positive given twice the weighting of other utilities (Base rate 3%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>(Pr)(HR)($U_a$)</th>
<th>(Pr)(1-HR)($U_a$)</th>
<th>(1-Pr)(FAR)($U_{\text{FAR}}$)</th>
<th>(1-Pr)(1-FAR)($U_{\text{cFAR}}$)</th>
<th>$U_{\text{Overall}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(0.03)(1.0)(1.0)=0.03</td>
<td>(0.03)(0)(0.5)=0</td>
<td>(0.97)(1.0)(0.5)=0.485</td>
<td>(0.97)(0)(0.5)=0</td>
<td>0.515</td>
</tr>
<tr>
<td>1</td>
<td>(0.03)(1.0)(1.0)=0.03</td>
<td>(0.03)(0)(0.5)=0</td>
<td>(0.97)(0.9639)(0.5)=0.467491</td>
<td>(0.97)(0.0361)(0.5)=0.0175085</td>
<td>0.515</td>
</tr>
<tr>
<td>2</td>
<td>(0.03)(0.9524)(1.0)=0.028572</td>
<td>(0.03)(0.0476)(0.5)=0.000714</td>
<td>(0.97)(0.8916)(0.5)=0.432426</td>
<td>(0.97)(0.1084)(0.5)=0.052574</td>
<td>0.514286</td>
</tr>
<tr>
<td>3</td>
<td>(0.03)(0.9524)(1.0)=0.028572</td>
<td>(0.03)(0.0476)(0.5)=0.000717</td>
<td>(0.97)(0.7108)(0.5)=0.344738</td>
<td>(0.97)(0.2892)(0.5)=0.140262</td>
<td>0.514286</td>
</tr>
<tr>
<td>4</td>
<td>(0.03)(0.7619)(1.0)=0.022857</td>
<td>(0.03)(0.2381)(0.5)=0.001455</td>
<td>(0.97)(0.5181)(0.5)=0.2512785</td>
<td>(0.97)(0.4819)(0.5)=0.2337215</td>
<td>0.5114285</td>
</tr>
<tr>
<td>5</td>
<td>(0.03)(0.7143)(1.0)=0.021429</td>
<td>(0.03)(0.2857)(0.5)=0.002855</td>
<td>(0.97)(0.3012)(0.5)=0.146082</td>
<td>(0.97)(0.6988)(0.5)=0.338918</td>
<td>0.5107145</td>
</tr>
<tr>
<td>6</td>
<td>(0.03)(0.4286)(1.0)=0.012858</td>
<td>(0.03)(0.5714)(0.5)=0.008571</td>
<td>(0.97)(0.0964)(0.5)=0.046754</td>
<td>(0.97)(0.9036)(0.5)=0.438246</td>
<td>0.506447</td>
</tr>
<tr>
<td>7</td>
<td>(0.03)(0.1429)(1.0)=0.004287</td>
<td>(0.03)(0.8571)(0.5)=0.0128565</td>
<td>(0.97)(0.012)(0.5)=0.00582</td>
<td>(0.97)(0.9880)(0.5)=0.47918</td>
<td>0.5021435</td>
</tr>
<tr>
<td>8</td>
<td>(0.03)(0)(1.0)=0</td>
<td>(0.03)(1.0)(0.5)=0.015</td>
<td>(0.97)(0)(0.5)=0</td>
<td>(0.97)(1.0)(0.5)=0.485</td>
<td>0.5</td>
</tr>
</tbody>
</table>

An alternative situation is one in which it is most important to correctly identify sexually victimised males who are not sexual perpetrators. This situation may arise if it is considered damaging to label a sexually victimised male who has not perpetrated as a potential perpetrator. A further decision analysis was conducted to examine the usefulness of the risk index under these circumstances. In this analysis
it was assumed that the benefit of a true positive was twice that of the other utilities; the results are summarised in table 13.20. The highest $U_{\text{Overall}}$ score is achieved for a cutting point of 8. This indicates that, under the circumstances describe above, the risk index would not be of use. Instead the appropriate course of action would be to label all the sexually victimised males as not at risk of perpetration.

Table 13.20 – Decision analysis: True negatives given twice the weighting of the other utilities (Base rate 3%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>$(Pr)(HR)(U_a)$</th>
<th>$(Pr)(1-HR)(U_a)$</th>
<th>$(1-Pr)(FAR)(U_{fa})$</th>
<th>$(1-Pr)(1-FAR)(U_{ca})$</th>
<th>$U_{\text{Overall}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.015</td>
<td>0</td>
<td>0.485</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>1</td>
<td>0.015</td>
<td>0</td>
<td>0.467491</td>
<td>0.035017</td>
<td>0.517508</td>
</tr>
<tr>
<td>2</td>
<td>0.014286</td>
<td>0.000714</td>
<td>0.432426</td>
<td>0.105148</td>
<td>0.552374</td>
</tr>
<tr>
<td>3</td>
<td>0.014286</td>
<td>0.000717</td>
<td>0.344738</td>
<td>0.280524</td>
<td>0.640265</td>
</tr>
<tr>
<td>4</td>
<td>0.0114285</td>
<td>0.0035715</td>
<td>0.2512785</td>
<td>0.467443</td>
<td>0.7337215</td>
</tr>
<tr>
<td>5</td>
<td>0.0107145</td>
<td>0.0042855</td>
<td>0.146082</td>
<td>0.677836</td>
<td>0.838918</td>
</tr>
<tr>
<td>6</td>
<td>0.006429</td>
<td>0.005851</td>
<td>0.046754</td>
<td>0.876492</td>
<td>0.937946</td>
</tr>
<tr>
<td>7</td>
<td>0.0021435</td>
<td>0.0128565</td>
<td>0.00582</td>
<td>0.95836</td>
<td>0.97918</td>
</tr>
<tr>
<td>8</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.985</td>
</tr>
</tbody>
</table>

The first decision analysis suggested that in circumstances in which the identification of a true positive is important, then the risk index should be ignored in favour of providing treatment for all sexually victimised males on the assumption that all were at risk of perpetration. This assumes it is possible to provide treatment for all sexually victimised males, but if it is possible to treat only a proportion then the risk index may yet be of use. The next three analyses examine this possibility for three different treatment efficacies (25%, 50% and 75%).
Table 13.21 summarises the analyses in which a treatment efficacy of 25% is assumed. The ‘proportion treated’ column indicates the proportion of sexually victimised males that would be treated if a certain cutting point were used. The third column indicates the proportion of this population that would perpetrate if the proportion treated were selected at random. The fourth column indicates the proportion of the population that would perpetrate if the proportion were selected using the cutting point on the risk index. The final column indicates the reduction in the proportion perpetrating using the risk index to make the selection compared to selection at random. As the results of this analysis indicate, the risk index makes only minimal reductions in the proportion of sexually victimised males subsequently perpetrating. The maximal reduction provided by the risk index is 0.3% of the total population of sexually victimised males, in situations in which treatment was given to one-third of this population.

Table 13.21 – Treatment efficacy of 25% (Base rate 3%)

<table>
<thead>
<tr>
<th>Cutting Point</th>
<th>Proportion (%) treated (Tr)</th>
<th>Proportion (%) perpetrating if Tr selected at random</th>
<th>Proportion (%) perpetrating if Tr selected using risk index</th>
<th>Risk index improvement over selection at random (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.00</td>
<td>2.25</td>
<td>2.25</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>96.50</td>
<td>2.27</td>
<td>2.25</td>
<td>0.02</td>
</tr>
<tr>
<td>2</td>
<td>89.34</td>
<td>2.33</td>
<td>2.29</td>
<td>0.04</td>
</tr>
<tr>
<td>3</td>
<td>71.80</td>
<td>2.46</td>
<td>2.29</td>
<td>0.18</td>
</tr>
<tr>
<td>4</td>
<td>52.54</td>
<td>2.61</td>
<td>2.43</td>
<td>0.18</td>
</tr>
<tr>
<td>5</td>
<td>31.36</td>
<td>2.76</td>
<td>2.46</td>
<td>0.30</td>
</tr>
<tr>
<td>6</td>
<td>10.64</td>
<td>2.92</td>
<td>2.68</td>
<td>0.24</td>
</tr>
<tr>
<td>7</td>
<td>1.59</td>
<td>2.99</td>
<td>2.89</td>
<td>0.10</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>3.00</td>
<td>3.00</td>
<td>0</td>
</tr>
</tbody>
</table>

These calculations were repeated for a treatment efficacy of 50%. The maximal risk index improvement was 0.6% of the total population of sexually victimised males, in circumstances in which one-third of the population was treated.
### Table 13.22 – Treatment efficacy of 50% (Base rate 3%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Proportion (%) treated (Tr)</th>
<th>Proportion (%) perpetrating if Tr selected at random</th>
<th>Proportion (%) perpetrating if Tr selected using risk index</th>
<th>Risk index improvement over selection at random (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.00</td>
<td>1.50</td>
<td>1.50</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>96.50</td>
<td>1.55</td>
<td>1.50</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>89.34</td>
<td>1.66</td>
<td>1.57</td>
<td>0.09</td>
</tr>
<tr>
<td>3</td>
<td>71.80</td>
<td>1.92</td>
<td>1.57</td>
<td>0.35</td>
</tr>
<tr>
<td>4</td>
<td>52.54</td>
<td>2.21</td>
<td>1.86</td>
<td>0.35</td>
</tr>
<tr>
<td>5</td>
<td>31.36</td>
<td>2.53</td>
<td>1.93</td>
<td>0.60</td>
</tr>
<tr>
<td>6</td>
<td>10.64</td>
<td>2.84</td>
<td>2.36</td>
<td>0.48</td>
</tr>
<tr>
<td>7</td>
<td>1.59</td>
<td>2.98</td>
<td>2.79</td>
<td>0.19</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>3.00</td>
<td>3.00</td>
<td>0</td>
</tr>
</tbody>
</table>

The final analysis assumed a treatment efficacy of 75%. If one-third of the population were treated, the risk index would lead to a reduction of 0.9%, compared to random selection.

### Table 13.23 – Treatment efficacy of 75% (Base rate 3%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Proportion (%) treated (Tr)</th>
<th>Proportion (%) perpetrating if Tr selected at random</th>
<th>Proportion (%) perpetrating if Tr selected using risk index</th>
<th>Risk index improvement over selection at random (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.00</td>
<td>0.75</td>
<td>0.75</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>96.50</td>
<td>0.83</td>
<td>0.75</td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td>89.34</td>
<td>0.99</td>
<td>0.86</td>
<td>0.13</td>
</tr>
<tr>
<td>3</td>
<td>71.80</td>
<td>1.38</td>
<td>0.86</td>
<td>0.53</td>
</tr>
<tr>
<td>4</td>
<td>52.54</td>
<td>1.82</td>
<td>1.29</td>
<td>0.53</td>
</tr>
<tr>
<td>5</td>
<td>31.36</td>
<td>2.29</td>
<td>1.39</td>
<td>0.90</td>
</tr>
<tr>
<td>6</td>
<td>10.64</td>
<td>2.76</td>
<td>2.04</td>
<td>0.73</td>
</tr>
<tr>
<td>7</td>
<td>1.59</td>
<td>2.96</td>
<td>2.68</td>
<td>0.29</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>3.00</td>
<td>3.00</td>
<td>0</td>
</tr>
</tbody>
</table>
13.8.2 Analysis assuming a base rate of 10%

As discussed in chapter 10, a base rate of 3% is at the lower end of the estimates of the proportion of sexually victimised males who subsequently perpetrate. The available evidence suggests that 10% may be a more realistic figure. The above analyses were repeated this time assuming that 10% of sexually victimised males would subsequently perpetrate if not treated.

Table 13.24 summarises the results of the decision analysis in which the benefit of identifying a true positive is given twice the weighting of the other utilities (benefit of identifying a true negative, cost of a false positive and cost of a false negative). The cutting points of 0 and 1 share the highest $U_{overall}$ score. This indicates that the risk index is not of use in situations in which it is considered important to correctly identify sexual perpetrators. In these circumstances all sexually victimised males would be identified as potential sexual perpetrators and all would be treated.

Table 13.24 – Decision analysis: True positive given twice the weighting of other utilities (Base rate 10%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>(Pr)(HR)(U_{m})</th>
<th>(Pr)(1-HR)(U_{m})</th>
<th>(1-Pr)(FAR)(U_{fa})</th>
<th>(1-Pr)(1-FAR)(U_{ca})</th>
<th>$U_{overall}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(0.1)(1.0)(1.0)=0.10</td>
<td>(0.1)(0)(0.5)=0</td>
<td>(0.9)(1.0)(0.5)=0.45</td>
<td>(0.9)(0)(0.5)=0</td>
<td>0.55</td>
</tr>
<tr>
<td>1</td>
<td>(0.1)(1.0)(1.0)=0.1</td>
<td>(0.1)(0)(0.5)=0</td>
<td>(0.9)(0.9639)(0.5)=0.433755</td>
<td>(0.9)(0.0361)(0.5)=0.16245</td>
<td>0.55</td>
</tr>
<tr>
<td>2</td>
<td>(0.1)(0.9524)(1.0)=0.09524</td>
<td>(0.1)(0.0476)(0.5)=0.00238</td>
<td>(0.9)(0.8916)(0.5)=0.40122</td>
<td>(0.9)(0.1084)(0.5)=0.04878</td>
<td>0.54762</td>
</tr>
<tr>
<td>3</td>
<td>(0.1)(0.9524)(1.0)=0.09524</td>
<td>(0.1)(0.0476)(0.5)=0.00238</td>
<td>(0.9)(0.7108)(0.5)=0.31986</td>
<td>(0.9)(0.2892)(0.5)=0.13014</td>
<td>0.54762</td>
</tr>
<tr>
<td>4</td>
<td>(0.1)(0.7619)(1.0)=0.07619</td>
<td>(0.1)(0.2381)(0.5)=0.011905</td>
<td>(0.9)(0.5181)(0.5)=0.233145</td>
<td>(0.9)(0.4819)(0.5)=0.216855</td>
<td>0.538095</td>
</tr>
<tr>
<td>5</td>
<td>(0.1)(0.7143)(1.0)=0.07143</td>
<td>(0.1)(0.2857)(0.5)=0.014285</td>
<td>(0.9)(0.3012)(0.5)=0.13554</td>
<td>(0.9)(0.6988)(0.5)=0.31446</td>
<td>0.535715</td>
</tr>
<tr>
<td>6</td>
<td>(0.1)(0.4286)(1.0)=0.04286</td>
<td>(0.1)(0.5714)(0.5)=0.02857</td>
<td>(0.9)(0.0964)(0.5)=0.04338</td>
<td>(0.9)(0.9036)(0.5)=0.40662</td>
<td>0.52143</td>
</tr>
<tr>
<td>7</td>
<td>(0.1)(0.1429)(1.0)=0.01429</td>
<td>(0.1)(0.8571)(0.5)=0.04285</td>
<td>(0.9)(0.012)(0.5)=0.0054</td>
<td>(0.9)(0.9880)(0.5)=0.4446</td>
<td>0.507145</td>
</tr>
<tr>
<td>8</td>
<td>(0.1)(0)(1.0)=0</td>
<td>(0.1)(1)(0.5)=0.05</td>
<td>(0.9)(0)(0.5)=0</td>
<td>(0.9)(0)(0.5)=0</td>
<td>0.50</td>
</tr>
</tbody>
</table>
Table 13.25 summarises the results of a second decision analysis, in which it was assumed that the benefit of identifying a true positive was twice that of the other utilities. The highest $U_{\text{Overall}}$ score is achieved for a cutting point of 8. This indicates that under these circumstances the risk index would not be of use. In this situation, the most appropriate course of action would be to identify all the sexually victimised males as not at risk for future sexually abusive behaviour.

Table 13.25 – Decision analysis: True negatives given twice the weighting of other utilities (Base rate 10%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>$(Pr)(HR)(U_a)$</th>
<th>$(Pr)(1-HR)(U_a)$</th>
<th>$(1-Pr)(FAR)(U_{fa})$</th>
<th>$(1-Pr)(1-FAR)(U_{ca})$</th>
<th>$U_{\text{Overall}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$(0.1)(1.0)(0.5)$= $0.05$</td>
<td>$(0.1)(0)(0.5)$= $0$</td>
<td>$(0.9)(1.0)(0.5)$= $0.45$</td>
<td>$(0.9)(0)(1.0)$= $0$</td>
<td>$0.50$</td>
</tr>
<tr>
<td>1</td>
<td>$(0.1)(0.9524)(0.5)$= $0.04762$</td>
<td>$(0.1)(0.0476)(0.5)$= $0.0238$</td>
<td>$(0.9)(0.9639)(0.5)$= $0.43755$</td>
<td>$(0.9)(0.0361)(1.0)$= $0.03249$</td>
<td>$0.516245$</td>
</tr>
<tr>
<td>2</td>
<td>$(0.1)(0.9524)(0.5)$= $0.04762$</td>
<td>$(0.1)(0.0476)(0.5)$= $0.0238$</td>
<td>$(0.9)(0.8916)(0.5)$= $0.40122$</td>
<td>$(0.9)(0.1084)(1.0)$= $0.09756$</td>
<td>$0.54878$</td>
</tr>
<tr>
<td>3</td>
<td>$(0.1)(0.7619)(0.5)$= $0.038095$</td>
<td>$(0.1)(0.2381)(0.5)$= $0.011905$</td>
<td>$(0.9)(0.7108)(0.5)$= $0.31986$</td>
<td>$(0.9)(0.2892)(1.0)$= $0.26028$</td>
<td>$0.63014$</td>
</tr>
<tr>
<td>4</td>
<td>$(0.1)(0.7619)(0.5)$= $0.038095$</td>
<td>$(0.1)(0.2381)(0.5)$= $0.011905$</td>
<td>$(0.9)(0.5181)(0.5)$= $0.233145$</td>
<td>$(0.9)(0.4819)(1.0)$= $0.43371$</td>
<td>$0.716855$</td>
</tr>
<tr>
<td>5</td>
<td>$(0.1)(0.7143)(0.5)$= $0.035715$</td>
<td>$(0.1)(0.2857)(0.5)$= $0.014285$</td>
<td>$(0.9)(0.3012)(0.5)$= $0.13554$</td>
<td>$(0.9)(0.6988)(1.0)$= $0.62892$</td>
<td>$0.81446$</td>
</tr>
<tr>
<td>6</td>
<td>$(0.1)(0.4286)(0.5)$= $0.02413$</td>
<td>$(0.1)(0.5714)(0.5)$= $0.02857$</td>
<td>$(0.9)(0.0964)(0.5)$= $0.04338$</td>
<td>$(0.9)(0.9036)(1.0)$= $0.81324$</td>
<td>$0.90952$</td>
</tr>
<tr>
<td>7</td>
<td>$(0.1)(0.1429)(0.5)$= $0.007145$</td>
<td>$(0.1)(0.8571)(0.5)$= $0.042855$</td>
<td>$(0.9)(0.012)(0.5)$= $0.0054$</td>
<td>$(0.9)(0.9880)(1.0)$= $0.8892$</td>
<td>$0.9446$</td>
</tr>
<tr>
<td>8</td>
<td>$(0.1)(0)(0.5)$= $0$</td>
<td>$(0.1)(1.0)(0.5)$= $0.05$</td>
<td>$(0.9)(0)(0.5)$= $0$</td>
<td>$(0.9)(1.0)(1.0)$= $0.9$</td>
<td>$0.95$</td>
</tr>
</tbody>
</table>

The performance of the risk index assuming a base rate of 10% was further assessed for the three rates of treatment efficacy. These are summarised in tables 13.26 to 13.28. At this base rate the risk index makes a moderate improvement over random selection in situations in which between approximately three-quarters (cutting point 3) and one-eighth (cutting point 6) of a population could be treated. As treatment effectiveness improved so too did the performance of the risk index. If a treatment efficacy of 25% is assumed, the reduction in the proportion perpetrating expressed as a percentage of the total victimised population is above 0.5% for situations in
which three-quarters to one eight of the population is treated. If the treatment has an efficacy of 50%, this figure rises to over 1.0%; and it rises further to over 1.5% if an efficacy of 75% is assumed. The point of maximum effectiveness for the risk index is a situation in which approximately one-third could be treated. In this circumstance the reduction in the proportion perpetrating as a proportion of the total victimised population is 0.9% for a treatment efficacy of 25%, 1.9% for an efficacy of 50% and 2.8% for a treatment efficacy of 75%.

Table 13.26 – Treatment efficacy of 25% (Base rate 10%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Proportion (%) treated (Tr)</th>
<th>Proportion (%) perpetrating if Tr selected at random</th>
<th>Proportion (%) perpetrating if Tr selected using risk index</th>
<th>Risk index improvement over selection at random (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.00</td>
<td>7.50</td>
<td>7.50</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>96.75</td>
<td>7.58</td>
<td>7.50</td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td>89.77</td>
<td>7.76</td>
<td>7.62</td>
<td>0.14</td>
</tr>
<tr>
<td>3</td>
<td>73.50</td>
<td>8.16</td>
<td>7.62</td>
<td>0.54</td>
</tr>
<tr>
<td>4</td>
<td>54.25</td>
<td>8.64</td>
<td>8.10</td>
<td>0.55</td>
</tr>
<tr>
<td>5</td>
<td>34.25</td>
<td>9.14</td>
<td>8.21</td>
<td>0.93</td>
</tr>
<tr>
<td>6</td>
<td>12.96</td>
<td>9.68</td>
<td>8.93</td>
<td>0.75</td>
</tr>
<tr>
<td>7</td>
<td>2.51</td>
<td>9.94</td>
<td>9.64</td>
<td>0.29</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>10.00</td>
<td>10.00</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 13.27 – Treatment efficacy of 50% (Base rate 10%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Proportion (%) treated (Tr)</th>
<th>Proportion (%) perpetrating if Tr selected at random</th>
<th>Proportion (%) perpetrating if Tr selected using risk index</th>
<th>Risk index improvement over selection at random (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.00</td>
<td>5.00</td>
<td>5.00</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>96.75</td>
<td>5.16</td>
<td>5.00</td>
<td>0.16</td>
</tr>
<tr>
<td>2</td>
<td>89.77</td>
<td>5.51</td>
<td>5.24</td>
<td>0.28</td>
</tr>
<tr>
<td>3</td>
<td>73.50</td>
<td>6.30</td>
<td>5.24</td>
<td>1.06</td>
</tr>
<tr>
<td>4</td>
<td>54.25</td>
<td>7.29</td>
<td>6.19</td>
<td>1.10</td>
</tr>
<tr>
<td>5</td>
<td>34.25</td>
<td>8.29</td>
<td>6.43</td>
<td>1.86</td>
</tr>
<tr>
<td>6</td>
<td>12.96</td>
<td>9.35</td>
<td>7.86</td>
<td>1.49</td>
</tr>
<tr>
<td>7</td>
<td>2.51</td>
<td>9.87</td>
<td>9.29</td>
<td>0.59</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>10.00</td>
<td>10.00</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 13.28 – Treatment efficacy of 75% (Base rate 10%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Proportion (%) treated (Tr)</th>
<th>Proportion (%) perpetrating if Tr selected at random</th>
<th>Proportion (%) perpetrating if Tr selected using risk index</th>
<th>Risk index improvement over selection at random (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.00</td>
<td>2.50</td>
<td>2.50</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>96.75</td>
<td>2.74</td>
<td>2.50</td>
<td>0.24</td>
</tr>
<tr>
<td>2</td>
<td>89.77</td>
<td>3.27</td>
<td>2.86</td>
<td>0.41</td>
</tr>
<tr>
<td>3</td>
<td>73.50</td>
<td>4.49</td>
<td>2.86</td>
<td>0.41</td>
</tr>
<tr>
<td>4</td>
<td>54.25</td>
<td>5.93</td>
<td>4.29</td>
<td>1.65</td>
</tr>
<tr>
<td>5</td>
<td>34.25</td>
<td>7.43</td>
<td>4.64</td>
<td>1.65</td>
</tr>
<tr>
<td>6</td>
<td>12.96</td>
<td>9.03</td>
<td>6.79</td>
<td>2.24</td>
</tr>
<tr>
<td>7</td>
<td>2.51</td>
<td>9.81</td>
<td>8.93</td>
<td>0.88</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>10.00</td>
<td>10.00</td>
<td>0</td>
</tr>
</tbody>
</table>

13.8.3 Analysis assuming a base rate of 17%

In the final set of calculations, the analyses were re-run with a high base rate (17%) of sexual perpetration among sexually victimised males. The results of the analysis mirror those already described, but here the usefulness of the risk index is further increased because of the high base rate of sexual perpetration. Table 13.29 and table 13.30 summarises the results of the two decision analyses. In the first (table 13.29), the benefit of identifying a true positive is given twice the weighting of the other utilities (benefit of identifying a true negative, cost of a false positive, cost of a false negative). In a situation in which is was possible to provide treatment to all sexually victimised males the risk index would not be of use if it was considered important to identify true positives. The results of the decision analysis indicate that in these circumstances the available cutting points ($U_{\text{Overall}} = 0.585$ or less) cannot improve on the decision to treat all of the population ($U_{\text{Overall}} = 0.585$). Table 13.20 summarises the results of the decision analysis in which the benefit of identifying a true negative is given twice the weighting of the other utilities. The highest $U_{\text{Overall}}$ score is achieved for a cutting point of 8. In this situation the most appropriate course of action would be to identify all the sexually victimised males as not at risk for future sexually abusive behaviour; therefore the risk index would not be of use.
Table 13.29 – Decision analysis: True positives given twice the weighting of other utilities
(Base rate 17%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>(Pr)(HR)(U\textit{O})</th>
<th>(Pr)(1-HR)(U\textit{O})</th>
<th>(1-Pr)(FAR)(U\textit{FA})</th>
<th>(1-Pr)(1-FAR)(U\textit{cR})</th>
<th>\textit{U}_{\text{overall}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(0.17)(1.0)(0.5)=</td>
<td>0.17</td>
<td>(0.83)(1.0)(0.5)=</td>
<td>(0.83)(0)(0.5)=</td>
<td>0.585</td>
</tr>
<tr>
<td>1</td>
<td>(0.17)(1.0)(0.5)=</td>
<td>0.17</td>
<td>(0.83)(0.9639)(0.5)=</td>
<td>(0.83)(0.0361)(1.0)=</td>
<td>0.585</td>
</tr>
<tr>
<td>2</td>
<td>(0.17)(0.9524)(1.0)=</td>
<td>0.161908</td>
<td>(0.83)(0.8916)(0.5)=</td>
<td>(0.83)(0.1084)(0.5)=</td>
<td>0.580954</td>
</tr>
<tr>
<td>3</td>
<td>(0.17)(0.9534)(1.0)=</td>
<td>0.161908</td>
<td>(0.83)(0.7108)(0.5)=</td>
<td>(0.83)(0.2892)(0.5)=</td>
<td>0.580954</td>
</tr>
<tr>
<td>4</td>
<td>(0.17)(0.7619)(1.0)=</td>
<td>0.129523</td>
<td>(0.83)(0.5181)(0.5)=</td>
<td>(0.83)(0.4819)(1.0)=</td>
<td>0.5647615</td>
</tr>
<tr>
<td>5</td>
<td>(0.17)(0.7143)(1.0)=</td>
<td>0.121431</td>
<td>(0.83)(0.3012)(0.5)=</td>
<td>(0.83)(0.6988)(0.5)=</td>
<td>0.5607155</td>
</tr>
<tr>
<td>6</td>
<td>(0.17)(0.4286)(1.0)=</td>
<td>0.072862</td>
<td>(0.83)(0.0964)(0.5)=</td>
<td>(0.83)(0.9026)(0.5)=</td>
<td>0.536431</td>
</tr>
<tr>
<td>7</td>
<td>(0.17)(0.1742)(1.0)=</td>
<td>0.024293</td>
<td>(0.83)(0.012)(0.5)=</td>
<td>(0.83)(0.9880)(1.0)=</td>
<td>0.5121465</td>
</tr>
<tr>
<td>8</td>
<td>(0.17)(0)(1.0)=</td>
<td>0.085</td>
<td>(0.83)(0)(0.5)=</td>
<td>(0.83)(1.0)(0.5)=</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Table 13.30 – Decision analysis: True negatives given twice the weighting of other utilities
(Base rate 17%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>(Pr)(HR)(U\textit{O})</th>
<th>(Pr)(1-HR)(U\textit{O})</th>
<th>(1-Pr)(FAR)(U\textit{FA})</th>
<th>(1-Pr)(1-FAR)(U\textit{cR})</th>
<th>\textit{U}_{\text{overall}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(0.17)(1.0)(0.5)=</td>
<td>0.085</td>
<td>(0.83)(1.0)(0.5)=</td>
<td>(0.83)(0)(1.0)=</td>
<td>0.5</td>
</tr>
<tr>
<td>1</td>
<td>(0.17)(1.0)(0.5)=</td>
<td>0.085</td>
<td>(0.83)(0.9639)(0.5)=</td>
<td>(0.83)(0.0361)(1.0)=</td>
<td>0.5149815</td>
</tr>
<tr>
<td>2</td>
<td>(0.17)(0.9524)(0.5)=</td>
<td>0.080954</td>
<td>(0.83)(0.8916)(0.5)=</td>
<td>(0.83)(0.1084)(1.0)=</td>
<td>0.5449886</td>
</tr>
<tr>
<td>3</td>
<td>(0.17)(0.9534)(0.5)=</td>
<td>0.080139</td>
<td>(0.83)(0.7108)(0.5)=</td>
<td>(0.83)(0.2892)(1.0)=</td>
<td>0.620103</td>
</tr>
<tr>
<td>4</td>
<td>(0.17)(0.7619)(0.5)=</td>
<td>0.0647615</td>
<td>(0.83)(0.5181)(0.5)=</td>
<td>(0.83)(0.4819)(1.0)=</td>
<td>0.6999885</td>
</tr>
<tr>
<td>5</td>
<td>(0.17)(0.7143)(0.5)=</td>
<td>0.0607155</td>
<td>(0.83)(0.3012)(0.5)=</td>
<td>(0.83)(0.6988)(1.0)=</td>
<td>0.790002</td>
</tr>
<tr>
<td>6</td>
<td>(0.17)(0.4286)(0.5)=</td>
<td>0.036431</td>
<td>(0.83)(0.0964)(0.5)=</td>
<td>(0.83)(0.9036)(1.0)=</td>
<td>0.874994</td>
</tr>
<tr>
<td>7</td>
<td>(0.17)(0.1429)(0.5)=</td>
<td>0.0121465</td>
<td>(0.83)(0.012)(0.5)=</td>
<td>(0.83)(0.9880)(1.0)=</td>
<td>0.88263</td>
</tr>
<tr>
<td>8</td>
<td>(0.17)(0)(0.5)=</td>
<td>0.085</td>
<td>(0.83)(0)(0.5)=</td>
<td>(0.83)(1.0)(1.0)=</td>
<td>0.915</td>
</tr>
</tbody>
</table>
Table 13.31 to 13.33 summarises the results of the treatment efficacy analyses assuming a base rate of 17%.

As with the previous analysis, when the proportion to be treated is between one-eighth and three-quarters, it would help to use the risk index to select the sub-sample for treatment. If a treatment efficacy of 25% is assumed, the reduction in the proportion perpetrating expressed as a percentage of the total victimised population is above 0.8% for these situations. If a treatment efficacy of 50% is assumed then this figure rises to above 1.6%, and rises further to over 2.5% for a treatment efficacy of 75%. Once again, the risk index would be of most use when approximately one-third of the population could be treated. In this circumstance the risk index improves on random selection by 1.5% for a treatment efficacy of 25%, 2.9% for an efficacy of 50% and 4.8% for an efficacy of 75%.

Table 13.31 – Treatment efficacy of 25% (Base rate 17%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Proportion (%) treated (Tr)</th>
<th>Proportion (%) perpetrating if Tr selected at random</th>
<th>Proportion (%) perpetrating if Tr selected using risk index</th>
<th>Risk index improvement over selection at random (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.00</td>
<td>12.75</td>
<td>12.75</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>97.00</td>
<td>12.88</td>
<td>12.75</td>
<td>0.14</td>
</tr>
<tr>
<td>2</td>
<td>90.19</td>
<td>13.17</td>
<td>12.95</td>
<td>0.21</td>
</tr>
<tr>
<td>3</td>
<td>75.90</td>
<td>13.77</td>
<td>12.95</td>
<td>0.82</td>
</tr>
<tr>
<td>4</td>
<td>55.95</td>
<td>14.62</td>
<td>13.76</td>
<td>0.86</td>
</tr>
<tr>
<td>5</td>
<td>37.14</td>
<td>15.21</td>
<td>13.96</td>
<td>1.46</td>
</tr>
<tr>
<td>6</td>
<td>15.29</td>
<td>16.35</td>
<td>15.18</td>
<td>1.17</td>
</tr>
<tr>
<td>7</td>
<td>3.43</td>
<td>16.85</td>
<td>16.39</td>
<td>0.46</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>17.00</td>
<td>17.00</td>
<td>0</td>
</tr>
</tbody>
</table>
**Table 13.32** – Treatment efficacy of 50% (Base rate 17%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Proportion (%) treated (Tr)</th>
<th>Proportion (%) perpetrating if Tr selected at random</th>
<th>Proportion (%) perpetrating if Tr selected using risk index</th>
<th>Risk index improvement over selection at random (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.00</td>
<td>8.50</td>
<td>8.50</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>97.00</td>
<td>8.75</td>
<td>8.50</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>90.19</td>
<td>9.33</td>
<td>8.90</td>
<td>0.43</td>
</tr>
<tr>
<td>3</td>
<td>75.90</td>
<td>10.55</td>
<td>8.90</td>
<td>1.64</td>
</tr>
<tr>
<td>4</td>
<td>55.95</td>
<td>12.24</td>
<td>10.52</td>
<td>1.72</td>
</tr>
<tr>
<td>5</td>
<td>37.14</td>
<td>13.84</td>
<td>10.93</td>
<td>2.91</td>
</tr>
<tr>
<td>6</td>
<td>15.29</td>
<td>15.70</td>
<td>13.36</td>
<td>2.34</td>
</tr>
<tr>
<td>7</td>
<td>3.43</td>
<td>16.71</td>
<td>15.79</td>
<td>0.92</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>17.00</td>
<td>17.00</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 13.33** – Treatment efficacy of 75% (Base rate 17%)

<table>
<thead>
<tr>
<th>Cutting point</th>
<th>Proportion (%) treated (Tr)</th>
<th>Proportion (%) perpetrating if Tr selected at random</th>
<th>Proportion (%) perpetrating if Tr selected using risk index</th>
<th>Risk index improvement over selection at random (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.00</td>
<td>4.25</td>
<td>4.25</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>97.00</td>
<td>4.63</td>
<td>4.25</td>
<td>0.38</td>
</tr>
<tr>
<td>2</td>
<td>90.19</td>
<td>5.50</td>
<td>4.86</td>
<td>0.64</td>
</tr>
<tr>
<td>3</td>
<td>75.90</td>
<td>7.36</td>
<td>4.86</td>
<td>2.51</td>
</tr>
<tr>
<td>4</td>
<td>55.95</td>
<td>9.87</td>
<td>7.29</td>
<td>2.58</td>
</tr>
<tr>
<td>5</td>
<td>37.14</td>
<td>12.26</td>
<td>7.90</td>
<td>4.75</td>
</tr>
<tr>
<td>6</td>
<td>15.29</td>
<td>15.05</td>
<td>11.54</td>
<td>3.52</td>
</tr>
<tr>
<td>7</td>
<td>3.43</td>
<td>16.56</td>
<td>15.18</td>
<td>1.39</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>17.00</td>
<td>17.00</td>
<td>0</td>
</tr>
</tbody>
</table>

**13.8.4 Dose-response relationships in treatment**

The risk index may also be of use in situations in which there is a dose-response relationship between the quality or quantity of treatment and the reduction in sexual perpetration. There may be more than one treatment available, and each may have a different balance between cost-effectiveness and treatment efficacy.
For example, an intensive one-to-one treatment may have an efficacy of approximately 50%, but be so expensive that it can be offered to only one-third of sexually victimised males. In contrast a treatment run by local social services may be cheap enough to offer to 75% of sexually victimised males, but only have an efficacy of 25%.

If we assume a base rate of 10%, we can identify the proportion of the total population that would perpetrate given the two treatment approaches. If the selection of either the one-third or the three-quarters to be treated were made at random then the proportion perpetrating would be approximately 8.2% in both cases (see cutting point 3, column 3 of table 13.26, and cutting point 5, column 3 of table 13.27). However, if the risk index was used to identify the two proportions, we could predict that the more expensive treatment of fewer victimised males would be superior. If the risk index were used to select the most at risk three-quarters of the population, a treatment efficacy of 25% would lead to a perpetration rate of 7.6% (see cutting point 3, column 4 of table 13.26). This could be improved upon if we used the risk index to select the most at risk one-third, and provided a treatment with a 50% efficacy, because this strategy would lead to a perpetration rate of 6.4% (see cutting point 5, column 4 of table 13.27).

The extent to which the risk index could be an aid to decision-making when there are two or more available treatments depends on the efficacy of the treatments, the proportion that can be treated and the base rate. Sometimes the improvement would be marginal. However, it would be possible to use the above calculations to make a prediction in advance about whether the risk index is likely to be of use.

13.8.5 Summary

The Seven-Factor Dichotomous Risk Index was selected for further analysis. The discriminative capacity of this index was examined assuming base rates of 3%, 10%
and 17%. For each of these base rates two utility analyses were conducted. The first analysis assumed that true positives were twice as importance as identifying true negatives. The second analysis assumed that true negatives were twice as important as true positives. These analysis suggested that the risk index may not be of use in these circumstances; instead it would be more appropriate to identify all sexually victimised males as potential perpetrators (if true positives are important), or identify all victims as not at risk (if true negatives are important). The analysis in which true positives are considered important makes the assumption that it is possible to treat all sexually victimised males. The next stage of the analysis examined the contribution of the risk index when it is possible to treat only a proportion of this population. These analyses suggested that the risk index may be of use when between one-eighth and one-third of the population could be selected for treatment, particularly when a medium or high base rate is assumed. Another application of the risk index is to decide whether to treat a smaller proportion of the population with a more effective treatment or a larger proportion with a less effective treatment.

13.9 Summary of results

The bivariate analysis of the risk factors indicated that the victim-perpetrators scored significantly higher than the victim-only group on a number of risk factors. Of the dichotomous risk factors these included: being sexually victimised by a female, neglect (failure to provide) and neglect (lack of supervision); witnessing intrafamilial physical abuse fell just short of significance. The victim-perpetrators also scored higher than the victim-only group on a number of variables using a more than dichotomous level of measurement, including: the severity of witnessing intrafamilial physical abuse, severity and duration of neglect (failure to provide) and the severity of neglect (lack of supervision).
A series of logistic regressions were run to control for potential confounds and to establish whether the risk factors independently predicted sexual perpetration when other risk factors were controlled for. In the first logistic regression the dichotomous variables that were statistically significant or close to significant were entered in one step along with age (on 01.05.99) and geographical location. In this regression only neglect (failure to provide) remained a significant predictor of sexual perpetration. In the second regression the more than dichotomous measures of the risk factors were entered along with the two potential confounds and the dichotomous measure of sexual victimisation by a female. In this regression only the severity measure of witnessing intrafamilial physical abuse remained a significant predictor of sexual perpetration.

Four risk indexes were generated to examine whether a combination of risk factors could predict which sexually victimised males were at risk of sexual perpetration. The first risk index (Seven-Factor Dichotomous Risk Index) used the dichotomous measure of seven risk factors selected \textit{a priori}. The second risk index (Seven-Factor Severity Risk Index) used more than dichotomous measures of the same risk factors, if such measures existed. Two further risk indexes (Four-Factor Dichotomous Risk Index and Four-Factor Severity Risk Index) used only those factors that were significant or close to significance at the bivariate level. The discriminative capacities of the four risk indexes were assessed using a variety of techniques, including the calculation of ROC-curves. According to this series of analyses, there were only marginal differences in the performance of the indexes.

The Seven-Factor Dichotomous Risk Index was selected for further analysis because its contents were selected \textit{a priori}, and so its predictive capacity is unlikely to suffer from cross-validation shrinkage. It was selected in favour of the seven-factor severity index because dichotomous measures are substantially easier to code than severity measures.
The further analysis of this risk index examined its performance in a variety of hypothetical situations, including three difference base rates (3%, 10% and 17%), different balances of utilities and different treatment factors. These analyses suggested that the risk index would not be of use if a substantial emphasis were placed on the correct identification of true negatives or the correct identification of true positives. However, this assumes it is possible to provide treatment for all referred sexually victimised males. If this is not possible then the risk index may be of use when between one-eighth and one-third of referred males are treated. The risk index may also be of use when we face a decision to treat more people with a less effective treatment or fewer people with a more effective treatment. However, as discussed in the next chapter, these further analyses of the risk index are at best exploratory.
14 Discussion

14.1 Introduction

Clinical experience suggests that sexually victimised male children are at heightened risk of subsequently developing sexually abusive behaviour. However, by no means every male victim of sexual abuse becomes a perpetrator, which suggests that other risk factors must be involved. This observation led to the two aims of the study. The first aim was to identify those additional risk factors that are causally implicated in the development of sexually abusive behaviour among male victims of sexual abuse. The second aim was much more applied; namely, to establish whether a risk index is able to predict which sexually victimised males are at risk of becoming perpetrators before they begin to abuse.

A catch-up longitudinal design provided a useful methodology to address these two aims. This design provides additional leverage on the causal criteria, while avoiding the ethical difficulties of a longitudinal study conducted in real time. It also provides a stringent test of the predictive capacity of a risk index, because the performance of the index cannot be artificially inflated by retrospective recall bias or the retrospective switching of temporal order.

The sample for the study consisted of sexually victimised males referred to Great Ormond Street Hospital between 1980 and 1992. Each person’s experience of the key risk factors was identified from contemporaneous social service and clinical case material gathered over the course of their childhood and adolescence. Evidence of subsequent sexual perpetration committed by the participants was gathered from social service and clinical case material, and police caution and conviction records.
This chapter will assess whether the thesis has met these two aims. It will first outline the limitations of the study and then discuss the methodological, theoretical, and policy and practice implications of the research.

14.2 Limitations

14.2.1 Introduction

Research in many areas of psychology is inherently problematic. As Meehl (1978) points out psychological researchers routinely face a number of difficulties (e.g. divergent causality, the idiographic problem) that are rarely if ever faced by other scientific disciplines. As chapter 4 argued, the study of sexual perpetration has a set of additional challenges. Limitations are to be expected, and the interpretation of the results of this study must be understood within this context. This section provides a detailed analysis of the study’s limitations, which is used as a basis for the interpretation of the results in the following section.

Limitations of the catch-up design are discussed first; these include the problems of contemporaneous case material, generalisability and researcher bias in coding. Other limitations include the use of police records and the definition of sexual perpetration.

14.2.2 Limitations of the catch-up design

Despite the undoubted strengths of the catch-up longitudinal design, it also suffers from a number of limitations. Many of the limitations that could be levelled at the design have been discussed in chapter 9. The chapter argued that some of the limitations could be overcome, and because these did not pose a problem for the study, they will not be revisited here. Instead, only problems that cannot be addressed or have only been partially addressed by the study will be discussed.
14.2.2.1 Problems of using contemporaneous case material

A central problem for the catch-up design is its reliance on information gathered for reasons other than the research, which leads to substantial heterogeneity in the information available on subjects. For this study, this had implications for the measurement of the risk factors, sexual perpetration and potential confounds.

The reliance on contemporaneous clinical and social service material means that it is only possible to measure those risk factors that were recorded in sufficient detail in a sufficient number of files. This restricted the study to examining maltreatment risk factors. However, a variety of other potential factors may be involved in the development of sexual offending, including other environmental risks and genetic factors. As Rutter and Maughan (1997) point out, conclusions about the causal status of risk factors must always be tentative if the study does not include a genetic component. One potential risk factor that could not be assessed was non-sexual antisocial behaviour. This is a particular limitation given the hypothesis that there may be two routes to sexual perpetration: one in which sexual perpetration is part of a wider pattern of antisocial behaviour and one unrelated to this behaviour (Watkins & Bentovim, 1992).

The study is also limited in that it did not examine whether protective factors moderated the relationship between the risk factors and sexual perpetration. Future research would benefit from examining the role of protective factors, but it is unclear whether contemporaneous case material will provide sufficient detail about these factors.

The nature of the contemporaneous information also limited the validity with which the maltreatment risk factors were measured. The coding of the care history variable provides an example. This variable was generated on the basis of the number of separations lasting a month or more from someone who had cared for the subject for one-year or more. Such quantitative information was routinely recorded in the social
service files, but information on the quality of the care received or the quality of the parent-child relationship was not recorded. It may be important to measure these factors, because they may serve to moderate the relationship between separations and developmental outcomes, including sexual perpetration.

The measurement of sexual perpetration is difficult because often the only people aware of the behaviour are the perpetrator and the victim, and both may be reluctant to reveal this knowledge to others. The use of case material written by professionals who were not specifically looking for this behaviour may further compromise validity. It is possible that a number of the subjects did sexually perpetrate, but this remained unidentified. The extent of this unidentified perpetration is, by definition, unknown; it is also unclear whether identified and unidentified perpetrators differ systematically from each other. One possibility is that unidentified perpetrators were simply better at avoiding discovery, which would mean that the risk factors do not index the risk of sexual perpetration as much as the risk of being caught.

An additional problem with the measurement of sexual perpetration is that even those who were identified as perpetrators may have started abusing earlier than the date of onset recorded in the file. As discussed in chapter 9, this has implications for both the temporal precedence and the non-spuriousness criteria. If a subject began perpetrating in 1990, but the first incident recorded in the files happened in 1992, then maltreatment documented to have occurred in 1991 would be included in the measurement of the risk factors. This maltreatment occurred after the onset of perpetration and so violates the temporal precedence criteria. It is also possible that the sexual perpetrator was more likely to invent or elaborate experiences of risk factors subsequent to the onset of sexual perpetration in an attempt to explain their behaviour. The catch-up design cannot ensure that the temporal precedence criterion nor this feature of the non-spuriousness criterion are met. However, as argued in chapter 9, it does increase the leverage on these criteria. The systematic exclusion of information gathered subsequent to the known onset of perpetration does decrease
the probability of confusing the temporal order of events and the invention of risk factors by perpetrators.

The measurement of potential confounds is also compromised by the use of a catch-up design. As discussed in chapter 4, if statistical controls are to be used to make causal inferences, two criteria must be met. First, all potential confounds must be identified and, second, each of these must be accurately measured. Both of these criteria are particularly problematic for the catch-up design. The reliance on information gathered for reasons other than the research study means that a number of potential confounds could not be measured. It is possible that the victim-only and victim-perpetrator groups differed on factors such as IQ or SES, and these may account for the relationship between the putative risk factors and sexual perpetration. The absence of information about such confounds in the files makes it impossible to assess whether this was the case. The use of information gathered for reasons other than the study also meant that a number of the nuisance variables could not be measured with the degree of accuracy possible in a genuine longitudinal study. For example, neither the social service nor the hospital files routinely recorded sufficient detail of type, quality or duration of the psychological treatment received at either the hospital or another institution. If treatment for sexual victimisation or other psychological problems does decrease the probability of sexual perpetration, then it would be necessary to control for this when we assess the relationship between risk factors and sexual perpetration. These considerations suggest that caution is needed when we interpret the results of the logistic regressions.

14.2.2.2 Generalisability

The problem of generalisability relates to both the use of a Great Ormond Street Hospital sample and the use of subjects who met the criteria imposed by the methodology of the catch-up design (e.g. social service file located, information predated perpetration).
The research study used a sample of male victims of child sexual abuse referred to Great Ormond Street Hospital between 1980 and 1992, which means that each person was known to professionals such as social workers. This is unlikely to compromise generalisability because the risk of sexual perpetration could only ever be assessed for those known to professionals. However, the extent to which the use of a Great Ormond Street Hospital sample limits generalisability is unclear. The hospital was one of the first in the UK to offer specialised assessment and treatment for sexually victimised children, and so its eminence may have led to referrals of more severe and difficult cases. In 1985 the hospital became a tertiary referral centre, which would also lead to the referral of complicated and possibly atypical cases. Without detailed information on the differences between these cases and those seen by other professionals it is difficult to make conclusions about the generalisability of the sample. It would have been helpful to use a sample that was not reliant on referrals to the hospital, such as all cases identified by social services within a particular geographical area and during a particular time frame. Certainly, it would be useful to replicate the results of the current study with this type of sample.

The methodology of the catch-up design imposed a number of additional inclusion-exclusion criteria that may limit generalisability. A potential subject could be included only if a social service file on that subject existed, was located, and permission was given by the relevant social service to view the file. Even if these criteria were met, to be of use the file had to contain information on the risk factors recorded before the age of 16 or the onset of perpetration (whichever occurred first).

The catch-up design introduces further problems for generalisability. In most designs the researcher is able to exert some control over what information is collected on subjects, including subjects who are subsequently excluded from the study, but in a catch-up design this may not be possible. In this study the information on the excluded subjects was often so sketchy that it was impossible to assess the extent to which the excluded subjects differed on key characteristics from the subjects in the final sample. It is possible that there are important differences
between the excluded subjects and the final sample, but the extent and nature of the bias is unknown.

The limits on generalisability are an inevitable consequence of the catch-up design, and reflect the tension between internal and external validity (Campbell, 1957). The catch-up design can provide leverage on a number of the causal criteria outlined in chapter 4, which increases internal validity, but the methodology of the design limits external validity. Campbell (1957) argues that both types of validity are important, but the primary objective must be to establish internal validity. It is only meaningful to ask whether a causal relationship is generalisable to other situations (external validity) if it is certain that there is a causal relationship in the observed situation (internal validity).

14.2.2.3 Researcher bias in coding

In a longitudinal study conducted in real time it is impossible for information on the outcome to bias the coding of information on the predictor variables, because the outcome has not occurred when the predictor variables are coded. This is not the case with the catch-up design, and so there is a possibility of researcher bias. Substantial efforts were made to remove this source of bias, including the division of tasks into three separate stages carried out by different researchers and the double coding of the risk factors under a genuine and a dummy condition. Despite these efforts some possible sources of bias remained. The researcher extracting the information from the case file would be aware of both the perpetration status of the subject and his likely risk scores. It is possible that knowledge of one may affect extraction of material about the other. For example, if a file stated that a subject had sexually abused others, this may have tempted the researcher to extract more quotes about risk factors. Although it would be possible for one researcher to go through a file dividing the material into that which referred to perpetration and that which referred to risk factors, the sheer quantity of information gathered on each subject prevented the use of this strategy. The only other approach would be to employ researchers blind to the study hypotheses, which was also impractical.
14.2.3 Use of police records

In addition to the use of contemporaneous social service and hospital case material, the study also relied on police records of cautions and convictions for sexual offences. A limitation of these data is that they suffer from a funnel effect. Of the sample who have sexually perpetrated, a proportion will be known to others, a proportion of those will be known to professionals, a proportion will be known to the police and a yet smaller proportion will receive a caution or conviction. This limitation is well known, and does not require further elaboration. This section summarises other limitations of the police data used in this study.

The probability of identifying a subject's sexual criminal record if one existed varied between subjects. The search for caution and conviction records was dependent on the accuracy with which a subject's name and date of birth were recorded in the hospital files. If there were inaccuracies in these, an existing record may not be identified. It is also possible that a subject may have changed his name since contact with the hospital, which would also lead to the search missing the record. While this may not be a major problem for most research populations, this sample experienced a substantial number of changes in carers, which is likely to increase the probability of name changes. The police forces conducting the searches insisted on imposing their own criteria of what constituted a 'hit'. These criteria were often stringent because the police did not want to release information about a sexual offence committed by someone not in the sample. However, this meant that a number of potential sexual perpetrators might have been missed.

At the time the study was conducted, caution records were held by local police forces, rather than centrally in a national database. For practical reasons it was not possible to search each subject in each of the local police forces. All subjects were searched in the Metropolitan Police Force database, and were also searched in the areas in which they were thought to be living over the last five years. The areas in which a subject was searched was based on the address of the subject's GP, obtained
from the Office of National Statistics. It is possible that subjects did not re-register with a GP after moving to a new location. If a subject received a caution for a sexual offence in this new area, the search would not have identified it. A number of subjects were not registered with a GP; these subjects were searched on the basis of the addresses recorded in the social service and clinical file. These addresses were current several years before the search was carried out, which may also have led to the search missing caution records.

These difficulties could mean that the search did not identify the criminal records of some subjects who had committed a sexual offence. This may not cause substantial problems for the study, because this misclassification would increase random error, rather than lead to a systematic bias. However, it still remains possible that the risk factors indicate the risk of someone being caught, not the risk of their perpetrating.

14.2.4 Definition of sexual perpetration

Many of the subjects classified as perpetrators had ages of onset before adulthood; therefore the definition of sexual perpetration in childhood and adolescence will substantially influence the results of the study. However, the definition of perpetration in this age group is problematic. First, as discussed in chapter 3, definitions of perpetration must rely on a vague societal consensus definition, rather than one based on developmental psychopathology. Secondly, there is a need to differentiate sexual perpetration from problematic sexual behaviour that deviates from normal sexual development but which does not constitute sexual perpetration. The available approaches to defining problematic sexual behaviours are underdeveloped. Therefore, to define sexual perpetration by young people it is necessary to differentiate vague social consensus definitions of sexual perpetration from underdeveloped, under researched definitions of other types of problematic sexual behaviour. It is difficult to have confidence in the definition of sexual perpetration by young people used in this or any study, but the definition will
substantially influence the results. This should be borne in mind when we interpret the findings of the study.

14.2.5 Limitations of the risk indexes

There are a number of limitations that apply specifically to the risk indexes. The risk factors were all coded from social service and clinical case material, as was the vast majority of the evidence for sexual perpetration. This raises the possibility that the predictive capacity of the indexes owes its success to shared method variance (Campbell & Fiske, 1959), rather than the predictive validity of the risk factors. Shared method variance is the extent to which an association is due not to the constructs of interest, but to the use of the same assessment method for the constructs. Even though another source of information on perpetration was available from police data, the base rate of perpetration in this source was so low that it would be impossible to assess predictive capacity from criminal records alone.

There are other alternative explanations of the success of the risk indexes. If Meehl’s (1978) critique of null hypothesis significance testing is taken seriously, we would expect that a risk index composed of any seven psychologically ‘bad things’, such as different forms of child maltreatment, would predict sexual perpetration, given a large enough sample size and valid measurement. A reasonable question is whether the same level of discrimination could also be achieved if we selected at random seven psychologically ‘bad things’. It may be more appropriate to ask not whether a risk index performs at a level significantly better than chance, but whether the area under the ROC-curve is significantly superior to the area achieved by a ‘random risk index’. This possibility cannot be examined in this study, but it may be helpful to use this more stringent test in future work.

Two issues of generalisability relate specifically to the risk index. The sample was a high-risk one because the majority had experienced severe sexual abuse and had also been exposed to additional risk factors. It is unclear how the index would perform
with a more representative sample of sexually victimised males. In situations in which a larger proportion had experienced no or few additional risk factors then variance in risk index scores would be increased, and this would lead to an increased discriminatory capacity of the risk index. Alternatively, in a sample in which there was a greater variety in the dimensions of sexual abuse (e.g. duration, severity), the measures of these features may become more important predictors of sexual perpetration than the additional risk factors.

The problem of generalisability also applies to the selection of the cut-off points to be used in a particular circumstance. The contents of the seven-factor risk indexes were decided on a priori, and this lessens to some extent the need for replication with an independent sample. However, the performance of a risk index at a given cutting point is based entirely on the results obtained with this particular sample. If a risk index was ever to be used in practice it would first be necessary to use several cross validation samples to identify the reliability of the sensitivity and specificity values obtained by the cutting points.

One important limitation of the current indexes is that the assessment of risk is made from information gathered up to the onset of this behaviour, which would not occur in clinical practice. Instead the risk assessment would be made at the point a sexually victimised male was referred to social services or a clinical team, which may be several years before the person begins to abuse others, if he does so. While this is a suitable end-point for when we examine causal criteria, it is not particularly well suited to testing the predictive capacity of a risk index. Ideally it would be necessary to calculate several risk scores for each subject, based on the risk factors experienced before several ages (e.g. 6 years, 8 years, 10 years etc.). Subjects who had began perpetrating before a particular end-point would be excluded from the analysis. The analysis of the predictive capacity of the index would be performed for each of these age points. This was not possible because such a procedure would have been too labour-intensive given the available research funds.
It is not clear how the end-point used in this study influenced its findings. One possibility is that certain risk factors may be particularly potent; once they occur so too does sexual perpetration. If the risk assessment were made some years prior to this point, then the risk index would not perform as well as it did in this study.

An additional limitation of the risk index is the use of z scores in the calculation of the severity risk indexes. These scores are sample dependent, because the score is expressed in terms of the mean value for the group; therefore they could not be used to generate a score for people other than the 104 subjects in the current sample. Further exploration of risk indexes that give the same weight to measures with different levels of measurement will require other weighting methods.

It would also have been useful to give different weights to the variables. In this study, none of the risk indexes used a weighting procedure to identify optimal weights for the risk factors. This is because it proved impossible to identify an appropriate statistical method to do this given the type of data used in this study. Commonly used statistical methods for identifying optimal weightings, such as discriminant function analysis, assume that the predictor variables are at least interval and normally distributed.

The versions of the risk indexes using dichotomous predictor variables were obviously unsuitable for this type of statistical treatment. It would also have been inappropriate to apply such an approach to the severity risk indexes. The variables were ordinal, but even if these were treated as interval there would still be substantial difficulties. The distribution of the variables was substantially non-normal and this was caused by a high proportion of subjects scoring 0. This meant transformation techniques could not be used to normalise the data, because no matter how the data were transformed, there would still be a pile up of scores on the transformed 0 value. Given these difficulties, optimal weighing methods were not used in the study.
It is also important to restate the caveats first mentioned in chapter 10. The further analysis of the Seven-Factor Dichotomous Risk Index must be considered exploratory. The decision analyses cannot be assumed to provide the appropriate balancing of the utilities; these analyses are more an example of how the risk index could be evaluated than a definitive evaluation. The assessment of treatment factors with treatment effectiveness set at 25%, 50% and 75% is also exploratory because no treatment to prevent the development of sexual perpetration has been developed as yet.

14.2.6 Summary

The findings of the research reported in this thesis can be understood only within the context of the study’s limitations. This section has elaborated these limitations and discussed their implications for interpreting the research. Limitations include the problems of using contemporaneous case material, generalisability, researcher bias in coding, and the definition of sexual perpetration. The following section will summarise the results of the current study and interpret them in the light of these limitations.

14.3 Summary and interpretation of results

The study reported in this thesis compared 21 sexually victimised males who had subsequently sexually perpetrated with 83 sexually victimised males who had not done so.

Descriptive measures were examined to assess whether the two groups differed on any variables that could confound the risk factor comparison. The only difference between the two groups to reach statistical significance was the amount of usable information available on subjects, when the start date was assumed to be the date of birth (Mann Whitney U = 343.00, z = -4.71, p < 0.001). The victim-only group had
approximately two years more usable information (mean = 14.97, sd = 2.13) than the victim-perpetrator group (mean = 13.0, sd = 2.14). This difference works against the study hypothesis, because the victim-perpetrator group had less time to build up a risk score, which made it more difficult for the victim-perpetrator group to have higher risk scores than the victim-only group. For this reason, this difference was not controlled for in the multivariate analyses. Two further variables, while not reaching conventional levels of statistical significance, were considered sufficiently close to significance to be included in the multivariate analysis as potential confounds. The sexual perpetrators were approximately 1 ½ years older than the victim-only group (t = -1.65, df = 24.62, p = 0.11) and the geographical location at the time of referral of the subjects also differed between the two groups ($\chi^2$ with continuity correction = 4.04, df = 2, p = 0.13).

The next stage of the analysis examined the exposure to risk factors in the two groups. At the dichotomous level, three of the seven risk factors discriminated between the two groups: sexual victimisation by a female, neglect (failure to provide) and neglect (lack of supervision). At this level, witnessing physical abuse fell just short of significance. The severity measures (six-point ordinal scale) of witnessing intrafamilial physical abuse, neglect (failure to provide), and neglect (lack of supervision) also differentiated between the two groups, as did the duration measure of neglect (failure to provide). Two additional variables approached statistical significant: severity of experiencing physical abuse (Mann Whitney U = 690.00, z = -1.52, p = 0.13) and duration of neglect (lack of supervision) (Mann Whitney U = 696.00, z = -1.56, p = 0.12).

Two logistic regressions were conducted to further explore the relationship between the risk factors and sexual perpetration. In these regressions the possible confounds of age and geographical location were entered in one step along with the risk factor measures that were significant or close to significance. In the first regression, in which dichotomous measures of the risk factors were used, only the adjusted odds ratio of neglect (failure to provide) remained a significant predictor of sexual...
perpetration (Adjusted OR = 3.73; 95% CI 1.12 to 12.42). The second regression used those severity and duration measures that were significant or close to significance. The dichotomous measure of sexual victimisation by a female was also included to establish whether the severity and duration measures still predicted sexual perpetration when this variable was controlled for. In this second regression the only risk factor measure to remain a significant predictor was the severity of witnessing intrafamilial violence (Adjusted OR = 1.53; 95% CI = 1.02 to 2.30).

The results of both the bivariate and the multivariate analyses must be interpreted within the context of the limitations discussed in the previous section. The catch-up design relies on information gathered for reasons other than the research question, and this may compromise the measurement of the risk factors, sexual perpetration and potential confounds. It may also mean that a number of potentially important confounds cannot be measured at all, because the contemporaneous data do not contain sufficient detail about these constructs. For this reason the results of the logistic regression must be interpreted with particular caution. None of the variables was measured with complete accuracy and a number of potential confounds were not measured at all. As outlined in chapter 4, the causal interpretations of statistical regression techniques must always be extremely tentative, particularly if measurement is compromised or all confounds are not identified. Therefore, it would be misleading to suggest that either neglect (failure to provide) or witnessing intrafamilial physical abuse are non-spurious risk factors, which have a genuine causal role in the development of sexual perpetration. (See appendix 8 for further elaboration of this point.) Other methods for addressing the non-spuriousness criterion will be discussed in section 14.5. However, the measurement of risk factors using information written before the onset of perpetration helps address the temporal precedence criterion; it also helps reject one source of spuriousness, because any association cannot be a result of perpetrators inventing experiences of the risk factors. Despite this, it does remain possible that the actual date of onset of perpetration may have been some time before the date recorded in the contemporaneous files.
The predictive capacity of four risk indexes was assessed to see whether it was possible to predict which sexually victimised males were at risk of sexual perpetration before the onset of this behaviour. Sensitivity, specificity, positive predictive values, negative predictive values and ROC-curves were calculated to examine the discriminatory capacity of the indexes.

Two risk indexes were developed on the basis of a priori predictions: one summated the dichotomous scores from seven risk factors (Seven-Factor Dichotomous Risk Index) and the other used more than dichotomous measures such as severity levels (Seven-Factor Severity Risk Index). Contrary to expectations, the severity risk index did not improve on the predictive capacity of the dichotomous risk index. To further examine the predictive capacity of the risk factors two additional indexes were generated. Shortened forms of the original risk indexes were derived using the four factors that were significant (or close to significance) at the bivariate level. Neither of these improved on the predictive capacity of the original seven-factor risk indexes.

The practical utility of the Seven-Factor Dichotomous Risk Index was further assessed under a number of hypothetical conditions. These included a condition in which the benefits of a true positive were twice that of the other utilities, and an alternative condition in which the benefits of a true negative were twice that of a true positive. The utility of the index was also examined for treatment efficacy of 25%, 50% and 75%. Each of these conditions was also examined for three different base rates (3%, 10% and 17%).

The results of these analyses suggest that the risk index would not be of use when a high value was placed on correctly identifying true negatives. It would be more appropriate to classify all victimised males as not at risk for perpetration rather than use the risk index. The risk index would also not be of use when a high value was placed on the identification of true positives and it was possible to offer treatment to
all identified male victims of sexual abuse. In these circumstances it would be more
appropriate to label all sexually victimised males as at risk and provide treatment to
all. However, in situations in which a high value was placed on identifying true
positives, but not all of the identified population could be treated, then the risk index
could make a valuable contribution, particularly at medium (10%) and high base
rates (17%) of subsequent sexually abusive behaviour.

In situations in which between three-quarters and one eight of the population could
be treated the analyses assuming a base rate of 10% indicated a considerable
advantage for the risk index compared to random selection. This held for treatment
efficacies of 25%, 50% and 75%. The reduction in perpetration achieved by using
the risk index compared to random selection expressed as a percentage of the total
population of sexually victimised males ranged from 0.5% to 2.8%. If a base rate of
17% is assumed, the figures range between 0.8% and 4.8%.

The risk index could also be of use in situations in which two or more competing
treatments exist. One treatment may be less effective, but could be offered to a
larger proportion of the population, because it is cost-effective. An alternative
treatment may be more effective, but the treatment may be offered to fewer people
because of its increased cost. The risk index could be used to decide which of the
two treatment approaches would lead to the greatest reduction in perpetration given
a fixed financial sum. The risk index could then be used to select the proportion for
treatment.

Although the results of the risk index are promising, it is important to bear in mind
the cautions discussed in the previous section. In particular, the risk index was
developed on a high-risk sample and it is unclear whether its performance would
increase or decrease if a more representative sample of sexually victimised males
was used in its evaluation. Also the performance of the risk index at a given cutting
point is based on the results obtained from this particular sample. As with any risk
index, without cross-validation the risk index should not be used. An additional
problem for the risk index is that the risk scores for the sexual perpetrators could include information about events that occurred close to the onset of perpetration. In clinical practice a risk assessment instrument is unlikely to have this advantage, and before the risk index was used it would be necessary to replicate the performance of the risk index with end-points that more closely match how it would be used in practice.

14.3.1 Summary

The catch-up longitudinal design, with its use of information gathered before the onset of sexual perpetration, helps address both the cause and prediction questions examined in this thesis. However, there are limitations with the study. The reliance on information gathered for reasons other than the research objectives may compromise the measurement of the risk factors, potential confounds and sexual perpetration. As discussed in chapter 4, regression techniques, as used in this study, are dependant on the accuracy with which these variables are measured. There are also a number of limitations with the study's attempt at addressing the prediction question. The performance of the risk index was evaluated on a high risk, unrepresentative sample; it is unclear what affect this had on the discriminatory capacity of the index. In addition, the cutting points used in the study have not been cross-validated, which could artificially inflate the performance of the index. The risk index scores could use information about events that occurred close to the onset of perpetration, which does not reflect how an index would be used in clinical practice. Despite these limitations there are a number of implications of the research, and these will be discussed in the following sections.
14.4 Methodological implications

14.4.1 Introduction

The catch-up longitudinal design is rarely used in sexual perpetration research, this study suggests it is a useful methodological strategy. A number of techniques were developed as part of the study to overcome the potential limitations of the catch-up design, and these techniques could be incorporated into future studies. The finding that the dichotomous measurement of maltreatment risk factors provided discrimination equal to the ordinal measurement of severity also has possible methodological implications.

14.4.2 The catch-up longitudinal design

Almost without exception sexual perpetration research has used a retrospective design, which limits causal interpretation. In contrast, the research reported in this thesis uses a catch-up longitudinal study. If research is to establish the causes of sexual perpetration, research designs must be used that can address the causal criteria outlined in chapter 4. In principle the catch-up design provides additional leverage on a number of these criteria and this study has demonstrated that such a design is feasible. One of the main implications of this study is that sexual perpetration research should make more use of the catch-up longitudinal design. However, it is not well placed to address the causal mechanism criterion, and changes to the design or even alternative research strategies will be needed (see section 14.5).

14.4.3 Computerised extraction program and Risk Factor Manual

One of the reasons the catch-up design may not have been more widely used is the practical difficulty for the researcher of managing large amounts of contemporaneous case material. This material is not recorded with a subsequent
research study in mind and it may be necessary, as in this study, to trawl through large amounts of information to identify relevant quotes. A further complication is that all relevant quotes must be extracted before the data are scored. Subsequent information may alter a risk score, but it may be impossible to decide how to alter it if the quote on which it is based is not recorded. For example, a researcher may at one point code a person as having being sexually abused by a mother-figure. Some of the evidence for that decision may have come from comments made by the father that were recorded in a social service file, but the researcher may subsequently read a document that raises doubts about the veracity of the father’s comments. If the researcher is unable to review all the evidence on which the original coding was made, it may be difficult to establish whether the coding of ‘sexual victimisation by a female’ can still be made without relying on the father’s evidence. This type of problem was frequently encountered in the case material used in this study. Even when not faced with this problem, the researcher may find it useful to review all the quotes relevant to a risk factor before making a coding.

The use of a computerised trawling program to facilitate the extraction process made a substantial contribution to the feasibility of a large-scale catch-up design. The computer program recorded each relevant quote along with important information about it, such as the date of the document and the risk factor to which the quote referred. All the quotes relevant to a risk factor could be reviewed simultaneously on one screen. The Risk Factor Manual (see appendix 5) was also crucial to the success of the study. The Manual was designed to give the researcher specific guidelines on how to code information despite the limitations of the data source. For example, in principle the measurement of witnessing intrafamilial violence requires evidence that a subject witnessed this type of event. While references to violence between family members was frequently made in the files, the files often did not explicitly state that the subject had witnessed the abuse. If a clear statement were required to code this variable, then it would be coded as absent for many subjects who had witnessed this type of abuse. The Manual provided a set of criteria that allowed the researcher to infer the presence of this variable in the absence of a clear statement.
that the subject had witnessed an event. These criteria asked the researcher to consider whether the subject was living with the victim and perpetrator at the time of the event, and whether there was evidence that the event had occurred more than once. Without the Manual it may have proved impossible to use contemporaneous case material, and, as argued throughout this thesis, contemporaneous information can be an extremely valuable asset. Future catch-up longitudinal studies using a similar design may benefit from using both the computer program and Manual.

14.4.4 Minimising researcher bias in the catch-up longitudinal design

A longitudinal study conducted in real time automatically minimises researcher bias, because the coding of the risk factor occurs before the outcome, and so knowledge of the outcome cannot bias the risk factor score. This automatic protection does not apply to a catch-up design. Two strategies were used in this study to minimise this source of bias. The task of translating the contemporaneous case material into quantitative risk scores was divided into three stages: extraction of information from the file, scoring the subject’s sexual perpetration status and scoring the subject’s experience of risk factors. For each subject, a different researcher completed each of the three tasks. This meant that the researcher coding the subject’s risk score was unaware of whether the subject had sexually perpetrated, and the researcher determining whether the subject had sexually perpetrated was unaware of the subject’s risk scores. Therefore, knowledge of one could not contaminate scoring of the other.

The date before which quotes had to be written if they were to be used in assigning a risk score differed between the victim-only and victim-perpetrator groups. Typically, quotes written before the age of 16 could be used for the victim-only group, whereas the onset-of-perpetration date, which usually occurred before the age of 16, was used for the sexual perpetrators. The risk factor coder could have used these differences to determine the sexual perpetration status of the subject, and this could have led to researcher bias. To address this problem the risk factors were
double-coded under a genuine and a dummy condition, which masked the perpetration status of the subject.

This study suggests that appropriate methodological strategies can reduce the possibility of researcher bias in a catch-up longitudinal design. Two possible strategies were used with some success in the current research and they may be applicable to other catch-up designs in the area of sexual perpetration.

14.4.5 The measurement of maltreatment

A number of authors have urged maltreatment researchers to abandon the dichotomous measurement of maltreatment for systems that take account of the severity, duration and other features of the experience (Cicchetti & Barnett, 1991; Newcombe & Locke, 2001; Wolfe & McGee, 1994). According to these authors, the limitation of dichotomous measurement is that it ignores variations in outcome that can be attributed to variations in the intensity of the maltreatment experience. This study compared the predictive capacity of a risk index derived from dichotomous measures of maltreatment with an index using measures of severity. Contrary to predictions, the dichotomous risk index performed equally as well as the severity index. This failure may be due to the lack of reliability with which the severity measures were coded. It is also possible that the contemporaneous case material did not contain sufficient information to code severity. However, if this finding were to be replicated, it may support the continued use of dichotomous measurement.

The study used the severity measures developed by Barnett, Manly and Cicchetti (1993). This measurement system has not received a thorough psychometric work-up, but neither has any other maltreatment measurement system. If there is to be a move away from dichotomous measurement, then further validation work on severity measures is needed. Until such a system has been developed, researchers may continue to use dichotomous measures, particularly because it is much easier to
measure maltreatment at the dichotomous level than to use severity and duration measures.

It is also possible that even a well-validated measure of severity may not substantially outperform dichotomous measures. The difference for a person between not experiencing a particular type of maltreatment and experiencing it may be large compared to the variation between levels of severity. Therefore, a dichotomous measure may capture substantial variation in outcome, making it difficult for severity measures to improve predictive capacity.

14.4.6 Summary

There are a number of methodological implications of the research reported in this thesis. Most importantly this study has demonstrated that the catch-up design is a feasible method for sexual perpetration research, and deserves wider use. When a genuine longitudinal study cannot be used, the catch-up design should be considered, because it can provide more leverage on the causal criteria than a retrospective study. It will also prove to be a useful design for research areas that are not as yet sufficiently developed to warrant the time and expense of a genuine longitudinal study. One reason that the catch-up design may not be more widely used is the need to rely on contemporaneous case material gathered for reasons other than the research study. The computerised trawling program and the Risk Factor Manual were developed to facilitate the extraction and scoring of information from this data source, and researchers may benefit from using these in future. In a genuine longitudinal design research bias cannot operate because the outcome is not known when the independent variable is coded. This is not the case with a catch-up design, but a number of innovative methods were developed as part of this research to minimise the possibility of researcher bias. Other researchers using the catch-up design could use similar strategies. Finally the comparable performance of the dichotomous risk index and the severity risk index suggests that calls to abandon dichotomous measurement of maltreatment may be premature.
14.5 Theoretical considerations

14.5.1 Introduction

The catch-up design is not well-placed to test causal mechanisms, and therefore is not well-placed to corroborate or refute theories of sexual perpetration. The aim of this thesis was much more modest. The study was designed to provide additional causal leverage on risk factors for sexual perpetration. The identification of these risk factors should facilitate the move to the next stage of research: the examination of risk mechanisms derived from theories of sexual perpetration. However, to test the predictions of these theories is not enough. As discussed in chapter 4, studies should be capable of corroborating and refuting risky predictions.

It is difficult to identify methodological techniques that can be used to test causal mechanisms, and even more difficult to identify methodologies that can test risky predictions made by theories of sexual perpetration. It is arguable that the initial stage of research conducted by this study is only worth pursuing if feasible methodologies exist to test risky predictions in future stages. In the final part of this section, a number of such methodological techniques will be evaluated, and recommendations will be made for future research.

14.5.2 The importance of studying causal mechanisms

The study reported in this thesis has sought to address three of the causal criteria: association, non-spuriousness and temporal precedence. However, it is usually necessary to move beyond these to establish the causal mechanisms linking a risk factor to sexual perpetration (Rutter, 1995). This is particularly important in areas of research, such as sexual perpetration, that are for ethical reasons non-experimental, because here the crud factor is prevalent and destructive (Meehl, 1990a).
Chapter 4 discusses these arguments in detail, and so only a brief summary is given here. Meehl (1990a) makes two apparently uncontroversial assertions, yet, when taken together, they lead to substantial problems for risk factor research. First, he points out that in psychology everything tends to correlate with everything else. Secondly, in psychology good tends to go along with good just as surely as bad goes along with bad. The implication of these two assertions is that there will (almost) always be a non-zero relationship between negative childhood experiences and psychopathological outcomes such as sexual perpetration, and these will always be detected as statistically significant given a large enough sample size and accurate measurement.

As argued earlier on in this chapter (section 14.3), this difficulty is not easily overcome by statistically controlling for potential confounds. This method relies on the accuracy with which the confounds are measured and the identification of all the relevant confounds, both of which are difficult to achieve in psychological research.

Given the ethical impossibility of conducting experimental studies, the method of choice for addressing this problem is to turn to the issue of risk mechanisms. If we can establish a generative mechanism linking these risk factors to the outcome, then we have substantially more support for a putative cause. The significance of the association is as nothing compared to the demonstration of a causal mechanism. The next stage of research, therefore, is to use theories of sexual perpetration to make predictions about possible risk mechanisms connecting risk factors to sexual perpetration.

If this strategy is to be successful it will be necessary for those predictions to be considerably more risky than those typically made by current theories of sexual perpetration. Before discussing methods of testing theories of sexual perpetration, it is worth recapping the epistemological need for risky predictions.
14.5.3 The importance of risky predictions

The crud factor not only causes problems for risk factor research, it also causes problems for the study of risk mechanisms, and for much the same reason. In non-experimental psychology, theories are typically corroborated by a study establishing a statistically significant relationship in the expected direction between two constructs. Given that everything correlates with everything else, in a non-experimental design two variables will (almost) always have either a positive or negative relationship, and this will be detected given an adequate sample size and adequate measurement. Therefore, even a theory lacking any verisimilitude has a 1 in 2 chance of being corroborated. If the risk mechanisms predicted by theories of sexual perpetration were tested in this way it would do little to move the area on from the study of risk factors.

To test a theory of a risk mechanisms is not enough: the test must be risky – it must expose the theory to the ‘grave danger of refutation’. This, however, is no easy task. In the final part of this section a number of methodological strategies will be discussed that may go some way to meeting these requirements.

14.5.4 Testing causal mechanisms

The catch-up design in its current form is not readily amenable to testing theories of sexual perpetration. Genuine longitudinal and experimental studies are also ruled out because of ethical difficulties. It may seem that the only possible strategy is to use a retrospective design. While this is certainly feasible, the methodological weakness of such an approach would do little to increase the causal status of the risk factors. One example should suffice to illustrate the problems with this design. A theory of sexual perpetration may predict that a sexually victimised male who in addition experiences physical neglect may be at heightened risk of sexual perpetration, because the neglect leads to severe low self-esteem, and this makes it difficult to form successful romantic relationships. To corroborate this theory it would be
necessary to demonstrate the expected temporal relationship between risk factor (physical neglect), mediating mechanism (low self esteem followed by a difficulty to form relationships) and sexual perpetration. Although it would be possible in a retrospective design to establish that each of these variables tends to co-occur, such a finding may do nothing more than prove Thorndike’s truism that ‘bad things tend to go together’ (Meehl, 1990a). The retrospective nature of the study would make it impossible to corroborate the predictions about the temporal relationship between events, but it is the predictions about temporal order that increase the riskiness of the test.

A longitudinal study could overcome these difficulties. It could also provide a riskier test of sexual perpetration theories; this is because one of the key advantages of the longitudinal design is its ability to focus on within-individual change (Farrington, 1988; 1991). Let us assume that a particular theory predicts that witnessing intrafamilial violence increases the risk of sexual perpetration among sexually victimised males through the operation of a particular psychological mechanism. It is possible to squeeze out of such a theory a risky prediction. If the risk factor does lead to the psychological mechanism, then we would expect to see the onset of or a marked increase in the proposed psychological mechanism after the experience of the risk factor. The temporal distance between the two events that would be considered corroborative would need to be determined *a priori* on the basis of the particular theory. If we carried out a longitudinal study in real time we would be able to obtain reasonably accurate timings of the occurrence of the risk factors, the psychological mechanism and the onset of perpetration. We could then take advantage of a within-individual research strategy to test these predictions. For each person experiencing the risk factor we would expect to see the proposed temporal relationship between the risk factor and the psychological mechanism. If other putative risk factors were also measured, and the temporal relationship between these and the proposed psychological mechanism was random, then this would put further money in the bank for the theory under test.
It is also worth noting that a randomisation procedure would further increase the riskiness of the test. If a sample were randomly assigned to experience a risk factor then any change in a hypothesised psychological mechanism would have few alternative explanations. While this would constitute a risky test, it would of course be ethically unacceptable to expose children to risk factors.

We face an apparently intractable methodological problem. For ethical reasons, we cannot use longitudinal or experimental studies; retrospective studies testing theories are unconvincing; and while the catch-up longitudinal design is well placed to address some of the causal criteria it cannot be used to test causal mechanisms. Methodological strategies that may provide solutions to this problem will be discussed in the following section.

14.5.4.1 Longitudinal studies of a single risk factor

Although it would be unethical to conduct a genuine longitudinal study of sexually victimised males exposed to additional risk factors, it may be permissible to conduct longitudinal studies of single risk factors in non-sexually victimised males. For example, a longitudinal study conducted in real time could focus on a sample at high risk for physical abuse. The psychological mechanism hypothesised to link physical abuse to sexual perpetration would be measured periodically. Cases of physical abuse identified through official sources, such as social service records, could then be examined in detail. The question examined would be whether there was an onset or a noticeable increase in the psychological mechanism after the onset of physical abuse. The level of change in this measure for the physically abused sub-sample could be compared to the variation in the psychological mechanism over time in a randomly selected proportion on the non-physically abused sample that had been extensively screened for physical abuse.

There are limitations with this design. The use of a single risk factor design is that the sample can be followed in real time without a substantial minority committing sexually abusive acts. However, this means that a crucial part of the theory cannot
be tested; namely the prediction that the psychological mechanism increases the likelihood of sexual perpetration. The focus on a single risk factor also ignores the possibility of interaction effects. It is possible that a risk factor will lead to a substantial increase in the hypothesised psychological mechanism only when other risk factors are present. For example, physical abuse may increase the hypothesised mechanism only for sexually victimised males. The most substantial problem for this design is its expense, particularly given these undoubted limitations.

14.5.4.2 Existing longitudinal study combined with a catch-up design

A comparatively inexpensive alternative to this design, which could also address some of these limitations, is to combine an existing longitudinal dataset with a catch-up longitudinal design. This study requires the identification of an existing longitudinal dataset that has measured one or more variables predicted to act as the psychological mechanisms connecting risk factors to sexual perpetration. The existing longitudinal study must also have used a high-risk sample of males, such as those known to social services. It would then be possible to identify those males in the sample who, according to social services, had experienced sexual abuse. The social service records of these people could then be used to time the experience of additional risk factors. The researchers would then examine the extent to which the temporal relationship between the occurrence of the risk factors and the psychological mechanisms corroborated the predictions of a theory of sexual perpetration.

The main difficulty for this approach is that an appropriate longitudinal dataset may be difficult to identify – it might not exist. Given the low base rate of sexually victimised males known to social services, the existing longitudinal study would need to have a very large sample size. The dataset must also contain information on the psychological mechanisms of interest.
14.5.4.3 Natural experiment of a single risk factor

An alternative is to identify a natural experiment of a single risk factor. This requires the researcher to identify a population in which the variable of interest has been introduced by nature in a pseudo-random fashion, rather than co-occurring with a number of other variables that may be causally linked to the psychological mechanism or the outcome (Rutter, 1995).

The reason we might want to use a natural experiment is that in a longitudinal design of a single risk factor the presence of other risk factors may complicate the interpretation of the results, particularly if the risk factors tend to occur in the same time period. The occurrence of a risk factor may be followed by an increase in the psychological mechanism, but one or more other risk factors may tend to occur at the same time as the risk factor of interest. The theory predicting a close temporal relationship between the risk factor and the mechanism would not be rigorously tested.

It can be difficult to identify situations that meet the criteria of a natural experiment, but there appears to be at least one suitable situation that could be used to address the psychological mechanisms hypothesised to link sexual victimisation to perpetration. This study would focus on victims of sexual perpetrators who were known to have operated in high-quality private day-care nurseries or schools. It is likely that exposure to additional risk factors is minimised in this population. The design would consist of researchers identifying such a sample and interviewing them to identify current psychological functioning. It would also be necessary to identify a suitable comparison group to test whether the psychological functioning in the sexually abused groups was significantly lower than in the non-abused group. One possible comparison group would be those children who were present at the nursery or school at the same time, but who were not sexually abused by the perpetrator.

A natural experiment does not provide the same degree of control as an experiment in which the researcher directly manipulates the independent variable, and so the
researcher using this approach must consider and seek to exclude alternative explanations of a finding (Cook & Campbell, 1979). One possibility that would need to be considered is that the sexual perpetrator selected the most vulnerable children and these may have shown elevated scores on the hypothesised mechanism regardless of whether they were sexually abused.

14.5.5 Summary

The current study has addressed some of the basic criteria for establishing a causal relationship between risk factors and sexual perpetration. To strengthen the causal status of these risk factors, it is important for research to test theories making predictions about the causal mechanisms connecting these experiences to sexually abusive behaviour. As argued in chapter 4, if research is to provide additional leverage on this criterion then there is a need to make riskier predictions than those made in much psychological research, including previous attempts to corroborate theories of sexual perpetration. However, it may prove difficult to generate and test these risky predictions. The experimental and longitudinal designs could provide situations in which risky tests could be corroborated, but both are ethically unacceptable in this area of research. This section has discussed a number of alternative strategies, including longitudinal studies of a single risk factor, the combination of an existing longitudinal dataset with a catch-up design, and the use of a natural experiment. It is clear from the discussion of these methodologies that they are not without their own problems. Given the need to corroborate risky predictions made by theories of sexual perpetration, further assessment of the feasibility of these designs is needed. The area also needs researchers to show inventiveness in generating methodologies that are ethically acceptable, but also provide risky tests of available theories.
14.6 Policy and practice implications

This thesis suggests that, in principle, the risk index could be of use in certain clinical situations. In particular, it would be of use if base rates of sexual perpetration were reasonably high, the identification of true positives was considered important and it was possible to treat between one-eighth and three-quarters of a population of sexually victimised males. However, this is predicated on the assumption of an effective treatment for those at risk of sexual perpetration. While treatments exist for children (Sirles, Araj, & Bosek, 1997), adolescents (Becker & Kaplan, 1993) and adults (Marshal, Laws, & Barbaree, 1990) who sexually abuse, no treatments have been developed – let alone evaluated – for children who have not sexually abused but who may be at risk of doing so. Some might argue that it is a futile, academic exercise to assess the performance of the index when no treatment exists. However, this section will argue that there are a number of important policy and practice implications of the index as it stands.

The implications discussed in this section all result from one observation: despite the use of a fairly crude risk index derived from a less than ideal data sources, it was possible to generate an index that could successfully discriminate between those who did and did not perpetrate before the onset of this behaviour.

Given that even a crude index could be used to predict which sexually victimised males require treatment for potential perpetration, it seems an apposite time to develop and evaluate treatments that are designed to reduce the risk of perpetration. This research will face the same difficulties as that assessing treatments for sexual recidivism. However, none of this research has taken advantage of the methodological literature on quasi-experimentation (Cook & Campbell, 1979). If studies of treatments for sexually victimised males were to draw upon this body of work, then it might be possible to make some reasonably definitive conclusions.
It may never be possible to compare treatment strategies with known efficacies of 25%, 50% and 75%, but it may be possible to identify more and less effective treatments. Each treatment could then be given a plausible range of treatment efficacies (e.g. least effective range = 10% to 25%, most effective range = 30% to 50%), and it would then be possible to compare the performance of the two indexes under these conditions.

An additional question is whether it is possible to improve the risk index and increase the accuracy of prediction. The performance of the current risk index may be best viewed as a starting point. The aim is to establish the extent to which prediction can be improved beyond that achieved by the index. This may involve adding other risk factors and incorporating into the risk assessment the ameliorating effects of protective factors (Cicchetti & Garmezy, 1993; Garmezy, 1985; Garmezy, Masten, & Tellegen, 1984; Rutter, 1985; 1987). One other alternative would be to incorporate what have been termed dynamic risk factors (Hanson & Harris, 1998). Current work in the prediction of sexual recidivism has argued for the need to incorporate such variables. These variables may be viewed as the psychological mechanisms hypothesised to link the risk factors to sexual perpetration, which once again highlights the need to develop robust methodological strategies to identify these causal mechanisms. Planning at the policy level should make available funds for the evaluation of these alternative risk indexes.

It will not only be necessary to compare the existing risk index with other actuarial methods of prediction; it will also be necessary to establish that the risk index performs equally or better than prediction by clinicians. A powerful source of resistance to the use of such a risk index in clinical practice is the notion that decisions about risk should be made by a clinician on a case-by-case basis using clinical skills to integrate the rich sources of qualitative information. Hidden behind this argument is the assumption that this idiographic approach will improve on the prediction made by the crude number crunching of the actuarial method. However, the question of which approach is superior is an empirical one. It is worth noting that
the omens are not good for claims about the superiority of clinical over actuarial prediction. Grove, Zald, Lebow, Snitz and Nelson (2000) review 129 studies that have compared the performance of the two approaches in a range of areas in psychology and medicine. The review found that the actuarial approach was less accurate than the clinical approach in 6% of studies and superior in approximately 46% of studies. In the remaining 46% of studies, the two methods were equal, but, as the authors note, this too is a 'win' for actuarial methods because they are substantially more cost effective than clinical methods. The available research in the area of sexual offending also questions the predictive ability of clinicians. One study found that psychiatrists and psychologists could not predict sexual and non-sexual recidivism at a statistically significant level in a group of sex offenders (Hall, 1988). Research examining the predictive capacity of clinicians and the risk index would certainly be of use. If clinicians are superior at predicting which males are at risk, then this would be an important finding, with substantial clinical implications. If not, this would add strength to the arguments for employing an actuarial instrument. This would leave the clinician free to do what no actuarial instrument could ever do: provide effective treatment to sexually victimised males to decrease the likelihood that they will repeat the cycle of victimisation.

14.6.1 Summary

Although the available treatments for sexually victimised males have not been evaluated, this does not mean that the risk index has no policy or practice implications. If a treatment were to exist, in certain situations the risk index could be of use in identifying which sexually victimised males were at heightened risk of sexual perpetration. Given this, it seems important that planning at the policy level should support the systematic evaluation of existing treatment programs. The successful performance of this rather crude index suggests that planning at the policy level should encourage the development of superior indexes, incorporating additional risk factor, dynamic risk factors and protective factors. The existing or future risk index should be compared with clinical prediction. Although social
workers and clinicians may prefer case-by-case prediction based on the integration of rich sources of qualitative material, it is by no means certain that the clinical approach to prediction is superior to the actuarial approach.

Which sexually victimised males receive treatment at the moment appears to have more to do with chance than anything else. These chance factors include knowledge at a local level that a treatment exists, availability of local funds to provide treatment and geographical location of an at-risk male and the service provider. The results of this study suggest that this state of affairs is unsatisfactory, because prediction can be made at a level better than chance.

14.7 Concluding remarks

This thesis had two aims. The first aim was to examine the causal status of putative risk factors for sexual perpetration among male victims of sexual abuse. The second aim was to develop a risk index that was able to predict which sexually victimised males were at risk of sexually perpetrating before they did so. Both of these aims have been well served by the catch-up longitudinal design. Genuine longitudinal research is ethically impossible in this area, and for this reason the retrospective design has been the dominant research strategy. The retrospective design cannot compete with the catch-up design in meeting these aims. The catch-up design deserves to be more widely used than it is, particularly in the area of sexual perpetration research.

Although the catch-up design can do much to provide additional leverage on the causal criteria, it is not well placed to test theories about the causal mechanism linking a risk factor to sexual perpetration. The challenge for future research is to move from risk factors to risk mechanisms. It is only when we have evidence of the mechanism that we can confidently begin to talk in terms of causation. As discussed in section 14.5, this will not be an easy task, but a number of plausible research
strategies could be used. One intriguing method consists of combining an existing longitudinal dataset with a catch-up design; this would allow the within-individual study of the temporal relationship between risk factors and hypothesised mechanisms. Corroboration of the predictions using this technique would constitute a risky test of the mechanisms, which is much needed in this area of research.

The use of a catch-up design has also demonstrated that we can predict which sexually victimised males are at risk of sexual perpetration before the onset of this behaviour. This research study is the first to provide evidence of this. If we can predict sexually abusive behaviour then this raises the possibility that we can prevent both the behaviour and its associated human costs.


Gunderson (Eds.), *Paul E Meehl: Selected philosophical and methodological papers*. Minneapolis: University of Minnesota Press.


Appendix 1

Statement of independent contribution
In accordance with section 7.2 (subsections f and j) and section 7.8 of the University College of London PhD Regulations and Procedures, this section summarises the independent contribution of the author to the research reported in this thesis. This statement has been certified by both supervisors (Dr Richard Hastings and Professor David Skuse).

The arguments developed in the literature review, methodology chapters and discussion chapter were the sole responsibility of the author. This included the arguments developed in chapter 4, 9 and 14 about the examination of causal criteria in this area of research. The discussion in chapter 10 about the assessment of the usefulness of a risk index was also developed by the author.

The author was solely responsible for the decision to use only contemporaneous information written before the onset of perpetration to code the risk factors. This methodological strategy allowed the study to examine the causal criteria of non-spuriousness and temporal precedence, which is a central feature of the research reported in this thesis.

The author was also solely responsible for the analysis strategy developed to examine the performance of the risk indexes. Part of this strategy included the development of equations to assess the performance of an index under different assumptions about base rates, treatment efficacy, the proportion selected for treatment and dose-response relationships in treatment. These equations were developed solely by the author.

Germain, & Brassard, 1987; Haugaard, 1991; McGee & Wolfe, 1991a&b; McGee, Wolfe, Yuen, Wilson, & Carnochan, 1995; Manly, Cicchetti, & Barnett, 1994; Mash & Wolfe, 1991; Wolfe, & McGee, 1994; Zuravin, 1991). As described in chapter 9, the use of contemporaneous case material can present substantial methodological difficulties, particularly the need to code data from vague, incomplete information. The manual provided the researcher with detailed guidelines about managing these problems. This and all other feature of the manual was the sole responsibility of the author.

The author was solely responsible for the development of the methodological strategies for blinding the data. This included the division of extraction and coding into three tasks (extraction of data, coding of perpetration status, coding of risk factors) and the use of a ‘dummy end-date’ to mask perpetration status.

All analyses reported in the results section were conducted by the author.

The author was involved in the extraction and coding of case material alongside four other researchers. The operational definition of sexual perpetration was developed with the help of other researchers.
Appendix 2

Approval letter from the Great Ormond Street Hospital for Children NHS Trust / Institute of Child Health Research Ethics Committee
Great Ormond Street Hospital
for Children NHS Trust
and the Institute of Child Health
(University College London Medical School)

19 September 1997

Professor D Skuse
Behavioural Sciences Unit
ICH

FILE COPY

30 Guilford Street
London WC1N 1EH
Telephone: 0171 242 9789
Direct Fax: 0171 813 8234

Dear Professor Skuse

95BS10  A prospective study of the onset of sexually abusive behaviour in boys who were sexually abused in early childhood: validation of risk index.

Notification of ethical approval

The above research has been given ethical approval after review by the Great Ormond Street Hospital for Sick Children NHS Trust / Institute of Child Health Research Ethics Committee subject to the following conditions.

1. Your research must commence within twelve months of the date of this letter and ethical approval is given for a period of 36 months from the commencement of the project. If you wish to start the research more than twelve months from the date of this letter or extend the duration of your approval you should seek Chairman's approval.

2. You must seek Chairman's approval for any proposed amendments to the research for which this approval has been given. Ethical approval is specific to this project and must not be treated as applicable to research of a similar nature, i.e. using the same procedure(s) or medicinal product(s). Each research project is reviewed separately and if there are significant changes to the research protocol, for example in response to a grant giving bodies requirements you should seek confirmation of continued ethical approval.

3. It is your responsibility to notify the Committee immediately of any information which would raise questions about the safety and continued conduct of the research.

4. Specific conditions pertaining to the approval of this project are:
   - That material arising from this research will be published under the auspices of the Institute of Child Health.

Research and Development Office
Yours sincerely

Anna Jenkins
Secretary to the Research Ethics Committee

cc Mark Richards
Project Team Leader
Appendix 3

Approval letter from the
Office for National Statistics
Dear Professor Skuse

THE ONSET OF SEXUALLY ABUSIVE BEHAVIOUR IN BOYS

Dr Fox has now given his approval for your study to proceed. This approval will be valid for 18 months. If you are unlikely to submit the work to NHSCR within this time please contact me.

NHSCR have provided an estimate to do the work. It is estimated that 62 per cent of the cases will automatically match, leaving 38 per cent requiring operator intervention. Assuming the data is received on disc the estimate is approximately £757.60 plus VAT. If the work is submitted on paper the cost will be approximately £897.85 plus VAT.

The estimate does not cover cases which require manual tracing, or any coding. The rates which will apply are those applicable in the financial year when the work is carried out, not the year in which the study is accepted. Any additions to the study, or changes to it may be subject to additional charges. On completion of flagging, future events will be charged annually.

Please let me know if you accept the estimate and wish to go ahead with the work.

Please complete and sign the attached Confidentiality Declaration which sets out the conditions under which OPCS can release data. One copy should be returned to me, the other is for you to keep. I need to receive one signed copy before any data can be provided to you.

I am attaching the specification for submitting the information on disc. If you have any technical queries regarding the creation or format of your disc please contact our NHSCR staff at Southport, Lynne Aindow on 0151 471 4306 or Wendy Eldridge on 0151 471 4229. General enquiries about your study should be directed to Jackie Gallagher on 0151 471 4456. If you decide to use the paper listing you have already sent, then please let me know. There seem to be quite a few duplicates on this list.

Please contact Frank Edwards on 0151 471 4306, before submitting your information. He will allocate a unique Medical Research number for your study which must be included on your disc or listing.
Your disc or listing should then be sent to:
Research Section
NHS Central Register
Smedley Hydro
Trafalgar Road
Southport
Merseyside
PR8 2HH

Please contact me if you require any further information.

Yours sincerely

[Signature]

Mrs P A Riach
Medical Research Unit
Appendix 4

Approval letter from the Crime Committee of the Association of Chief Police Officers
Dear Professor,

Sexual Abuse and Criminal Careers

Thank you for your letter on this subject dated 19th October 1995.

This is clearly an interesting and worthwhile piece of research. Tony Butler's personal support is noted, you will be aware that he takes a lead in this subject area within Crime Committee.

In the circumstances I therefore have no problem, on behalf of Crime Committee, in supporting the research.

The question of access to case files is, of course, a matter for each Chief Constable. I would have no objections to you enclosing a copy of this letter when you write to them.

Yours

[Signature]

HONORARY SECRETARY

Professor D. P. Farrington,
Institute of Criminology,
University of Cambridge,
7 West Road,
Cambridge.  CB3 9DT
Appendix 5

Risk Factor Manual
Risk Factor Manual
Trawling Guidelines

This section provides directions to enable the researcher trawling the file to correctly extract the relevant information for coding. The process of file trawling consists of four stages:
- Stage 1: Before file trawling
- Stage 2: Completion of the onscreen front sheet
- Stage 3: Reading the file and extracting quotes
- Stage 4: Construction of a care table.

Stage 1: Before file trawling

Before the research trawls the file it is essential that this document is read in its entirety. The trawler must not only be familiar with the information discussed in this trawling section, but must also be familiar with the coding sections, including the contents of each of the risk factor chapters. Coding can be made only if a series of criteria are met; it is crucial that the trawler is aware of these and extracts the necessary information.

Stage 2: Completing the onscreen cover sheet

Refer to the General Coding Guidelines for full details, under the heading 'Role of the perpetrator coder'.

Stage 3: Reading the file and extracting quotes

The file should be read word for word and in a chronological order. The majority of the file is likely to consist of reports or documents in either handwritten or typed format, and these should be read in full. An attempt should be made to read all handwritten documents, even if the handwriting is difficult to interpret.

The trawler should read the file until a quote of interest is reached. A quote of interest is defined as one that the trawler thinks may change any scoring assigned to a subject.

The next step is to determine whether the quote meets the criteria for extraction. In general, the trawler should extract any quote that s/he feels will change the coding, and should also extract a quote if s/he is unsure about whether the quote would affect the scoring. The general rule is: If in doubt, extract.

However, the trawler should not extract a quote if the same piece of information has already been extracted and the scoring would not be affected by the additional quote. (E.g. Quote already in the database: ‘The children’s clothes are unclean but adequate’; New quote: ‘Both boys were dressed in dirty clothes’. The new quote would not be recorded.)

This final point is designed to limit the amount of information extracted by the coder, but there are exceptions and these are described below.

Source of authority

If the source of the information is more authoritative, then the quote should be extracted, even if it refers to the same information. An ‘authority table’ exists to guide this decision (see below).
The authority table

<table>
<thead>
<tr>
<th>Authority level</th>
<th>People</th>
</tr>
</thead>
<tbody>
<tr>
<td>High authority</td>
<td>Psychologists, Psychiatrists, Psychotherapists, Psychiatric social workers</td>
</tr>
<tr>
<td>Low authority</td>
<td>All others, including results of court cases, lawyers and other members of the judiciary, police officers, and putative perpetrators</td>
</tr>
</tbody>
</table>

The trawler should extract information if the source of authority is a clinician or social worker, even if the subject or anyone else has reported the same information. In the same way, the trawler should extract information when the source is the subject, even when the same information has been reported by anyone else other than clinicians or social workers.

**Contradictory information**

Information that contradicts extracted information should also be extracted.

**Veracity of allegation**

The opinion of a professional as to the veracity of an allegation and a professional reporting an allegation without commenting on its veracity are not the same type of information. (E.g., 'The child alleges that he was forced to masturbate the father'; 'There is no doubt in my mind that the child's allegation are true.') The opinion of the professional provides vital evidence that the event did or did not occur, but a report of an incident without an opinion often contains important details about the experience, such as severity and duration.

Extraction of additional information

If the quote meets the criteria for extraction then it is to be typed on screen. Some additional pieces of information are also needed:

**The source of the information**

It is important to include their position/status.

**The date of the document from which the quote is taken**

**Other information necessary for understanding the meaning of the quote**

It is difficult to specify what additional information will be needed. As a general rule, you should establish what you understand by the quote, reread it and then check that you would arrive at the same understanding without additional information. Remember that if any additional information is recorded be sure to place [square brackets] around it, so that it can be distinguished from the quote itself.

**Contact**

The file trawler must remember to record information on the contact between a subject and a perpetrator if that individual was never or is no longer a carer. If you do not do this, it will be impossible to estimate duration of contact for any perpetrator other than a carer.
Stage 4: Construction of a care table

The trawler is responsible for extracting quotes relevant to the subject's care history, and using this information to generate a care table for the subject. The completed care history provides a 'time line' that can be referred to for a summary of the subject's life, as well as for more detailed information to be used during the coding process (e.g. no. of separations, duration of child sexual abuse).

Definitions

Main caregiver
A person who has ultimate responsibility for the subject or who shares this responsibility with one other person (e.g. biological father/mother, foster father/mother, adopting father/mother, step-father/mother, grandfather/mother).

There are additional characteristics of the main caregiver role:
- The role can pass from one person to another as people leave and enter the subject's life.
- If a subject moves from one main caregiver or set of main caregivers to a completely different set, then the new carers are main carers from day one. (E.g. Subject lives with mother, social services move the child to a foster family. The foster family are main carers from day one.)
- If the subject is living with a main caretaker, and a partner of the main caretaker begins to live with them, then the new individual must be present for one month or more before they are defined as main caretaker. However, once a person has cared for the subject for a month or more, and has become a main caretaker, that person is assumed to have been a main carer from the first day s/he began to live with the subject. Any calculation of duration is to be taken from the first day s/he began to live with the subject. (E.g. The mother is living with the subject, a boyfriend of the mother moves in on 1 January 1986. The boyfriend remains there until 2 November 1986. Once the boyfriend has lived there for a month, he becomes a main carer, and is assumed to have been a main carer since 1 January 1986.)
- A partner of a main caretaker cannot achieve either main carer or primary status on a cumulative basis. The caretaker has to have been caring for the subject continuously for a month.
- If a partner of a main caretaker achieves main carer status at some point in the child's care history, they maintain this status throughout the span of the child's history. For example: if a partner of a main caretaker leaves a care unit, and then resumes caring for the child in a later unit, they would automatically achieve main carer status in this new unit.
- If a subject lives with several people, say his mother and both grandparents, the main caretaker must always be assumed to be the closest familial relation to the subject. Grandparents represent closer relations than uncles or aunts, who are closer than all other family relations, who are closer than non-relatives.

Primary Carer
A primary carer is a main carer who has cared for the subject for one year or more.

The following are additional aspects of the primary caretaker role:
- If the subject is living with a primary caretaker, and a partner of the primary caretaker begins to live with them, then that new person must be there for one year or more before they are defined as a primary caretaker. However, once a person has cared for the subject for a year or more that person is assumed to have been a primary carer from the first day s/he began to live with the subject.
- If a carer achieves primary carer status at some point in the child's care history, they maintain this status from that time onwards. For example: if a primary carer leaves a care unit, and then resumes caring for the child in a later unit, they would automatically achieve primary carer status in this new unit.
- A carer cannot achieve primary carer status on a cumulative basis. The carer must have continuously cared for the subject for a year or any separations during that year were for less than one month.
Care Unit
A care unit refers to a period of care by the same person or pair of people meeting the definition for main carer or primary carer, which lasts at least one month. Any change in the carers within one unit, whether an addition or subtraction, marks the end of that care unit. This change may also mark the beginning of a new care unit provided that the new arrangement lasts at least one month.

Separation
A separation occurs if the subject is separated from one or more primary carers for a period of one month or more. Note that this refers only to primary carers. Simultaneous separation from two primary carers counts as two separations.

Trawling instructions
Quotes should be extracted that contain information about any change in carers that would constitute a new care unit. For each change in carers that meet the definition for a new care unit, a care unit will need to be constructed. The care units should be placed in temporal order. The care unit should record the following; the trawler should extract sufficient information to complete each of these sections:
- The start date and end-date
- The name of each of the people caring for the subject
- The role of all of the people caring for the subject (e.g. foster parent, boyfriend of the mother, grandparent)
- The reason the care unit ended. Use the following categories: divorce of carers, death of carer(s), sexual abuse of subject, other subtype of maltreatment of subject, drug use / mental illness of previous carers, carers unable to cope, carers abandoned the subject, more suitable placement found, other (specify) or don’t know.
- Details of any contact with previous primary carers (extract whether there was contact, the amount of contact and the nature of the contact)
- Details of contact with any perpetrators of any form of maltreatment who were not carers
General Coding Guidelines

This chapter discusses difficulties that apply to all or many of the risk factors. Rather than replicate this information in each of the chapters, it is summarised here, and should be referred to as necessary.

Levels of certainty

When coding the maltreatment subtypes, particularly Child Sexual Abuse, it is often necessary to base a decision on a degree of probability rather than absolute certainty. There may also be contradictory information about whether a subtype of maltreatment did or did not occur.

A statement made by a clinician or social worker that a subtype of maltreatment occurred is usually sufficient on its own to make a coding. However, the statement must make it clear either that the abuse did occur or that it probably occurred. A statement that it possibly occurred is not sufficient.

If someone other than a clinician or social worker, makes such a claim, this is also enough to make a coding, provided that the information given is not contradicted by anyone on a higher level in the authority table. In such cases, the statement of the higher authority is the one to be coded (see authority table above).

Occasionally two or more clinicians or social workers disagree about whether a type of maltreatment occurred. In this situation, it will be necessary for the coder to make a decision based on available information. It will be necessary to consider whether one of the professionals had more access to important information when arriving at their decision.

The role of the Perpetrator Coder

The perpetrator coder must complete the following tasks:
• Establish perpetration status of the subject
• Establish when perpetration by the subject began and set additional cut-off date for the other coder
• Complete the coding sheet for ‘subject a perpetrator of sexual abuse’.
• Remove all references to perpetration status from quotes tagged for risk factors
• Remove all references to perpetration status that appear in the care history table
• Complete the coding of the onscreen cover sheet.

Perpetration status

Details are given in the chapter ‘Subject as a Perpetrator of Sexual Abuse’.

Date of first perpetration

Details are given in ‘Onset of Perpetration or Sexualised Behaviour’.

Additional cut-off date

The longitudinal integrity of the study requires risk factors to be coded using quotes from documents that were dated before the onset of perpetration. However, if the risk factor coder was required to code up to only one of these ‘cut-off’ dates, s/he would be immediately aware of the perpetration status of the subject. With this knowledge, the risk factor coders might be tempted to adjust their risk factor scores in the direction of the hypothesis.
An alternative coding strategy will be used to avoid this problem. All subjects are to be coded twice: once using a cut-off date of the subject’s sixteenth birthday and once using another cut-off date. If the subject perpetrated prior to his sixteenth birthday, this alternative cut-off will be the date he began abusing. If the subject has not perpetrated before this date, the alternative cut-off will be a randomly selected date sometime between the subject’s sixth and sixteenth birthday.

**Random date:** The perpetrator coder will use the random number generator to select a number between 1 and 119 to represent the month. The perpetrator coder will then calculate which month this number relates to using the method described below, and assign the first day of this month as the additional cut-off date.

\[ X = \text{Random Number generated by computer, } 1 < X < 119 \]

\[
\text{Equivalent age of subject} = \frac{6 + X}{12}
\]

Then by adding this number of years and months to the subject’s date of birth, calculate the random month and year. Finally, assign the date as the first day of that month.

The risk factor coder will be unaware whether this second cut-off date has been randomly selected or is the date of first perpetration.

**Removal of references to perpetration status**

The perpetrator coder must establish that no quotes that refer to risk factors also refer to the perpetration status. The quickest way of doing this is to double click on the ‘Subject a perpetrator of sexual abuse’ tag and double click on each of the quotes that appear in this list to establish. If one or more of these have been tagged for risk factors, the perpetrator coder should split up the one quote into two or more quotes, making sure that perpetration status is not mentioned in any quote that refers to risk factors.

**Removal of quotes from the Care History Table**

The Perpetrator Coder should, for subjects who became perpetrators, remove all references to perpetration status from the care history table. Thus, if a subject started perpetrating at age 13 years and 4 months, his care history table should, as it currently stands, end at this point. However, to mask this fact, the perpetrator coder should delete all references to the subject’s perpetration and indicate in his/her own words that the care unit (which currently ends at 13 yrs 4 months) continued until the subject was 16 yrs.

**Coding of the Onscreen Cover Sheet**

The onscreen cover sheet requires basic information on the subject and is divided into 5 sections.
Onset of Perpetration

It is important to clearly establish the date of first perpetration. This is so that the coder can decide whether a recorded quote can be used to assess risk exposure.

This date of onset tends to be recorded in a limited number of ways. Although rare, a precise date (e.g. 14 February 1989) is sometimes given for a first offence. If this happens, take the first day of the month as the date (e.g. 1 February 1989), so that the risk factor coder cannot distinguish genuine from dummy end-dates. It is more common for the file to provide only an inexact date of onset, such as the subject's age at the time. Here the perpetrator coder should use the following guidelines.

Subject's Age

When the subject's age is used to describe the onset of perpetration, the day the subject turned that age is taken as the date of first perpetration. For example:

'The subject was ten when he first sexually abused his sister'
The day the subject turned ten would be taken as the date of perpetration.

'The subject was between eight and nine years old when he first sexually abused his sister'
In this case, the subject's eighth birthday would be taken as the onset of perpetration.

Number of years ago

When a quote from a dated document states that the subject began perpetrating X number of years ago, the coder should subtract X years from the document date and use the mid-point of that year (i.e. 1st July) as the onset of perpetration.

'The subject began perpetrating six years ago (document date 28/5/88)' In this case the coder would determine that first perpetration was on 1/7/82.

Should the quote give a range of years, the earliest first date of perpetration should be used.

Should the document be undated and this is the only quote from which the onset of perpetration can be estimated, the case should be discussed with another member of the research team.

Vague terminology

"A couple": subtract 2 years from the document date and use the mid-point of that year
"A few"/ "Several": subtract 3 years from the document date and use the mid-point of that year
Calculation of duration

Dealing with vague information
Sometimes the information on the start or end of contact or on the duration of contact will be vague. The general rule is to take the midpoint of the range, using the most accurate information possible on the range.

Start point during or between a number of years:
Sometimes the start point may be expressed as occurring at some point between or during a number of years. (E.g. 'Between 1980 and 1983 James went to live with his father'.) In these situations take the midpoint of the range as the starting point. Assume that the range is from the 1st of January for the earliest year to the 31st December for the later year. (E.g. In the above example the range would be 01.01.80 to 31.12.83, which would make the midpoint 01.01.82).

Start point in a year
Sometimes the start point may be expressed as occurring during a year. (E.g. 'In 1980 James went to live with his father'). Take the midpoint, which would be the 1st of July 1980.

Start point between months
Sometimes the start point may be expressed as occurring at some time between or during a number of months. (E.g. Between March and September 1980 James went to live with his father.) In these situations take the midpoint of the range as the starting point. Assume that the range is from the 1st day of the early month to the last day of the later month. (E.g. In the above example the range would be from the start of March to the end of November, which would make the midpoint the middle of July.)

Start point in a month
Sometimes the start point may be expressed as occurring sometime during a month. (E.g. 'In November 1980 James went to live with the father.') Take the midpoint, which would be 16 November.

End-point - years and months
Apply the same principles given above for start points.

Starting point or end-point expressed as age
Here we make different assumptions about the range. The range is estimated from the first day of the earlier point to the first day of the last point. E.g. Between three and four years. First day of earlier point is the day at which the subject turned three; first day of the later point is the day at which the subject turned four. Making the mid point when the subject was exactly three and a half years.

Duration expressed as years or months
Here the range is estimated from the first day of the earlier point to the first day of the last point. E.g. 'X lived with the father for between five and eight months.' This would score as six and a half months.') E.g. 'X lived with his father for between five and seven years.' This would score as six years.

Rounding up and rounding down
It is possible that your final calculation will involve a proportion of a month (e.g. Seven years, and two and a half months; six months and five days). If the proportion of the month is expressed as a fraction, round up half a month or a larger fraction, round down anything else. If the proportion of a month is expressed in days round up 15 days or more (regardless of the number of days in the month) round down anything else.
Dealing with contradictory information
If quotes in the care table are contradictory then take the midpoint.

For non-carers, the duration of maltreatment may be reported, and there may be a disagreement as to how long the maltreatment lasted. In these cases, the authority table should be consulted, and the individual with the highest authority is to be believed. If there is no difference in terms of authority then calculate the midpoint.

**Coding Duration**

The contact duration (minimum) is made on the basis of the length of contact between the subject and the perpetrator, taking account of any periods in which the subject and perpetrator were not in contact. Note that it is not designed as an estimation of the duration of maltreatment. Contact duration (minimum) can be established from the care history table and from the quotes tagged to the relevant maltreatment factor.

**More than one perpetrator**

If the maltreatment of the subject involves more than one perpetrator, the 'contact duration' will be calculated according to the length of contact with all of the perpetrators, but the same period of time is not to be counted twice.

**Total contact duration**

This is the length of time from when the subject was first in contact with any of the perpetrators of the maltreatment to the point when they last had contact with any of them. Take care not to count the same period of time twice. Subtract from the duration any periods of time in which the subject was in contact with none of the perpetrators. If there are a number of perpetrators, who criss-cross in their contact it may be necessary to plot the contact out on a 'Duration time line sheet'.

If the contact with a perpetrator is not known, ignore this perpetrator when making the calculation.

If contact information is known for none of the perpetrators, code as insufficient information.
Duration expressed in years and months is to be converted by the coder to decimal years. The following table gives the decimal conversion for months:

<table>
<thead>
<tr>
<th>Number of months</th>
<th>Decimal Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td>0.17</td>
</tr>
<tr>
<td>3</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>0.33</td>
</tr>
<tr>
<td>5</td>
<td>0.42</td>
</tr>
<tr>
<td>6</td>
<td>0.50</td>
</tr>
<tr>
<td>7</td>
<td>0.58</td>
</tr>
<tr>
<td>8</td>
<td>0.67</td>
</tr>
<tr>
<td>9</td>
<td>0.75</td>
</tr>
<tr>
<td>10</td>
<td>0.83</td>
</tr>
<tr>
<td>11</td>
<td>0.92</td>
</tr>
</tbody>
</table>
Subject as a Perpetrator of Sexual Abuse

Introduction

The definition of perpetration is split into two age dependent categories:

- Perpetration by subjects who are aged six or over but under twelve. These will be referred to as 'Younger perpetrators'.
- Perpetration by subjects who were twelve or older. These will be referred to as 'Older perpetrators'.

Definitions that are specific to only 'Younger' or 'Older' perpetrators are written below in the sections dealing with these two groups. The operational definitions that are common to both categories of abuse are described at the end of this chapter.

Perpetration by 'Younger' subjects aged six or over but under the age of 12.

Definition of Perpetration by 'Younger' subjects

The following three categories represent behaviours that are to be classified as perpetration of sexual abuse by 'younger' subjects. Below these categories are two additional criteria to be considered when assessing perpetration status.

1. Genital Contact

Direct contact with the genitals of the victim when the victim is not clothed or their clothing has been partially removed counts as perpetration. Alternatively the victim may be required to touch the genitals of the unclothed subject under duress. Under both circumstances, the following criteria must be met:

EITHER
- There is an indication that the act was repeated more than twice (this could either be with different victims or the same victim on different occasions) and there is a two-year age gap between perpetrator and victim.

OR
- If physical coercion of any kind is used

OR
- If the use of an offensive weapon is threatened

Contact with genitals can be assumed if the word 'masturbated', or a phrase with similar meaning, is used in relation to a physical act involving the subject and his victim, i.e. one person must have been masturbating the other.

'More than twice' would be suggested by phrases such as 'several times', 'many times' or even 'a few times'.

2. Penetrative acts

- A penetrative act only has to occur once to be considered as perpetration.
- A two-year age gap must be apparent between perpetrator and victim, unless physical coercion of any kind has been used.
For definitions of penetration, physical coercion, and two-year age gap see the 'Operational Definitions' section at the end of this chapter.

3. General Statements of acts

If a clinician or social worker state that a subject under the age of twelve 'has sexually abused another child' this would be counted as perpetration.

Coding of duration of perpetration

As a result of the young age of the subjects who would qualify as 'younger' perpetrators, it is likely that the only source of data concerning the duration of their perpetrating behaviour would be from their hospital or social services files. The coder should attempt to establish from these files the date on which perpetration commenced and the date of the last recorded episode of perpetration.

For the date of last known perpetration, enter the latest date that the subject was identified as having perpetrated abuse. This is rarely stated, but can often be calculated from the information given. If the perpetration is identified as a one-off, put the same date for last identified date as for the date of onset. If there is more than one incident of perpetration by the subject, put the latest incident in. If the quote states that he perpetrated for a specific number of months, add this number to the date of onset and give the first day of the month just calculated. If the quote states that he perpetrated for a specific number of years, add this number to the date of onset and give the date as the 1st July in that year.

If it is not possible to make such an estimate from the information given, the coder should indicate that there is insufficient information to answer this question, but if possible answer the subsequent duration questions.

Perpetration by 'Older' subjects 12 years of age or more.

The following three criteria must be met:

1) The victim was under the age of 16 years at the time of the abuse
2) The abuse involved sexual contact between the subject and the victim (see definition of sexual contact given below)
3) EITHER
   the subject was two years or more older than the victim
   OR
   the subject employed physical coercion (see definition of physical coercion)
   OR
   the subject threatened the victim with an offensive weapon.

Sexual Contact

The criterion for sexual contact is met if the act involves physical contact between a victim and a perpetrator and that contact involves attempts, successful or otherwise, to perform sexual acts with the victim. The acts that meet this criterion are described extensively in the 1 to 5 severity levels for this section. Note that contact is assumed to have taken place if the phrases 'sexual abuse', 'molested' or 'indecent assault' is used. In addition, the criterion for sexual contact is met if an individual uses physical coercion to permit another individual to perform sexual acts on the victim, either successfully or unsuccessfully.
Severity Levels - Perpetration against children or adults by 'older' subjects aged 12 or more

The following are behaviours that are classified as perpetration. If the subject meets the criteria for one of the perpetration definitions given above and has performed one of the following behaviours, he can be classified as a perpetrator of sexual abuse.

Severity score 1: Non-genital contact

- Physical contact between the subject and the victim that contains a sexual element, but involves neither the touching of the victim's genital area by the subject nor the touching of the subject's genital area by the victim. See the definition of genital given above. Note that unsuccessful attempts to perform such acts also meet the criteria as long as some form of physical contact was involved.
- This level is to be scored if the file states that the victim was sexually abused by the subject, but no other information is contained in the file.
- If the file states that the victim of the subject was on the Child Protection Register for 'Sexual Abuse' or 'Actual Sexual Abuse' and no other information is given, then this level is to be scored. If the file states that the victim of the subject was on the Child Protection Register for 'Potential Sexual Abuse' or any variation of that phrase (e.g. 'Grave Concern' or 'Likely') then, in the absence of further information, this is not to be coded as sexual perpetration.
- Sometimes a word or phrase will be used to describe a sexual behaviour, but that word or phrase does not make it clear what the act involved. Examples include 'molested' and 'indecent assault'. If this is the only evidence contained in the file then score at level one.

Examples
- The subject grabs at the victim and tries to fondle his bottom.
- The subject holds the victim and tries to 'French kiss' him.
- The file states that the victim was sexually abused by the perpetrator; no other information is given.
- The subject has a conviction for indecently assaulting the victim; no other information is given.
- The victim reports that the subject 'sexed' him; no other information is given.

Severity score 2: Non-penetrative genital contact

- The subject touches the victim's genital area (clothed or unclothed) and/or the victim touches the subject's genital area (clothed or unclothed), but no penetration is involved. Note that unsuccessful attempts to perform such acts also meet the criteria as long as some form of physical contact is involved.
- This level includes: masturbation of the victim by the subject, masturbation of the subject by the victim, and simulated intercourse that does not involve penetration.
- If penetration is involved score at a higher level.
- If it is unclear whether penetration did or did not take place then score at this level.

Examples
- The subject forces the victim to caress his penis and testicles.
- The subject engages in mutual masturbation with the victim.
- The victim is forced to caress the subject's anus, but no penetration is involved.
- The subject rubs his penis up against the victim's penis.
Severity score 3: One form of penetration

- Three is to be scored if one form of penetration is involved. See the definition of penetration given above.
- The definition of one form of penetration is penetration involving one type of orifice.
- If the penetration consisted of more than one type of penetration then score it at level four.

Examples
- A perpetrator has oral sex with his younger brother.

Severity score 4: More than one form of penetration

- Four is to be scored if more than one form of penetration is involved. See the definition of penetration given above.
- More than one form of penetration is defined as penetration involving more than one type of orifice.

Examples
- It is reported that the victim had oral and anal sex with the subject.
- The victim reports that he performed oral sex on the subject; he later reveals that the subject had annals sex with him.

Severity score 5: Penetration plus physical abuse, defecation and urination

- The subject has used physical abuse, as defined above, while performing intercourse or other forms of sexual penetration. For a 5 to be scored, the physical act must meet or exceed the minimum criteria for 'experiencing physical abuse', e.g. use of restraint. Note that physical injuries that are a direct result of the sexual penetration do not count as physical abuse.
- Note that if the word 'raped' is used it cannot be assumed that physical abuse was also performed and so a five cannot be scored unless there is additional information indicating that it was.
- Five can also be scored if the sexual act involves the subject defecating or urinating on the victim or the victim having to defecate or urinate on the subject. Note that to be scored in the sexual abuse section, defecation or urination must be part of a sexual act or must involve a sexual element. For example, urinating on the victim as part of other acts with a sexualised element such as kissing the victim; however, urinating on the victim solely for the purpose of demeaning him would be scored in rejection by carers, not this section.

Examples
- The subject ties the victim to the bed and has forced sexual intercourse with him.
- The subject sodomizes the victim having beaten him, until he agreed to the act.
- After a series of sexual acts, such as the subject touching the victim's genitals, the subject defecates on the victim's chest.
Coding the duration of perpetration

Information relating to the duration of their perpetrating behaviour may be available either from their hospital or social services files or from criminal record data.

In order to code duration the coder will attempt to establish from these data sources the date on which perpetration commenced and the date of the last recorded episode of perpetration. This could be found in quotes in which these dates are clearly stated or from dates of cautions or convictions. It is possible that this information is not available in either of the ways stated above. If these dates cannot easily be calculated on the basis of the information given, it is important that the coder does not attempt to make a ‘best guess’ from this information.

For the date of last known perpetration, enter the latest date that the subject was identified as having perpetrated abuse. This is rarely stated, but can often be calculated from the information given. If the perpetration is identified as a one-off, put the same date for last identified date as for the date of onset. If there is more than one incident of perpetration by the subject, put the latest incident in. If the quote states that he perpetrated for a specific number of months, add this number to the date of onset and give the first day of the month just calculated. If the quote states that he perpetrated for a specific number of years, add this number to the date of onset and give the date as the 1st July in that year.

If it is clear from the criminal record data that the subject was cautioned or convicted for a sexual offence, and that this was later than any offences recorded in the GOSH or social service file, the coder should enter the latest caution or conviction date. It is possible that the subject was convicted for an offence a number of months or years after the offence was originally mentioned in the social services file. If it is not clear that the two data sources are referring to two different events (i.e. victims are of different genders, ages), assume that they refer to the same event and enter the earlier date.

If it is not possible to make such an estimate from the information given, the coder should indicate that there is insufficient information to answer this question, but if possible answer the subsequent duration questions.
Operational Definitions common to both Younger and Older Perpetrators

The following are definitions that are relevant in the classification of both Younger and Older perpetrators.

Physical Coercion

The concept of 'physical coercion' is necessary to classify behaviours as sexually abusive when there is less than a two-year age gap between the perpetrator and the victim. Physical coercion is defined as physical behaviour designed to ensure compliance in a sexual act against the subject’s wishes. This must include contact either between the perpetrator and the subject or an object and the subject, provided that the perpetrator has deliberately caused this contact.

Actions that meet the criteria for physical abuse would, of course, also meet the criteria for physical coercion if they were designed to ensure compliance against the victim’s wishes. If the word ‘forced’ is used, e.g. the subject was forced to have sex with his twin brother, assume that the force involved physical coercion and therefore code as sexual abuse.

Genital Area

For the purpose of this section, the genital area is defined as the penis, testicles and anus for males and the vagina and anus for females. Note the breasts are not considered part of the genital area.

Penetration

Penetration consists of penetration, including attempted penetration, of the vagina, anus or mouth by a penis, or the penetration, including attempted penetration, of the vagina or anus by a finger or objects, such as the handle of a hairbrush or a stick. Note that this definition includes both penetration of the subject as well as penetration by the subject.

Two Years Difference in Age

This can be established from any quote that indicates that there is a two-year difference between the subject and the victim(s). Alternatively, any clear statement of the age of the victim(s) or statements from which this age can be logically deduced could be used to establish the two-year age difference. An example of such a statement might be that the victim is clearly identifiable as being of pre-school age (under 5) whilst the subject is an adolescent (over 12). A two-year age difference can also be assumed if the subject is described as the victim’s uncle, or the victim is described as the subject’s niece or nephew. If the subject is described as an older boy with reference to the victim, the coder should interpret this as involving a two year difference.

*If the coder is in any doubt about the age differential, assume that there is not a two-year gap.*
Discontinuity of care

Definitions

The definitions of main caregiver, primary carer, care unit and separation are given in the chapter on ‘Trawling Guidelines. These definitions are important for the coding of this risk factor. Please refer to them before coding.

Potential problems

Establishing duration of contact or separation

A separation can be coded only if the subject is separated from someone who has cared for the subject for one year or more and when that separation lasts for one month or more. Sometimes these lengths of time are not clearly stated in the file. Duration of one month or one year can be inferred in the following circumstances:

Start date and end-date are clearly stated, and the difference between the two is equal to or greater than 28 days or 365 days
Examples:
Mother left home on 20 February 1984 and returned home on the 21 March 1984.  
Mother left home on 20 February 1984 and returned home on 20 February 1985.

One date is given, and it can be reasonably inferred that there is one month or more / one year or more difference between the two dates
Examples:
Mother left home on 1 February 1984, and returned home mid April 1984.
Mother left home on 1 February 1984, and returned home mid February 1985.

The month of both dates is given, and the difference is greater than one month or one year
Examples:
Father left home in February 1984 and returned home in April 1984.
Father left home in February 1984 and returned home in March 1984.

A duration is given, and that duration is equal or greater than one month / one year.
Examples:
The father left the home for 28 days / four weeks / one month

References to early / mid / late
References to early, mid or late portions of a month or year may sometimes be sufficient
Examples
The mother left the home in mid February 1984 and returned home late March 1985
Dichotomous question: Did the subject experience one or more separations?

Code as present if the following criteria are met:
- At least one of the experiences occurred before the subject's 16th birthday or the date of the subject’s first perpetration, whichever took place first.
And for at least one experience, the following criteria are also met:
- The subject was separated from someone meeting the criteria for primary carer
- The separation lasted for one month or more

Number of separations by age sixteen / onset of perpetration

Sum the number of separations experienced by age sixteen or the date of first perpetration, whichever took place first.

Before including a separation in the total, ensure that the following criteria are met:
- The experience occurred before the subject’s sixteenth birthday or the date of the subject’s first perpetration, whichever took place first
- The person the subject was separated from meets the criteria for primary carer
- The separation lasted for one month or more

Note that the simultaneous separation from two primary carers counts as two separations.
**Child Sexual Abuse**

**Definitions**

**Child sexual abuse**

The following three criteria must be met: 1) the subject was under the age of 16 years at the time of the abuse; 2) the abuse involved sexual contact between the subject and the perpetrator (see definition of sexual contact given below); and 3) EITHER the perpetrator was two or more years older than the subject OR the perpetrator employed physical coercion (see definition of physical coercion below).

**Sexual Contact**

The criterion for sexual contact is met if the act involves physical contact between a victim and a perpetrator and that that contact involves attempts, successful or otherwise, to perform sexual acts with the victim. The acts that meet this criterion are described extensively in the 1 to 5 severity levels for this section. Note that contact is assumed to have taken place if the phrases ‘sexual abuse’, ‘molested’ or ‘indecent assault’ are used. In addition, the criterion for sexual contact is met if a person uses physical coercion to allow another person to perform sexual acts with the subject, either successfully or unsuccessfully. (E.g. the mother pins the subject down whilst the father has anal sex with him.)

**Physical coercion**

The concept of ‘physical coercion’ is necessary to classify behaviours as sexually abusive when there is less than a two-year age gap between the perpetrator and the victim. Physical coercion is defined as physical behaviour designed to ensure compliance in a sexual act against the subject’s wishes. This must include contact either between the perpetrator and the subject or an object and the subject, provided that the perpetrator has deliberately caused this contact.

Actions that meet the criteria for physical abuse would, of course, also meet the criteria for physical coercion if they were designed to ensure compliance against the victim’s wishes. If the word ‘forced’ is used, e.g. the subject was forced to have sex with his twin brother, assume that the force involved physical coercion and therefore code as sexual abuse.

**Genital area**

For the purpose of this section, the genital area is defined as the penis, testicles and anus for males and the vagina and anus for females. Note the breasts are not considered part of the genital area.

**Penetration**

Penetration consists of penetration, including attempted penetration, of the vagina, anus or mouth by a penis, or the penetration, including attempted penetration, of the vagina or anus by a finger or objects, such as the handle of a hairbrush or a stick. Note that this definition includes both penetration of the subject as well as penetration by the subject.

**Perpetrator criteria**

Any person, no matter what relationship they have with the subject, can be a perpetrator of sexual abuse.
Entering information in the correct factor

Child sexual abuse and physical abuse

Physical injuries that occur as a direct result of sexual interaction (e.g. rectal tears) are coded solely under sexual abuse. Physical coercion before, during or after the sexual interaction can be scored in the experiencing physical abuse section, if the criteria for physical abuse are met. If physical abuse occurs in the context of sexual penetration, this is to be coded as level five severity of sexual abuse, and the physically abusive element is also to be coded under physical abuse.

Potential problems

Reference to a group being maltreated

If the file refers to a group being maltreated and this group includes the subject, then it is permissible to infer that the subject has experienced this form of maltreatment.

Dichotomous question: Did the subject experience child sexual abuse?

Code as present if the following criteria are met:
• At least one of the experiences occurred before the subject’s 16th birthday or the date of the subject’s first perpetration, whichever took place first.
And for that or those experiences, the following criteria are also met:
• Sexual Contact was involved (see definition)
• EITHER the perpetrator of the abuse was two or more years older than the subject OR physical coercion was used (see definition)

If the criteria are not met, code no.
Severity levels for Child Sexual Abuse

Severity score 1: Non-genital contact

- Physical contact between the perpetrator and the subject that contains a sexual element, but involves neither the touching of the subject’s genital area by the perpetrator nor the touching of the perpetrator’s genital area by the subject. See the definition of genital given above. Note that unsuccessful attempts to perform such acts also meet the criteria as long as some form of physical contact was involved.
- This level is to be scored if the file states that the subject was sexually abused, but no other information is contained in the file.
- If the file states that the subject was on the Child Protection Register for ‘Sexual Abuse’ or ‘Actual Sexual Abuse’ and no other information is given, then this level is to be scored. If the file states that the subject was on the Child Protection Register for ‘Potential Sexual Abuse’ or any variation of that phrase (e.g. ‘Grave Concern’ or ‘Likely’) then, in the absence of further information, this is not to be coded as sexual abuse.
- Sometimes a word or phrase will be used to describe a sexual behaviour, but that word or phrase does not make it clear what the act involved. Examples include ‘molested’ and ‘indecent assault’. If this is the only evidence contained in the file then score at level one.

Examples
- The subject is made to suck the perpetrator’s breasts.
- The perpetrator grabs at the subject and tries to fondle his bottom.
- The perpetrator holds the subject and tries to ‘French kiss’ him.
- The file states that the subject was sexually abused; no other information is given.
- A father has a conviction for indecently assaulting the subject; no other information is given.
- The subject reports that the father ‘sexed’ him; no other information is given.

Severity score 2: Non-penetrative genital contact

- The perpetrator touches the subject’s genital area (clothed or unclothed) and/or the subject touches the perpetrator’s genital area (clothed or unclothed), but no penetration is involved. Note that unsuccessful attempts to perform such acts also meet the criteria as long as some form of physical contact is involved.
- This level includes: masturbation of the subject by the perpetrator, masturbation of the perpetrator by the subject, and simulated intercourse that does not involve penetration.
- If penetration is involved score at a higher level.
- If it is unclear whether penetration did or did not take place then score at this level.

Examples
- The perpetrator forces the subject to caress his penis and testicles.
- The perpetrator engages in mutual masturbation with the subject.
- The subject is forced to caress the perpetrator’s anus, but no penetration is involved.
- The subject is forced to rub his penis over the perpetrator’s vagina, but no penetration takes place.
- The perpetrator rubs his penis up against the subject’s penis.

Severity score 3: One form of penetration

- Three is to be scored if one form of penetration is involved. See the definition of penetration given above.
- The definition of one form of penetration is penetration involving one type of orifice.
- If the penetration consisted of more than one type of penetration then score it at level four.
• (Note for global coding: If one perpetrator unit consists solely of one type of penetration, and another unit consists solely of another type, then both would score at level 3, and so the overall global coding of the most severe would be a 3, not a 4, even though overall two types of penetration were involved.)

Examples
• The subject has venereal disease. No information regarding the sexual contact is known.
• A father has oral sex with his son.

Severity score 4: More than one form of penetration

• Four is to be scored if more than one form of penetration is involved. See the definition of penetration given above.
• More than one form of penetration is defined as penetration involving more than one type of orifice.
• (Note for global coding: If one perpetrator unit consists solely of one type of penetration, and another unit consists solely of another type, then both would score at level 3, and so the overall global coding of the most severe would be a 3, not a 4, even though overall two types of penetration were involved.)

Examples
• It is reported that the subject had oral and anal sex with the perpetrator.
• The subject reports that he performed oral sex on the perpetrator; he later reveals that the same perpetrator had anal sex with him.

Severity score 5: Penetration plus physical abuse, defecation and urination

• The perpetrator has used physical abuse, as defined above, while performing intercourse or other forms of sexual penetration. For a 5 to be scored, the physical act must meet or exceed the minimum criteria for ‘experiencing physical abuse’, e.g. use of restraint. Note that physical injuries that are a direct result of the sexual penetration do not count as physical abuse.
• Note that if the word ‘raped’ is used it cannot be assumed that physical abuse was also performed and so a five cannot be scored unless there is additional information indicating that it was.
• Five can also be scored if the sexual act involves the perpetrator defecating or urinating on the subject or the subject having to defecate or urinate on the perpetrator. Note that to be scored in the sexual abuse section, defecation or urination must be part of a sexual act or must involve a sexual element. For example, urinating on the subject as part of other acts with a sexualised element such as kissing the subject; however, urinating on the subject solely for the purpose of demeaning him would be scored in rejection by carers, not this section.

Examples
• The perpetrator ties the subject to the bed and has forced sexual intercourse with him.
• The perpetrator sodomizes the subject having beaten him, until he agreed to the act.
• After a series of sexual acts, such as the perpetrator touching the subject’s genitals, the perpetrator defecates on the subject’s chest.
**Duration of child sexual abuse**

Refer to the section in the introduction.

Duration of contact with a particular perpetrator should be included in the calculation only if it is possible to code ‘Yes’ for the Dichotomous question solely on the basis of this perpetrator.

---

**Number of perpetrators of child sexual abuse**

Sum the number of perpetrators of child sexual abuse.

A particular perpetrator should be included in the calculation only if it is possible code ‘Yes’ for the Dichotomous question solely on the basis of this perpetrator.

---

**Gender of perpetrators of child sexual abuse**

A particular perpetrator should be used in assigning a score only if it is possible code ‘Yes’ for the Dichotomous question solely on the basis of this perpetrator.

If all of the perpetrators of the child sexual abuse against the subject are male then code ‘1’.

If all of the perpetrators of the child sexual abuse against the subject are female then code ‘2’.

If perpetrators of the child sexual abuse includes both males and females then code ‘3’.

---

**Closeness to perpetrator**

A particular perpetrator should be used in assigning a score only if it is possible code ‘Yes’ for the Dichotomous question solely on the basis of this perpetrator.

The levels of the closeness to perpetrator scoring system are based on the social role the perpetrator has in relation to the subject and are described in the next section.
The Closeness Levels

1. **Non-family: stranger, acquaintance, peer, person living with the subject in an institutional setting, lodger**
   - This level refers to people who are not family members and have no responsibility for the care of the subject.
   - A stranger is a person who was not known to the subject until the day of the abuse and who cannot be better described by a higher level.
   - An acquaintance is a person who is known to the subject but whose relationship cannot be better described by a higher level. A friend of the family would score at this level. A boyfriend of the mother or girlfriend of the father would also score at this level if they were not resident or, if living with the subject, it is clear that they had no care responsibilities for the subject.
   - People who live with the subject in an institutional setting and who are not authority figures also score at this level.
   - A peer is an acquaintance or friend similar in age to the subject.
   - Lodgers and boarders who reside with the subject, but who clearly do not have a caring role, score at this level.

2. **Family members and authority figures who do not have a caring role: same generation siblings and cousins, uncles and aunts and grandparents**
   - This level includes family members who are from the same generation as the subject, which is defined as the familial relationship between the subject and the perpetrator rather than proximity in age.
   - Full siblings, stepsiblings, foster siblings, adopted siblings and cousins are to be coded at this level whether or not they are resident with the subject.
   - Uncles and aunts who are not carers score at this level. This includes uncles and aunts who have never been resident with the subject or have been resident but have never had a caring role. If the criteria for carer are met, then they would score at level 3.
   - Grandparents who are not carers score at this level. This includes grandparents who have never been resident with the subject or have been resident but have never had a caring role. If the criteria for carer are met, then they would score at level 3.
   - Any person who is in a position of recognised authority over the child scores at this level. The term includes people such as teacher, residential carer, scout leader, vicar etc. Note that an adult is not automatically considered an authority figure by virtue of being an adult. Also note that an individual with a social role such as teacher, residential carer, scout leader or vicar is only to be considered an authority figure if s/he has that role in relationship to the subject.
   - This level also includes parents of the subject who have never been main carers or primary carers for the subject.

3. **All main carers and primary carers**
   - This level includes all those individuals who are main carers or primary carers.
   - Any person, regardless of their previous or subsequent social role or familial relationship would score at this level if they have been a main carer or primary carer for the subject. Therefore, uncles, aunts, grandparents or even people who were at one time strangers could score at this level.
   - See the definition of main caregiver and primary caregiver before making this coding.
Physical Abuse

Definitions

Physical abuse

Physical abuse is defined as physical behaviour that does or has the potential to do physical harm to the subject. This must include contact between either the perpetrator and the subject or an object and the subject, provided that the perpetrator has caused this contact. Detailed descriptions of behaviours that constitute physical abuse are given in the severity section below.

Restraint

Restraint is defined as the use of an object such as rope or flex that the perpetrator has used to prevent the subject’s free movement.

Perpetrator Criteria

Any person, no matter what relationship they have with the subject, can be a perpetrator of physical abuse.

Entering information in the correct factor

Experiencing physical abuse and child sexual abuse

Physical injuries that occur as a direct result of sexual interaction (e.g. rectal tears) are coded solely under sexual abuse. Physical coercion before, during or after the sexual interaction can be scored in the experiencing physical abuse factor, (i.e. in both factors) if the criteria for physical abuse are met. If it is unclear whether physical abuse was involved then the criterion is not met.

Experiencing physical abuse and witnessing intrafamilial physical abuse

If a family member simultaneously physically abuses both the subject and another family member then this event is to be coded as both physical abuse and witnessing intrafamilial physical abuse. If the subject accidentally becomes caught up in physical abuse involving two family members, and is himself physically abused in some way, then this too is to be recorded in both factors.
Potential problems

Marks to the body

For many of the scores there must be evidence of a mark to the body. E.g. ‘The caregiver inflicted minor marks on the child’s body during a spanking’. For some scores evidence of mark is not needed:

- Restraint of a person
- Behaviours that involve a person being hit with an instrument, or if there is a mention of him/her being ‘kicked’, ‘beaten’, ‘beaten up’, ‘thumped’ or ‘punched’
- A report that physical abuse took place or that a person is on the Child Protection Register under the category of ‘Physical Abuse’ or ‘Actual Physical Abuse’
- A person is smothered or choked
- A person is hospitalised because of a physically abusive act.

Cause of marks to the body

Note that it must be clearly stated or the coder must be able to reasonably infer that the mark was a result of the physical abuse by a perpetrator, rather than as a result of some non-abusive source.

Clearly stated
E.g. ‘The child had a black eye, which he got from being punched by the father’
Here it clearly states that the mark was a result of the physically abusive act. This factor could be scored as present on the basis of this quote, and the severity level referring to marks above the neck could also be scored.

Reasonably inferred
E.g. ‘The father was violent towards the child, and the child would often turn up at casualty with injuries.’
The proximity of the two clauses (father’s violence, hospitalisation) means that we can interpret the injuries as resulting from the father’s violence. This factor could be scored as present using this quote, and the severity level referring to hospitalisation for less than 24 hours could also be scored.

Neither clearly stated nor reasonably inferred
E.g. ‘The child had a number of bruises on his back.’
Here there is no indication that the bruises were caused by physical abuse. Therefore, this factor cannot be scored using this evidence alone. Also the severity level referring to bruising received below the neck could not be scored on the basis of this quote.

E.g. Social worker reports that the subject was physically abused by the father. In a separate document a schoolteacher reports that the subject turned up at school with a bruise on his legs. It is not clear that the bruise resulted from the father’s violence, and so the severity level referring to bruising below the neck could not be scored using this quote alone.

Physical abuse vs. Punishment

It is important to distinguish physical abuse from non-abusive physical punishment. Punishment is regarded as non-abusive if it involves appropriate setting of boundaries and does not result in marks. However, if an action referred to as punishment is deemed to be inappropriate by a social worker or a clinician, it is to be coded as physical abuse. Physical abuse can also be coded if the behaviours in the severity section below are described in the file as punishment.

Some words could be interpreted as either punishment or physical abuse: hit, slapped, spanked or smacked. For these words physical abuse can be coded only if there is evidence of marks.
If the person is hit with an instrument or if there is a mention of him/her being kicked, beaten or punched, then physical abuse can be coded without the mention of marks.

**Verbal and physical aggression**

Some words and phrases are used to refer to both physical and verbal aggression, such as aggressive/ly, forceful/ly, hostile/hostility, stormy (relationship) and turbulent (relationship). These are indicators of possible physical abuse and the quote containing the phrase should be extracted, but they are not sufficient in themselves for a coding of physical abuse.

Although words such as violence, attacked and assaulted are not completely clear about the nature of the act, they are sufficient for a coding of physical abuse. This is because in most cases they are used to refer to physical abuse (e.g. “he was violent to her”).

**Reference to a group being maltreated**

If the file refers to a group being maltreated and this group includes the subject, then it is permissible to infer that the subject has experienced this form of maltreatment.

---

**Dichotomous question: Did the subject experience physical abuse?**

Code as present if the following criteria are met:

- At least one of the experiences occurred before the subject’s 16th birthday or the date of the subject’s first perpetration, whichever took place first.

And for that or those experience/s, the following criteria are also met:

- The criteria for physical abuse are met (see definition)
- The behaviour is to be coded in this rather than another risk factor (see ‘Entering information in the correct factor’)
- If marks to the body are needed to make a coding, there is evidence of marks
- If marks to the body are needed to make a coding, the marks are a result of the perpetrator
- The behaviour is not classified as punishment
- The behaviour is not classified as verbal aggression.

If the criteria are not met, code no.
Severity levels for Experiencing Physical abuse

Severity score 1: Minor marks below the neck / no clear description

- The caregiver inflicted minor marks on the child’s body during a spanking; there were no marks to the neck or head. The term ‘minor marks’ includes redness or use of the term minor marks, but not bruising of any kind.
- The perpetrator held the child in order to restrain him, but not by the head or neck. For a definition of ‘restraint’ see the top of this section.
- Reports indicated that the caregiver had kicked, beaten, or punched the child; no other information was given.
- The child received injuries that were documented to have occurred by non-accidental means. The details of the report were not specific enough to warrant a higher rating.
- The perpetrator was reported to have spanked the child with an open hand or an object likely to inflict only minor marks in most cases (e.g. a soft belt, a ruler, a table tennis bat), with the child sustaining marks on or below the shoulder.
- This level is also to be scored if there is a report that physical abuse took place, but no other information is contained in the file.
- If the file states that the subject was on the Child Protection Register for ‘Physical Abuse’ or ‘Actual Physical Abuse’ and no other information is given, then this level is to be scored. If the file states that the subject was on the Child Protection Register for ‘Potential Physical Abuse’ or any variation of that phrase (e.g. ‘Grave Concern’ or ‘Likely’) then, in the absence of further information, this is not to be coded as physical abuse.

Examples
- Redness on the child’s bottom was reported following a spanking with the belt.

Severity score 2: Non-minor marks below neck

- The caregiver inflicted a bruise or bruises to the child’s body from an incident.
- The caregiver spanked the child with an object likely to leave a non-minor mark (e.g. a hairbrush, a belt buckle, an electric cord).

Examples
- The child sustained welts on the back after being beaten with a hairbrush.
- The child was beaten with an electric cord, resulting in numerous marks.

Severity score 3: Marks or restraint to neck or above / serious bruising, minor lacerations, minor burns

- The caregiver inflicted marks on the child’s head, face or neck (e.g. a black eye).
- The perpetrator used restraint (see definition) involving contact with the subject on the head or neck.
- The caregiver’s rough handling of the child resulted in serious bruising or minor lacerations (e.g. require stitches or minor medical attention).
- The caregiver inflicted minor burns (e.g. minor cigarette burns) to the child’s body.

Examples
- The child received a handprint on the neck after the parent grabbed him.
- The child had a black eye from being punched in the face.
- Small circular burns on the child’s hands were identified as cigarette burns.
Severity score 4: Serious injuries but not hospitalised / <24 hrs hospitalisation / asphyxiation
- The caregiver hit the child with an object (e.g. a baseball bat, a telephone) likely to result in serious injury (e.g. non-minor lacerations, second degree burns, fractures or concussion), or threw the child against the wall, but injuries that were sustained did not require hospitalisation, according to available medical information.
- The caregiver attempted to choke or smother the child, but no emergency medical care was required.
- The caregiver inflicted serious burns (second degree) to the child’s body, but the injury did not require hospitalisation.
- The caregiver inflicted an injury that required some hospital care, such as treatment in a casualty department, but did not require hospitalisation for more than 24 hours (e.g. stitches, fractures, non-minor sprain). Also score at this level if the subject was hospitalised, but the file does not state the duration of the hospitalisation.

Examples
- The child was beaten with a board that had nails in it. The child received bruises and cuts.
- The child was thrown downstairs, and fractured one arm.
- The child was severely burned by the parent and was treated in a casualty department.

Severity score 5: Twenty four hours or more hospitalisation / permanent damage
- The caregiver inflicted an injury to the child that either required hospitalisation for more than 24 hours (e.g. internal injuries) or was permanently physically damaging or disfiguring (e.g. brain damage, severe scarring).

Examples
- The child was set on fire; the burns were permanently disfiguring.
- The child was hospitalised for one week for internal injuries and evidence of a shaken infant syndrome.

Duration of physical abuse
Refer to the section in the introduction.

Duration of contact with a particular perpetrator should be included in the calculation only if it is possible to code ‘Yes’ for the Dichotomous question solely on the basis of this perpetrator.

Number of perpetrators of physical abuse
Sum the number of perpetrators of physical abuse

A particular perpetrator should be included in the calculation only if it is possible code ‘Yes’ for the Dichotomous question solely on the basis of this perpetrator.
Witnessing Intrafamilial Physical Abuse

**Definitions**

**Definition of witnessing**

Witnessing is defined as the subject being either visually or aurally exposed to, but neither a recipient nor a perpetrator of, an event of intrafamilial physical abuse. The file must clearly state or the scorer must be able to reasonably infer that the physical abuse was not directed towards or perpetrated by the subject.

Witnessing can be inferred only if at least one of the following constellations of criteria is met:

**EITHER**

Both these criteria must be met:
- It is clearly stated that the subject witnessed intrafamilial physical abuse. (E.g. ‘The subject saw the mother being punched in the stomach by the father’.)
- Perpetrator and victim criteria (see below) are met.

**OR**

All three of the criteria must be met:
- There is a general statement about a violent relationship between a victim and a perpetrator. (Examples of general statements: ‘The father used to hit the mother’; ‘The father said he saw nothing wrong with disciplining the children using a walking stick.’)
- Perpetrator and victim criteria (see below) are met
- The subject, perpetrator and victim have all lived together at some point in the subject’s life.

NOTE: If there is a clear statement or it can be reasonably inferred that the subject did not witness intrafamilial physical abuse, despite the above criteria being met, then obviously do not code it as witnessing.

**OR**

If the only evidence is statement about a specific event or series of events then more stringent criteria apply. All three criteria must be met:
- There is a specific statement about a violent event between a victim and a perpetrator (e.g. ‘On 6th June, Mrs X was taken to hospital after been beaten up by her husband)
- Perpetrator and victim criteria (see below) are met
- The subject, perpetrator and victim were living together at the time of the event.

NOTE: If there is a clear statement or it can be reasonably inferred that the subject did not witness intrafamilial physical abuse, despite the above criteria being met, then obviously do not code it as witnessing.

**OR**

Witnessing can also be inferred if the following three criteria are met:
- There is a specific statement about a violent event between a victim and a perpetrator (e.g. ‘On 6th June, Mrs X was taken to hospital after been beaten up by her husband)
- Perpetrator and victim criteria (see below) are met
- Although the subject, victim and perpetrator were not living together during the event, it is clear that at least one of the events took place in the house where the subject was living. (Note that this criteria is met even if neither the victim nor perpetrator are also resident; it is only necessary for the subject to be living there.)

NOTE: If there is a clear statement or it can be reasonably inferred that the subject did not witness intrafamilial physical abuse, despite the above criteria being met, then obviously do not code it as witnessing.
Physical abuse

Physical abuse is defined as physical behaviour that does or has the potential to do physical harm to the subject. This must include contact between either the perpetrator and the subject or an object and the subject, provided that the perpetrator has caused this contact. Detailed descriptions of behaviours that constitute physical abuse are given in the severity section below.

Restraint

Restraint is defined as the use of an object such as rope or flex that the perpetrator has used to prevent the subject’s free movement.

Perpetrator Criteria

Perpetrator and victim criteria

A violent event between two people meet the criteria for this factor if EITHER one of the victims OR one of the perpetrators is a family member. A family member is anyone who, if a perpetrator, would score at level two or above for closeness to perpetrator, except for authority figures. The following is a brief overview of people who meet the criteria; if in doubt consult the closeness to perpetrator ratings.

Anyone referred to in familial terms, including extended family: parents, siblings, grandparents, uncles, aunts and cousins. Note that it also includes foster parents and siblings. People who have cared for the subject for one month or more also meet the criteria. Partners of the mother and father are considered family members only if they have cared for the subject for a month or more.

Occasionally the physical abuse between family members will not have a clearly defined perpetrator and a clearly defined victim. (E.g. ‘The mother and father would often have extremely violent fights involving punching and kicking.’) Here both people are classed as perpetrators.

Entering information in the correct factor

Experiencing physical abuse and witnessing intrafamilial physical abuse

If a family member simultaneously physically abuses both the subject and another family member then this event is to be coded as both physical abuse and witnessing intrafamilial physical abuse. If the subject accidentally becomes caught up in physical abuse involving two family members, and is himself physically abused in some way, then this too is to be recorded in both factors.
Potential problems

Marks to the body

For many of the scores there must be evidence of a mark to the body. E.g. ‘The caregiver inflicted minor marks on the child’s body during a spanking’. For some scores evidence of mark is not needed:

- Restraint of a person
- Behaviours that involve a person being hit with an instrument, or if there is a mention of him/her being ‘kicked’, ‘beaten’, ‘beaten up’, ‘thumped’ or ‘punched’
- A report that physical abuse took place or that a person is on the Child Protection Register under the category of ‘Physical Abuse’ or ‘Actual Physical Abuse’
- A person is smothered or choked
- A person is hospitalised because of a physically abusive act.

Cause of marks to the body

Note that it must be clearly stated or the coder must be able to reasonably infer that the mark was a result of the physical abuse by a perpetrator, rather than as a result of some non-abusive source.

Clearly stated
E.g. ‘The child had a black eye, which he got from being punched by the father’
Here it clearly states that the mark was a result of the physically abusive act. This factor could be scored as present on the basis of this quote, and the severity level referring to marks above the neck could also be scored.

Reasonably inferred
E.g. ‘The father was violent towards the child, and the child would often turn up at casualty with injuries.’
The proximity of the two clauses (father’s violence, hospitalisation) means that we can interpret the injuries as resulting from the father’s violence. This factor could be scored as present using this quote, and the severity level referring to hospitalisation for less than 24 hours could also be scored.

Neither clearly stated nor reasonably inferred
E.g. ‘The child had a number of bruises on his back.’
Here there is no indication that the bruises were caused by physical abuse. Therefore, this factor cannot be scored using this evidence alone. Also the severity level referring to bruising received below the neck could not be scored on the basis of this quote.

Witnessing Intrafamilial physical abuse vs. Witnessing Punishment

It is important to distinguish physical abuse from non-abusive physical punishment. Punishment is regarded as non-abusive if it involves appropriate setting of boundaries and does not result in marks. However, if an action referred to as punishment is deemed to be inappropriate by a social worker or a clinician, it is to be coded as physical abuse. Physical abuse can also be coded if the behaviours in the severity section below are described in the file as punishment.

Some words could be interpreted as either punishment or physical abuse: hit, slapped, spanked or smacked. For these words physical abuse can be coded only if there is evidence of marks.
If the person is hit with an instrument or if there is a mention of him/her being kicked, beaten or punched, then physical abuse can be coded without the mention of marks.

**Verbal and physical aggression**

Some words and phrases are used to refer to both physical and verbal aggression, such as aggressive/ly, forceful/ly, hostile/hostility, stormy (relationship) and turbulent (relationship). These are indicators of possible physical abuse and the quote containing the phrase should be extracted, but they are not sufficient in themselves for a coding of physical abuse.

Although words such as violence, attacked and assaulted are not completely clear about the nature of the act, they are sufficient for a coding of physical abuse. This is because in most cases they are used to refer to physical abuse (e.g. “he was violent to her”).

---

**Dichotomous question: Did the subject witness intrafamilial physical abuse?**

Code as present if the following criteria are met:

- At least one of the experiences occurred before the subject’s 16th birthday or the date of the subject’s first perpetration, whichever took place first.

And for that or those experience/s, the following are also met:

- The general criteria for witnessing are met and at least one of the witnessing criteria of constellations are met.
- The criteria for physical abuse are met (see definition).
- The perpetrator and victim criteria are met.
- The behaviour is to be coded in this rather than another risk factor (see ‘Entering information in the correct factor’)
- If marks to the body are needed to make a coding, there is evidence of marks
- If marks to the body are needed to make a coding, the marks are a result of the perpetrator
- The behaviour is not classified as punishment
- The behaviour is not classified as verbal aggression.

If the criteria are not met, code no.
Severity Levels for Witnessing Intrafamilial Physical Abuse

Severity score 1: Minor marks below the neck / no clear description

- The perpetrator inflicted minor marks on the victim's body during a spanking; there were no marks to the neck or head. The term 'minor marks' includes redness or use of the term minor marks, but not bruising of any kind.
- The perpetrator held the victim in order to restrain him/her, but not by the head or neck. For a definition of 'restraint' see the top of this section.
- Reports indicated that the perpetrator had kicked, beaten, or punched the victim; no other information was given.
- The victim received injuries that were documented to have occurred by non-accidental means. The details of the report were not specific enough to warrant a higher rating.
- The perpetrator was reported to have spanked the victim with an open hand or an object likely to inflict only minor marks in most cases (e.g. a soft belt, a ruler, a table tennis bat), with the victim sustaining marks on or below the shoulder.
- This level is also to be scored if there is a report that physical abuse took place between family members, but no other information is contained in the file.

Examples
- Redness on the victim's bottom was reported following a spanking with the belt.

Severity score 2: Non-minor marks below neck

- The perpetrator inflicted a bruise or bruises to the victim's body from an incident.
- The perpetrator spanked the victim with an object likely to leave a non-minor mark (e.g. a hair brush, a belt buckle, an electric cord).

Examples
- The victim sustained welts on the back after being beaten with a hairbrush.
- The victim was beaten with an electric cord, resulting in numerous marks.

Severity score 3: Marks or restraint to neck or above / serious bruising, minor lacerations, minor burns

- The perpetrator inflicted marks on the victim's head, face or neck (e.g. a black eye).
- The perpetrator used restraint (see definition) involving contact with the subject on the head or neck.
- The perpetrator's rough handling of the victim resulted in serious bruising or minor lacerations (e.g. require stitches or minor medical attention).
- The perpetrator inflicted minor burns (e.g. minor cigarette burns) to the victim's body.

Examples
- The victim received a handprint on the neck after the perpetrator grabbed her.
- The victim had a black eye from being punched in the face.
- Small circular burns on the victim's hands were identified as cigarette burns.

Severity score 4: Serious injuries but not hospitalised / <24 hrs hospitalisation / asphyxiation

- The perpetrator hit the victim with an object (e.g. a baseball bat, a telephone) likely to result in serious injury (e.g. non-minor lacerations, second degree burns, fractures or concussion), or threw the victim against the wall, but injuries that were sustained did not require hospitalisation, according to available medical information.
- The perpetrator attempted to choke or smother the victim, but no emergency medical care was required.
- The perpetrator inflicted serious burns (second degree) to the victim's body, but the injury did not require hospitalisation.
- The perpetrator inflicted an injury that required some hospital care, such as treatment in a casualty department, but did not require hospitalisation for more than 24 hours (e.g. stitches, fractures, non-minor sprain). Also score at this level if the subject was hospitalised, but the file does not state the duration of the hospitalisation.

**Examples**
- The victim was beaten with a board that had nails in it. The victim received bruises and cuts.
- The victim was thrown downstairs, and fractured one arm.
- The victim was severely burned by the perpetrator and was treated in a casualty department.

**Severity score 5: Twenty four hours or more hospitalisation / permanent damage**
- The perpetrator inflicted an injury to the victim that either required hospitalisation for more than 24 hours (e.g. internal injuries) or was permanently physically damaging or disfiguring (e.g. brain damage, severe scarring).

**Examples**
- The victim was set on fire; the burns were permanently disfiguring.
- The victim was hospitalised for one week for internal injuries and evidence of a shaken infant syndrome.

---

**Duration of witnessing intrafamilial physical abuse**

Refer to the section in the introduction.

Duration of contact with a particular perpetrator should be included in the calculation only if it is possible to code 'Yes' for the Dichotomous question solely on the basis of this perpetrator.

---

**Number of perpetrators of witnessing intrafamilial physical abuse**

Sum the number of perpetrators of witnessing intrafamilial physical abuse.

A particular perpetrator should be included in the calculation only if it is possible code 'Yes' for the Dichotomous question solely on the basis of this perpetrator.

---

**Was a mother figure a victim of the intrafamilial physical abuse?**

For this section, mother is defined as a female who meets the criteria for primary carer (i.e., has cared for the subject for a period of one year or more). This section is only to be coded for people who meet this criterion.

Code as present if at least one of the victims of the intrafamilial physical abuse was a mother figure.
Neglect (Failure to Provide)

Definitions

Definition of neglect (failure to provide)

Physical neglect (failure to provide) is coded when a caregiver (see definitions of main and primary caregiver) fails to exercise a minimum degree of care in meeting the child's physical needs. When families are below the poverty level, physical neglect is scored if children's physical needs are not met because the parents fail to access available community resources for the well being of their children. For example, parents are unable to provide food for their children; however, they have not taken the necessary steps to seek sources of emergency sustenance.

Neglect (failure to provide) covers 5 domains:
- Supplying the child with adequate food
- Ensuring that the child has clothing that is sanitary, appropriate for the weather and permits the child freedom of movement
- Providing adequate shelter
- Ensuring adequate medical, dental and mental health care
- Ensuring the child’s adequate hygiene.

Restraint

Restraint is defined as the use of an object such as rope or flex that the perpetrator has used to prevent the subject’s free movement.

Perpetrator criteria

The perpetrator of neglect has to be the main caregivers for the subject at the time of the experience of neglect. See the definition of main caregiver given in the introduction.

Entering information in the correct factor

Neglect failure to provide and neglect lack of supervision

An event may be coded for either or both forms of neglect as well as another risk factor. Regardless of the other form of maltreatment, neglect (failure to provide) concerns the lack of essential necessities: food, clothing, shelter, health care and hygiene; whereas neglect (lack of supervision) concerns the inadequate supervision of a child in either a safe or an unsafe environment.
Potential problems

Vague reference to neglect

If the file refers to the neglect of the subject and no other information is given, then this quote should be coded in neglect (failure to provide); it is not to be coded under neglect (lack of supervision). The same principle applies if the subject is on the Child Protection Register under ‘Neglect’ or ‘Actual Neglect’. If any other form of wording is used in reference to the Child Protection Register, then the criteria for neglect are not met.

Neglect by Proxy

The coder should not infer that the subject is being neglected solely because one or more of his siblings have experienced neglect. However, if in doubt, the trawler should record the item that describes the neglect of the sibling and let the coder decide about the neglect of the subject.

Reference to a group being maltreated

If the file refers to a group being maltreated and this group includes the subject, then it is permissible to infer that the subject has experienced this form of maltreatment.

Illness and neglect (failure to provide)

Illness in itself does not indicate neglect (failure to provide). Code only if it is clearly stated or it can be reasonably inferred that the illness resulted from a failure to provide by the caregivers or a failure to follow a course of treatment.

Dichotomous question: Did the subject experience neglect (failure to provide)?

Tick yes if the following criteria are met:
• At least one of the experiences occurred before the subject’s 16th birthday or the date of the subject’s first perpetration, whichever took place first. And for that or those experience/s, the following are also met:
  • The experience meets the criteria described in the definition of neglect (failure to provide)
  • The perpetrator criteria are met. Note that for neglect (failure to provide), the perpetrator is by definition the main caregiver; it is impossible for it to be anyone else
  • The behaviour is to be coded in this rather than another risk factor (see ‘Entering information in the correct factor’).

If the criteria are not met, tick no.
Severity Levels for neglect (failure to provide)

Severity score 1

- The caregiver does not ensure that food is available for occasional meals. The subject (less than age 10) often has to fix his own supper and/or occasionally misses meals because of parental negligence. Also code this severity level if the following words are used to describe the number of meals missed: once in a while, at times, now and then, now and again, every now and again, and from time to time. To score at this level, it must be clear or inferable that it is the odd meal that has been missed - that this is a rarity.

- The caregiver fails to provide clothing for the subject that is adequately clean and that allows freedom of movement (e.g. the clothing is so small that it restricts movement or so large that the subject often trips or has difficulty keeping the clothing on).

- The caregiver does not attempt to clean the house. Garbage has not been removed, dirty dishes are encrusted with food, and floors and other surfaces are very dirty. An unpleasant odour from garbage and debris permeates the living quarters.

- The caregiver has missed several of the subject’s medical or dental appointments, and often fails to take the subject to the doctor or dentist for ‘check-ups’ or ‘well baby’ appointments. The caregiver does not ensure that the subject is taken to the health clinic for adequate immunisation, and medical personnel have expressed concern.

- The caregiver does not attempt to keep the subject clean. The caregiver bathes the subject or washes the subject very infrequently. The subject brushes teeth only infrequently or not at all, and signs of tooth decay or discolouration are evident.

- This level is also to be scored if there is a report that neglect took place, but no other information is contained in the file. Note that neglect (lack of supervision) should not be coded as well solely using this information.

- If the file states that the subject was on the Child Protection Register for ‘Neglect’ or ‘Actual Neglect’ and no other information is given, then this level is to be scored. If the file states that the subject was on the Child Protection Register for ‘Potential Neglect’ or any variation on that phrase (e.g. ‘Grave Concern’ or ‘Likely’) then, in the absence of further information, this is not to be coded as neglect (failure to provide).

Examples

- A 9-year-old subject often fixes dinner for himself because the caregivers are sleeping.

- The subject always wears clothing that is so small it restricts movement.

- The caregiver has failed to sign papers for evaluation of a behaviour problem that has been reported at school.

- The subject is dirty and frequently scratches matted hair.

Severity score 2

- The caregiver sometimes fails to ensure that food is available. This can be assessed by the number of meals missed or the general words used. Thus, score ‘2’ if: two or more meals are missed per week; the caregiver does not feed the subject for 24 hours; meals are described as being missed often, recurrently, repeatedly, or time and again. To score at this level, it must be clear or inferable that it is neither the odd meal that has been missed nor the ‘norm’ for meals to be missed.

- The caregiver does not dress the subject in clothing that is appropriate for the weather (e.g. lightweight clothing during the winter).

- The caregiver is aware that the house is infested with roaches or other vermin and has not attempted to improve the conditions.

- The caregiver does not ensure adequate sleeping arrangements for the subject (e.g. there are no beds or mattresses, or the mattresses are filthy and sodden with urine or other substance likely to promote the growth of mould or mildew.)
• The caregiver seeks medical attention but does not follow-through consistently with medical recommendations for a minor illness or infection (e.g., prescribed medicine is not administered for a mild infection, chronic head lice is not treated).
• The caregiver does not change the infant's nappies frequently, often leaving the soiled nappy unchanged for several hours, resulting in nappy rash.

Examples
• A subject has walked to school several consecutive days wearing only a thin jacket without hat or gloves, throughout winter or during cold weather.
• A social worker has visited the home several times when no food has been available. The subject reports that he does not have lunch or dinner two or three times per week.
• The subject has been diagnosed with an ear infection, but the parent does not follow through with administration of the prescribed antibiotic.

Severity score 3
• The caregiver does not provide meals on a usual basis, thereby perpetuating a pattern of regularly missed meals: four or more periods of at least two consecutive meals per week are unavailable to the subject. Also code the severity as ‘3’ if the meals are described as being missed all the time, continually, usually, regularly, perpetually, generally, more often than not. To score at this level, it must be clear or inferable that it is the ‘norm’ for meals to be missed.
• The caregiver fails to make adequate provisions for shelter for the family. For example, the caregiver does not acquire or maintain public assistance, resulting in a loss of residence or loss of financial assistance for seven days or more.
• The caregiver does not seek or follow through with medical treatment for moderately severe medical problems (e.g., the caregiver does not follow preventative measures for a chronic heart condition, or moderately elevated blood lead levels are left untreated), or the caregiver administers medical treatment that is inappropriate without consulting a doctor (e.g., caregiver gives the subject mild sedatives to control the subject, without a doctor’s consultation).
• The caregiver maintains an unsanitary living situation, where spoiled food or garbage are frequently present and where rat or vermin infestation is untreated.
• The expectant mother jeopardises the health of her unborn subject by using alcohol or drugs during pregnancy, but no foetal or alcohol or drug symptoms are evident.

Examples
• The subject is regularly not fed. He has missed two consecutive meals an average of four times a week for the last several months.
• The family has been evicted because the parent did not take appropriate actions to maintain public assistance and made no other arrangement for making rent payments. The family had no stable living arrangements for 2 weeks.
• The parent has been drunk several times during pregnancy.
• The subject has come to school with an infected cut. Despite notes from the school nurse recommending medical attention, the cut continues to be untreated.

Severity score 4
• The caregiver has made no arrangements for adequate shelter (e.g., the caregiver has not sought heat during the winters; the family is living in a car because no alternative housing was sought). The conditions continue for prolonged periods.
• The caregiver maintains the home environment such that living conditions are extremely unhealthy (e.g., faeces and urine are present in the living areas).
• The caregiver does not seek or comply with medical treatment for potentially life threatening illness or injury (e.g., the subject is not taken to the Emergency Room for severe bleeding, third degree burn, fractured skull).
• The caregiver has provided such poor nourishment that the subject fails to gain weight or grow at the rate expected for their development. The failure to grow as expected is not due to any identifiable organic failure. In addition, the failure to grow must be stated by a health care worker as being caused specifically by the carer’s neglect.

**Examples**
- The subject lives in an unheated home because the parents have failed to ensure that heating was available. During the winter, the subject came to school with frostbite.
- The subject was hit by a car, receiving a fracture and severe cuts and bruises. The subject came to school complaining of pain and stated that the parents would not take him to the hospital.

**Severity score 5**
- The caregiver has provided such poor nourishment or care to the subject that physical consequences have ensued such as weight loss in an infant, severe malnutrition, or severe non-organic failure-to-thrive.
- The caregiver has abused alcohol or drugs during pregnancy to the extent that the infant is born with foetal abnormality syndrome or a congenital drug addiction.
- The caregiver provided such gross inattention to the subject’s medical needs that the subject was permanently disabled as a result of lack of medical attention (e.g. severe starvation, dehydration).
- The caregiver does not seek professional help for the subject’s life threatening emotional problems (e.g. suicidal or homicidal attempts).

**Examples**
- At birth the subject is addicted to heroin.
- The subject is diagnosed as being severely malnourished.

---

**Duration of neglect (failure to provide)**

Refer to the section in the introduction.

Duration of contact with a particular perpetrator should be included in the calculation only if it is possible to code ‘Yes’ for the Dichotomous question solely on the basis of this perpetrator.

---

**Number of perpetrators of neglect (failure to provide)**

Sum the number of perpetrators of witnessing intrafamilial physical abuse.

A particular perpetrator should be included in the calculation only if it is possible code ‘Yes’ for the Dichotomous question solely on the basis of this perpetrator.
Neglect (lack of supervision)

**Definitions**

**Lack of supervision**

Within this system, lack of supervision is coded when a main caregiver (see definition in introduction) does not take adequate precautions to ensure a child's safety in and out of the home, given the child's particular emotional and developmental needs. The parent's failure to ensure the child's safety may include both permitting the child to be exposed to dangerous situations (e.g. allowing the child to play in an unsafe area, permitting the child to accompany someone with a known history of violent acts) as well as failing to take adequate precautions to evaluate the conditions pertaining to the child's safety (e.g. neglecting to screen the background or competency of alternative caregivers, failing to ascertain the child's whereabouts).

There are four broad elements that caregivers may violate to jeopardise children's physical safety:

1: **Supervision**: failing to take steps to ensure that the child is engaging in safe activities.

According to this dimension, as the number of hours that the child is unsupervised increases so does the potential harm. Therefore severity scores for lack of supervision increase with more prolonged periods of inadequate supervision. To assist coders in making distinctions about the relative seriousness of particular instances of lack of supervision, we have provided approximate durations of inadequate supervision that are there to serve as guidelines rather than firm criteria. We recognise that these cut-off points are somewhat arbitrary and that exact times are frequently unavailable in the records; however, we felt that establishing ranges of time was necessary to clarify coding decisions and, thus, reliability among coders.

2: **Environment**: failing to ensure that the child is playing in a safe area.

This dimension is distinguished from lack of hygiene or medically unhealthy conditions of the living environment covered under failure to provide. In the case of lack of supervision, environment refers to immediate physical dangers inside or outside the home such as broken glass, unguarded electrical fixtures, toxic chemicals and firearms.

3: **Substitute care**: failing to provide for adequate substitute care in the caregiver's absence or mental or physical incapacity.

In this respect, lack of substitute care includes situations when auxiliary supervision is not obtained, when parents do not ensure that substitute caregivers are able to adequately supervise the child, when caregivers are unable to adequately monitor the child's safety because the caregivers are intoxicated with alcohol or drugs, or when caregivers have a severe psychiatric condition that makes appropriate supervision of children highly unlikely (e.g. caregiver has delusions or hallucinations).

4: **Developmental needs**: failing to recognise the developmental needs of the child in providing adequate supervision to ensure the child's safety.

Because, in general, the consequences of failing to supervise younger children are potentially more serious, the influence of the child's developmental level should be considered when making decisions about the severity of parental failure to provide adequate supervision.

Additionally children who have a history of dangerous, impulsive or immature behaviour require more intensive supervision, and may be given a higher supervision rating if they are unsupervised. For example, an adolescent who is known to exhibit poor judgement and to engage in impulsive and
destructive behaviour would require more supervision than most children of the same age. It is difficult to quantify the amount of supervision that is required at each developmental level. The examples provided give some guidelines of relative severity, but the information available for each case must be considered with regard to the age and particular needs of each child.

In summary, when making individual decisions about severity, the coder should take into account the length of time the child was left unsupervised, the amount of danger present in the physical environment, the physical environment, the adequacy of potential substitute caregivers, and the developmental needs of the child.

### Perpetrator criteria

The perpetrator of neglect has to be the main caregivers for the subject at the time of the experience of neglect. See the definition of main caregiver given above.

### Entering information in the correct factor

**Neglect failure to provide and neglect lack of supervision**

An event may be coded for either or both forms of neglect as well as another risk factor. Regardless of the other form of maltreatment, neglect (failure to provide) concerns the lack of essential necessities: food, clothing, shelter, health care and hygiene; whereas neglect (lack of supervision) concerns the inadequate supervision of a child in either a safe or an unsafe environment.
Potential problems

Vague reference to neglect

If the file refers to the neglect of the subject and no other information is given, then this quote should be coded in neglect (failure to provide); it is not to be coded under neglect (lack of supervision). The same principle applies if the subject is on the Child Protection Register under ‘Neglect’ or ‘Actual Neglect’. If any other form of wording is used in reference to the Child Protection Register, then the criteria for neglect are not met.

Neglect by Proxy

The coder should not infer that the subject is being neglected solely because one or more of his siblings have experienced neglect. However, if in doubt, the trawler should record the item that describes the neglect of the sibling and let the coder decide about the neglect of the subject.

Reference to a group being maltreated

If the file refers to a group being maltreated and this group includes the subject, then it is permissible to infer that the subject has experienced this form of maltreatment.

Dichotomous question: Did the subject experience neglect (lack of supervision)?

Tick yes if the following criteria are met:

- At least one of the experiences occurred before the subject’s 16th birthday or the date of the subject’s first perpetration, whichever took place first.
  And for that or those experience/s, the following are also met:
  - The experience meets the criteria described in the definition of neglect (lack of supervision)
  - The perpetrator criteria are met. Note that for neglect (lack of supervision), the perpetrator is by definition the main caregiver; it is impossible for it to be anyone else.
  - The behaviour is to be coded in this rather than another risk factor (see ‘Entering information in the correct factor’).

If the criteria are not met, tick no.
Severity levels for neglect (lack of supervision)

Severity score 1

- The caregiver fails to provide adequate supervision or arrange for alternative supervision for short periods of time (e.g. up to a morning, an afternoon or an evening i.e. approximately less than three hours) with no immediate source of danger in the environment.
- If the file refers to the neglect of the subject and no other information is given, then this quote should be coded in neglect (failure to provide) not neglect (lack of supervision).
- If the file states that the subject was on the Child Protection Register for 'Neglect' or 'Actual Neglect' and no other information is given, then this is to be scored under neglect (failure to provide); it is not to be coded under neglect (lack of supervision). If the file states that the subject was on the Child Protection Register for 'Potential Neglect' or any variation on that phrase (e.g. 'Grave Concern' or ' Likely') then, in the absence of further information, this is not to be coded under either neglect (failure to provide) or neglect (lack of supervision).

Examples
- An eight year old is left alone during the day for a few hours.
- Pre-schoolers play outside unsupervised or are left in the care of an 8 year old. (In this case, the pre-schoolers who are unsupervised in an environment with a few hazards reported would receive a score of 1. Similarly supervision of pre-schoolers by a slightly older subject would represent mildly inadequate alternate supervision, which would also be coded 1.
- Subject is left in the care of questionably suitable baby-sitters (e.g. pre-adolescent, mildly impaired older person).

Severity score 2

- The caregiver fails to provide supervision or arrange for adequate alternative supervision, or provides poor supervision for a 'working day' (i.e. approximately three to eight hours), with no immediate source of danger in the environment.
- The caregiver fails to provide supervision for short periods of time (e.g. up to a morning, an afternoon or an evening i.e. approximately less than three hours) when the subject is in an unsafe play area.
- The subject receives inadequate supervision despite a history of problematic behaviour (e.g. impulse behaviour, hyperactivity).

Examples
- The subject is frequently left alone for the day, without a responsible caregiver available.
- An infant is left in the care of an 8 year old for several hours (in this case the infant is given a code of 2. The eight year old would be given a code of 1, similar to the example under level 1).
- A six year old is locked out of the home alone, and the caregiver does not return until the evening.
- The subject is allowed to play in an unsafe play area (e.g. broken glass present, old basement garage cluttered with toxic chemicals, power tools or an old refrigerator) unsupervised.
- Subject gets into trouble with neighbours because of lack of supervision.

Severity score 3

- The caregiver fails to provide adequate supervision for a 'whole day' (i.e. the time between when the subject gets up in the morning to when he goes to bed for the night) or a 'whole night' (i.e. the time between when the subject goes to bed for the night and when he gets up in the morning). If it is not clear whether the subject has been left for a 'working day' or a 'whole day', rate the severity as a 2.
• The caregiver allows the subject to play in an unsafe play area for up to a 'working day' (approximately 3 to 8 hours).

Examples
• The subject is left unsupervised for the night.
• The subject often sees no one for the whole day, other than his younger siblings.
• The subject is left in the care of an unreliable caregiver (e.g. who is known to drink, or is extremely inattentive) for the whole day.

Severity score 4
• The caregiver provides no supervision for more than a whole day/night, but less than 24 hours. If it is not clear from the text whether the subject has been left for a 'working day', a 'whole day', or between a whole day and 24 hours, rate the severity as a '2', being the most conservative option.
• The caregiver allows the subject to play in an area that is very dangerous (i.e. high probability that the subject will be hit by a car or fall out of a window, get burned or drowned).
• A subject with a known history of destructive or dangerous acts (e.g. fire-setting, suicidal ideation) is left unsupervised.

Examples
• A carer goes to work one morning, leaving his school-age son alone, only to return drunk at three the next morning.
• The subject is allowed to play by highway, or on the roof of a condemned building.
• The subject is allowed to go with a caregiver who has known history of violence or who has a restraining order prohibiting contact with the subject.

Severity score 5
• The caregiver fails to provide adequate supervision for 24 hours or more.
• The caregiver places the subject in a life-threatening situation or does not take steps to prevent the subject from being in a life-threatening situation.

Examples
• A pre-school subject is left alone for 24 hours.
• The subject is kicked out of the home with no alternative living arrangements.
• The caregiver keeps loaded firearms in a location that is accessible to the subject.
• A toddler plays near a swimming pool unsupervised.
Duration of neglect (lack of supervision)

Refer to the section in the introduction.

Duration of contact with a particular perpetrator should be included in the calculation only if it is possible to code 'Yes' for the Dichotomous question solely on the basis of this perpetrator.

Number of perpetrators of neglect (lack of supervision)

Sum the number of perpetrators of witnessing intrafamilial physical abuse.

A particular perpetrator should be included in the calculation only if it is possible code 'Yes' for the Dichotomous question solely on the basis of this perpetrator.
Rejection by carers

**Definitions**

This risk factor is often underlined by the carer’s dislike of or hostility towards the child. This dislike of or hostility towards the child may be based on the child being seen as bearing the same negative characteristics as a disliked or hated person. Another context in which this subtype may exist is where the child is deemed to be responsible for a misfortune that has befallen the carer, such as illness associated with childbirth or loss of employment. It may also be the case that, for some parents, the child was, and is, actually unwanted, and that feelings of dislike and hostility ensue from the carer’s obligation to look after the child.

The dislike may be manifest behaviourally in several ways:

- **Blaming:** including references to blaming the subject, ‘scapegoating’, ‘targeting’ and negative misattribution
- **The subject is perceived as deserving harsh discipline and punishment**
- **Hostility:** expression of hostility towards the subject
- **Degradation:** including references to degrading the subject, criticising the subject
- **Rejection:** including references to rejection, hostility towards the subject, threats to remove the subject from the home or kill the subject, preference for another person, disliked because resembles another person
- **Locking up or containing the child**

**Examples**

**Blaming**

‘Both the father and mother blamed the child for the break-up of their marriage.’
‘The mother blames her son [subject] for the suicide of her boyfriend.’
‘The child is very much scapegoated within the family.’

**Hostility**

‘The father shouts and screams at the child.’
‘The mother has told the child that she dislikes him.’
‘The father states that he wishes the child was dead’

**Degradation**

‘The father belittles the child, referring to him as a wimp and a loser.’

**Rejection**

‘The child was rejected from the anti-natal stage’

**Threats to have the subject removed from the home or kill the subject**

‘The parents have told him [the subject] that they are going to get rid of him and then have a new baby to replace him.’
Stepfather ‘has requested him to be accommodated. There have been two visits on an emergency basis when the stepfather has threatened to kill the child if he is not removed’.

**Preference for another person**

‘The mother made it clear that if it came down to a choice between her son [subject] and her boyfriend, she would choose the boyfriend.’
The father made it clear that he much preferred his natural son to his adoptive son [subject].

Disliked because resembles another person

Mother told the social worker 'how she wished she never had the child, how she wished she had aborted, and would like him taken away - she said he was too much like his father. She threatens the child with the social worker taking him away if he is not good'.

Perpetrator criteria

The perpetrator of rejection by carers has to be the main caregiver for the subject prior to or at the time of the experience of rejection. See the definition of main caregiver given in the introduction.

Potential problems

Feelings and attitudes

Occasionally there is a statement of the carer's feelings or attitudes to the subject, but no description of a behaviour (e.g. the father has admitted to hating the subject.) In these situations it is permissible to code rejection by carers solely on the basis of a feeling or attitude; the assumption being that the subject is likely to be aware of these feelings. Also note that it is permissible to make this coding whether or not it is stated that the subject heard or was aware of the feeling or attitude.

Single instance of the behaviour

Rejection by carers is conceptually defined as interactional patterns that are a pervasive characteristic of a parent-child relationship. Often in the files there is a description of a single instance of a behaviour that, if it were a pervasive characteristic of the relationship, would be rejection by carers. It is permissible, in the absence of evidence to the contrary, to assume that this single instant was representative of a pervasive feature of the relationship, and therefore to code the rejection by carers on that basis. If it is stated or can be reasonably inferred that the single instant was not representative of the relationship, then do not code rejection by carers.

Dichotomous question: Did the subject experience Rejection by carers?

Code as 'present' if the following criteria are met:
- At least one of the experiences occurred before the subject's 16th birthday or the date of the subject's first perpetration, whichever took place first.

And for that or those experience/s, the following are also met:
- The experience meets the criteria described in the definition of rejection by carers
- The perpetrator criteria are met. Note that for rejection by carers the perpetrator is by definition the main caregiver; it is impossible for it to be anyone else.

If the criteria are not met, code 'no'.

409
Duration of Rejection by carers

Refer to the section in the introduction.

Duration of contact with a particular perpetrator should be included in the calculation only if it is possible to code 'Yes' for the Dichotomous question solely on the basis of this perpetrator.

Number of perpetrators of Rejection by carers

Sum the number of perpetrators of witnessing intrafamilial physical abuse.

A particular perpetrator should be included in the calculation only if it is possible code ‘Yes’ for the Dichotomous question solely on the basis of this perpetrator.
Appendix 6

Genetic theories
There is a sizeable literature on the role of genetic factors in aggression and criminal behaviour more generally (e.g., Ebstein & Belmaker, 2002), but as yet there has been no application of this research to sexual offender populations. One interesting possibility is the idea that genetic factors may play a moderating role (Baron & Kenny, 1986) in the relationship between environmental risk factors and psychopathological outcomes. Caspi et al. (2002), for example, found that a functional polymorphism in the gene encoding monoamine oxidase A (MAOA) moderated the relationship between the experience of maltreatment and subsequent aggressive behaviour in a large sample (N = 1037) of male children. For those children who had not experienced maltreatment, those high or low in MAOA activity had comparable levels of aggressive behaviour. For those children who had experienced severe maltreatment, those with low MAOA activity were more likely to show aggressive behaviour than those with high MAOA activity. This may explain why not all people exposed to environment risk factors, such as childhood maltreatment, subsequently develop adverse outcomes (Curtis & Cicchetti, 2003). It is a reasonable hypothesis – though untested as yet – that genetic factors may serve to moderate the relationship between sexual offending and the environmental risks discussed in this chapters 5 and 6.
Appendix 7

Correlation matrices
Table A1 gives the odds ratios and 95% CIs for the relationship between dichotomous risk factors. The association (Kendall’s tau-b) between the more than dichotomous risk factors is given in table A2.

### Table A1 - Odds ratios and 95% CI for relationship between dichotomous risk factors

<table>
<thead>
<tr>
<th></th>
<th>Sexual abuse by a female</th>
<th>Physical abuse</th>
<th>Witnessing physical abuse</th>
<th>Neglect (failure to provide)</th>
<th>Neglect (lack of supervision)</th>
<th>Emotional maltreatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discontinuity of care</td>
<td>2.27 (0.27-19.2)</td>
<td>1.58 (0.40-6.27)</td>
<td>6.89 (1.35-35.1)</td>
<td>1.96 (0.46-8.29)</td>
<td>7.5 (0.90-62.41)</td>
<td>1.73 (0.41-7.31)</td>
</tr>
<tr>
<td>Sexual abuse by a female</td>
<td>-</td>
<td>2.08 (0.69-6.16)</td>
<td>2.41 (0.81-7.16)</td>
<td>2.24 (0.83-5.91)</td>
<td>2.04 (0.78-5.31)</td>
<td>1.01 (0.39-2.61)</td>
</tr>
<tr>
<td>Physical abuse</td>
<td>-</td>
<td>-</td>
<td>7.07 (2.89-17.37)</td>
<td>4.56 (1.89-11.17)</td>
<td>2.56 (1.09-6.01)</td>
<td>2.12 (0.92-4.92)</td>
</tr>
<tr>
<td>Witnessing abuse</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.87 (0.83-4.19)</td>
<td>2.63 (1.14-6.06)</td>
<td>1.31 (0.59-2.92)</td>
</tr>
<tr>
<td>Neglect (failure to provide)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.76 (1.24-6.13)</td>
<td>1.07 (0.49-2.31)</td>
</tr>
<tr>
<td>Neglect (lack of supervision)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.81 (0.83-3.95)</td>
</tr>
</tbody>
</table>

Although a number of the associations were statistically significant, the majority of these were in the small to moderate range. Three sets of associations were large (Kendall’s tau-b > 0.60): the severity and duration measures of witnessing intrafamilial violence (Kendall’s tau-b = 0.62, p < 0.001), the severity and duration measures of neglect (failure to provide) (Kendall’s tau-b = 0.73, p < 0.001), and the severity and duration measures of neglect (lack of supervision) (Kendall’s tau-b = 0.78, p < 0.001). The strength of these associations may have caused multicolinearity problems if both variables were entered into a logistic regression, but, as described in the method chapter, an a priori decision was made to use only severity, because of concerns about the validity of the duration measures. The size of the association between variables that were entered into the regression is unlikely to lead to multicolinearity, and the results of the two regressions suggested that this was not a problem. Multicolinearity is typically indicated by unexpectedly large adjusted odds ratios and unexpectedly wide confidence intervals for those odds.
ratios. In both regressions the adjusted odds ratios and confidence intervals appeared reasonable given the unadjusted odds ratios and the sample size.

Table A2 – Association between more than dichotomous risk factors (Kendall’s tau-b)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Disc care</td>
<td>0.03</td>
<td>0.03</td>
<td>0.31***</td>
<td>0.29**</td>
<td>0.29**</td>
<td>0.21**</td>
<td>0.13</td>
<td>0.01</td>
<td>0.30***</td>
<td>0.21**</td>
</tr>
<tr>
<td>Sexual abuse (sev)</td>
<td>-</td>
<td>0.13</td>
<td>0.01</td>
<td>-0.06</td>
<td>0.07</td>
<td>0.10</td>
<td>0.01</td>
<td>0.03</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td>Sexual abuse (dur)</td>
<td>-</td>
<td>-</td>
<td>0.02</td>
<td>0.17*</td>
<td>0.10</td>
<td>0.15*</td>
<td>0.08</td>
<td>0.03</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>Phys abuse (sev)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.55***</td>
<td>0.35***</td>
<td>0.18*</td>
<td>0.27</td>
<td>0.21</td>
<td>0.21</td>
<td>0.13</td>
</tr>
<tr>
<td>Phys abuse (dur)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.41***</td>
<td>0.39***</td>
<td>0.28</td>
<td>0.33</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>Wit phys (sev)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.62***</td>
<td>0.20</td>
<td>0.16</td>
<td>0.28</td>
<td>0.19</td>
</tr>
<tr>
<td>Wit Phys (dur)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.12</td>
<td>0.16</td>
<td>0.26</td>
<td>0.29***</td>
<td>0.29</td>
</tr>
<tr>
<td>Neglect (fp) (sev)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.73***</td>
<td>0.30</td>
<td>0.31***</td>
<td></td>
</tr>
<tr>
<td>Neglect (fp) (dur)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.23**</td>
<td>0.33***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neglect (ls) (sev)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.78***</td>
<td></td>
</tr>
<tr>
<td>Neglect (ls) (dur)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.05, ** p < 0.01, *** p < 0.001

Abbreviations: dur = duration, fp = failure to provide, ls = lack of supervision, phys = physical abuse, sev = severity, wit = witnessing
Appendix 8

Limitations of regression techniques
Regression techniques are a mainstay of non-experimental psychology, including sexual offending research. It is not hard to see why. If researchers want to understand the development of sexual offending, they must study it as it occurs in the real world, but the real world does not use an experimental design. Pre-existing groups (e.g., those victims who subsequently perpetrator and those who do not) are likely to differ on potential confounds and, as a further confusion, putative risks are likely to overlap. Regression offers the researcher a beguiling possibility: the association between variables can be statistically 'adjusted' for these confounds; we can then establish the unique or independent contribution of a risk factor to an outcome. The use of these techniques is, however, controversial and there remain questions about their interpretability (e.g. Briere, 1988; 1992; Briere & Elliot, 1993; Cook & Campbell, 1979; Lord, 1969; Tabachnick & Fidell, 1989). The analysis of the data reported in this thesis relied in part on these techniques, and so it is worth discussing these limitations.

The reliance on statistical regression in sexual offending research suggests an unfortunate paradox: the times we need our statistical armamentarium the most are precisely those times it is likely to let us down. A statistical 'equivalent maker' would be most useful in the softer regions of psychology, which typically rely on non-experimental designs, but it is in these regions that the assumptions of statistical tests are likely to be heavily violated. The harder areas of psychology are characterised by a virtuous cycle in which strong theories and strong methodologies feed off each other. In the softer areas, strong methodologies are ruled out for ethical reasons and so theory testing never really gets off the ground. As discussed in the chapter on theories of perpetration, sexual offending research is a notable example of this problem.

The interpretability of statistical regression depends on the accurate measurement of each variable and the identification of relevant confounds (Cook & Campbell, 1979). Strong theories can guide our decisions about what to measure and how to measure it; they can also give us clues about what confounds may be important.
Weak theories are not of help here, and this can lead to problems. It is easy to envisage data in which a variable that is not a risk factor appears to be so because a relevant confound has not been identified or has been identified but poorly measured. Alternatively, a genuine risk factor, once it has been statistically controlled for, may appear to make no contribution to the outcome, but only because it was not measured with any accuracy. There are yet further difficulties. The standard threats to statistical conclusion validity (Cook & Campbell, 1979) are likely to be more pervasive in the softer than the harder areas of psychology.

The research reported here provides example of each of these problems. Ethical problems rule out both experimental and longitudinal designs, but the use of a catch-up design meant that risk factors, confounds and outcomes could not be measured with accuracy. A lack of theory also limits our knowledge about what confounds it may be important to control for. The catch-up design also led to the exclusion of a sizeable proportion of the initial sample, which limits the study's statistical power. Perhaps the most telling indication of these problems is that the results changed depending on which measurement system was used. In the regression using dichotomous risk factors, only neglect (failure to provide) predicted sexual perpetration, but when ordinal measures were used witnessing intrafamilial violence (severity) was the only significant predictor.

In these circumstances, the results may tell us more about measurement error and the vagaries of statistical power in a particular study than they do about the relationship between risk factors and an outcome. As a consequence, findings fail to replicate across studies and progress in the research area does not occur.

There are some potential solutions to these difficulties, but they may not be easy to implement. For example, rather than use one fallible indicator of a construct, as in standard regression techniques, measurement error could be reduced by using multiple indicators of the same underlying construct, as in Structural Equation
Modelling. However, these techniques require large samples, which may be difficult to obtain in this area of research.
Appendix 9

Practice implications
Whatever the methodological strengths of the catch-up design, the study, time and again, faced the challenge of using social service material to code maltreatment. Typically information was scattered throughout unwieldy files, hidden away in a mass of irrelevant detail. To make matters worse, important information was often absent. Maltreatment classification systems (e.g. Glaser, 2002), for example, recognise the need to distinguish between conceptually distinct forms of emotional maltreatment, but this study was unable to use the categories because of a lack of information. Some risk factors were recorded, but still lacked basic descriptions of important details, such as the duration of the experience. This study had to rely on a proxy measure (length of contact with perpetrators), because a clear statement of the actual duration was rarely recorded. These were not the problems of a handful of social service departments; they were a defining characteristic of nearly every department.

The files used in the study were old, stretching back from the early 1990s to the 1970s, but it is unclear whether file keeping has improved in the intervening years. Social services are overstretched, and it is understandable that keeping well-managed files is seen as secondary to the other demands made on a social worker’s time. However, a summary of a child’s experience of maltreatment is crucial in the planning of services for that child. The results of this study could contribute to the development of a detailed and systematic method of summarising this information. The Risk Factor Manual, a document nearly 20000 words long, would not be of use, but some of its features could form the basis of a standardised summary form.

The form would identify different forms of maltreatment along with brief definitions of each type. To provide a more detailed overview, the standard categories used in social work (i.e. sexual abuse, physical abuse, neglect and emotional maltreatment) may be supplemented with further subdivisions; for example, neglect could be categorised as ‘failure to provide’ or ‘lack of supervision’ and emotional maltreatment could be divided into Glaser’s (2002) five categories. The document would also be used to record basic information about each type of maltreatment,
such as severity, duration, date of onset, number of perpetrators and perpetrator identifying information (e.g. age, gender, relationship to child). Another important feature, again derived from the study reported here, would be a standardised method of summarising the child's care history. This would involve a social worker noting the start date, end date of a care unit, along with information about the identity of carers and the reasons a care unit ended. This basic information, so often missing from social service files, would provide a detailed but brief overview of a child's maltreatment experience, which would be of help in planning for that child.