The Relationship between Investigative Interview Quality, Trial Process, and Outcome in Cases of Child Sexual Abuse

by

Anne Sophie Pichler
BA(Hons) / LLB(Hons)

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Deakin University

March, 2018
I am the author of the thesis entitled “The Relationship between Investigative Interview Quality, Trial Process and Outcome in Cases of Child Sexual Abuse” submitted for the degree of Doctor of Philosophy

This thesis may be made available for consultation, loan and limited copying in accordance with the Copyright Act 1968.

‘I certify that I am the student named below and that the information provided in the form is correct’

Full Name: Anne Sophie Pichler

Signed: [Signature Redacted by Library]

Date: 03 July 2018
DEAKIN UNIVERSITY
CANDIDATE DECLARATION

I certify the following about the thesis entitled “The Relationship between Investigative Interview Quality, Trial Process and Outcome in Cases of Child Sexual Abuse” submitted for the degree of Doctor of Philosophy

a. I am the creator of all or part of the whole work(s) (including content and layout) and that where reference is made to the work of others, due acknowledgment is given.

b. The work(s) are not in any way a violation or infringement of any copyright, trademark, patent, or other rights whatsoever of any person.

c. That if the work(s) have been commissioned, sponsored or supported by any organisation, I have fulfilled all of the obligations required by such contract or agreement.

d. That any material in the thesis which has been accepted for a degree or diploma by any university or institution is identified in the text.

e. All research integrity requirements have been complied with.

'I certify that I am the student named below and that the information provided in the form is correct'

Full Name: Anne Sophie Pichler
Signed: [Signature Redacted by Library]
Date: 1 March 2018
Acknowledgments

First and foremost, I would like to sincerely thank my supervisor, Dr Stefanie Sharman. Her guidance and encouragement have been invaluable in helping me complete this thesis, and her attention to detail has helped me bring my writing to the next level. I am extremely grateful to Stef for always making time for me – day and night - to offer support and constructive feedback. A special thankyou also goes to my late external supervisor Dr Nina Westera. The short time I knew Nina she was an inspiration and mentor, always smiling and positive, enthusiastically encouraging me in my research and supporting me through the challenges of the Royal Commission Project.

Next I would like to thank Professor Don Thomson, whose insightful and helpful comments on my work were much appreciated. Thank you also to Professors Martine Powell and Jane Goodman-Delahunty for offering me the opportunity to work on this once-in-a-lifetime research project with the Royal Commission into Institutional Responses to Child Sexual Abuse, and for helping me gain access and permission to use data from the RC project for my PhD. On this note, I am extremely grateful to the Royal Commission for allowing me to use this data for my thesis. I hope the benefit of having done so will become apparent, and that these studies will help to inform procedure and policy in the area of child witnesses in child sexual abuse trials.

I would also like to express my gratitude to Sarah Zydervelt, Maddie Bearman, and Mairi Benson for their assistance with data coding, and for being a sounding board through challenging stretches of the Royal Commission project. I extend my thanks also to all my friends and colleagues at Deakin University who have supported and encouraged me throughout my candidature. You made my
time at Deakin enjoyable and rewarding, and I hope the connections we made will persist well into our careers.

Last but not least, I would like to thank my family for your continuing support. Thank you for your patience and understanding when I was not at my best, and for sharing this journey with me.
Publications

Abstract

Child sexual abuse (CSA) is one of the most difficult crimes to prosecute and it suffers from high attrition rates. The low rate of successful prosecution may partially be due to the type of evidence that is available in these cases; there are rarely any eyewitnesses or corroborating evidence (such as physical injuries). As a result, the complainant’s evidence is typically central to these cases. In Australia, a complainant’s evidence-in-chief usually takes the form of a pre-recorded child witness interview, conducted by specially trained investigative interviewers. After a child’s investigative interview is played in court they are then cross-examined on their evidence, usually via CCTV. Obtaining reliable and accurate accounts from child witnesses can be extremely difficult; however, after more than three decades of research, there is now a broad international consensus on how children should be interviewed. The quality of interviews is thus usually assessed by whether or not they abide by these “best practice” standards. Given the centrality of child witness interviews in CSA cases, it is conceivable that the quality of the interview might influence prosecutions and trials. However, prior to this thesis, very limited research had examined how these interviews are utilised in court, and how – or whether – their quality influences trials.

Four original studies were conducted that examined different facets of this complex relationship between child witness interviews and trials. The aim of Study 1 was to provide a detailed and current description of the CSA cases currently being prosecuted in Australian courts, and introduce the sample of cases that form the basis of Studies 2 to 4. Study 1 also examined the quality of complainants’ child witness interviews, including the relationship between these interviews and other case characteristics. Given the exploratory nature of the
study, no specific hypotheses were made. Seventy-three CSA trials with 85 complainants from three Australian jurisdictions were read and coded for a number of case, evidence, offence and trial-related variables. Child witness interview transcripts used in these trials were also coded for four measures of interview quality: interviewer questions, interviewer behaviour, interview length, and investigative questions. Results showed that the majority of complainants (78%) were female and most defendants (97%) were male. Trials had between 1 and 4 complainants, who were aged between 7 and 19 at trial ($M = 13.56$ years, $SD = 3.35$). Corroborative evidence was absent in the majority of cases, and contact offences were the most common type of offences prosecuted. Complainants generally underwent long cross-examinations, and 40% of trials resulted in a guilty verdict on at least one charge. Interview quality was generally poor, with only slightly more open than leading questions being asked. None of the case characteristics examined were related to interview quality, but older children were found to be more likely the victims of penetrative offences than younger children. Overall, the findings tentatively suggested that prosecutors do not take interview quality into account when deciding whether or not to prosecute a case.

Study 2 undertook a qualitative examination of the influence of child witness interviews during CSA trials. It examined how legal professionals discussed child witness interviews throughout trials, with a particular focus on interviews’ usefulness as evidence-in-chief. Trial transcripts for the sample of cases introduced in Study 1 were read and discussions regarding the child witness interview were recorded. These discussions were then analysed using open coding to identify the core themes. Results showed that prosecutors, defence counsel and
judges discussed the child witness interview in three different contexts: in terms of their usefulness as evidence-in-chief; during legal discussions about judicial directions and the admissibility of evidence; and in the context of planning and organising the trial. Regarding the child witness interview’s usefulness as evidence-in-chief, discussions showed that legal professionals were at times concerned about interview procedure (including how the child was questioned, and the behaviour and actions of the interviewer and complainant), technological issues, and interview structure (including lack of clarity, lack of coherence, and being overly long). These issues had consequences in the trials such as complainant and juror fatigue, undermining of complainant credibility, delays, and confusion for legal professionals and complainants alike. Overall, findings supported some of the concerns that prosecutors had raised in past studies about the utility of the child witness interview as evidence-in-chief, and highlighted some of the downstream effects that child witness interviews can have at trial.

Study 3 concentrated on an area that prosecutors have highlighted as specifically problematic: the relationship between the amount and type of detail elicited during interviews and how such detail is used to undermine the complainant’s credibility during cross-examination. Specifically, the study determined whether questioning in child witness interviews is related to inconsistencies raised during cross-examination, and to explore the nature of any inconsistencies raised. The cross-examinations of 73 complainants (15 male, 58 female; aged 7 to 19 years, \( M = 13.34, SD = 3.35 \)) were coded for inconsistencies, which were then classified as being either central or peripheral and falling into one of 7 evidential categories. For those inconsistencies that involved the child witness interview, the question that originally elicited the information during the
interview was identified and coded as being either open, specific, or leading. Results showed there were a total of 1087 inconsistencies involving the child witness interview. Two competing hypotheses were made regarding the relationship between questioning and inconsistencies, given the different theories on the purpose of cross-examination. Results supported the first hypothesis (which was based on legal literature), that open-questions should lead to more inconsistencies than specific or leading questions. The alternate hypothesis—based on the investigative interviewing literature—that open-ended questions should result in fewer inconsistencies, was thus not supported. It was also found that significantly more inconsistencies concerned peripheral than central details. Finally, evidential coding showed that a large proportion of inconsistencies occurred in content that was not necessary from an evidential perspective. Overall, the findings support prosecutors’ concerns that long interviews containing minute details increase the number of inconsistencies raised during complainants’ cross-examinations, and invite a rethink of the benefits of open-ended questions when used in the context of a CSA trial.

The final study in this thesis examined the extent to which adherence to best practice interviewing is related to trial outcomes. Of the 85 complainants described in Study 1, eight were excluded because the outcome of their case was missing, resulting in a sample of 77 complainants (16 male, 61 female; aged 7 to 19 years, \(M = 13.83, SD = 3.30\)). A number of variables coded for in Studies 1 and 3 were utilised in Study 4: child witness interview inconsistencies, verdict, corroborative evidence present, number of victims, and all measures of interview quality. The contribution of interview quality on outcome was examined in three scenarios: individually, while controlling for the influence of the number of
corroboration categories present and number of victims (the “control variables”), and collectively. Overall, results did not support the hypothesis that, once the control variables were taken into account, better-quality interviews should be associated with more guilty verdicts. None of the interview variables was significantly predictive of outcome, nor made a significant individual contribution to the model, in any of the scenarios examined. However, in support of the second hypothesis, evidence was found that more cross-examination inconsistencies were associated with fewer guilty verdicts. A hierarchical regression model that included the control variables, two interview quality variables and child witness interview inconsistencies, was significant, and child witness interview inconsistencies made a unique statistically significant contribution to the model. The findings from this study suggest that child witness interview quality may influence trial outcomes indirectly; that is, interview quality influences how many inconsistencies are raised during cross-examination which, in turn, influence trial outcomes.

Overall, the findings of this thesis suggest that once a case reaches trial, interviews have the most influence during cross-examination. Interview quality appears to have little direct effect on trial outcomes, likely because there are numerous factors that can influence trials. Nevertheless, results suggest that there are a number of changes that could be made to current interviewing practices that should improve complainants’ trial experiences, including shortening interviews, reducing the amount of detail requested, and adopting more focused rather than broad open-ended questions. The thesis concludes with a discussion of limitations, and the implications of these findings for researchers, policy makers, interviewers, and interview trainers.
Table of Contents

Acknowledgments...........................................................................................................i
Publications .................................................................................................................. iii
Abstract ......................................................................................................................... iv
Table of Contents ......................................................................................................... ix
List of Figures .............................................................................................................. xiii
List of Tables ................................................................................................................. xiii

CHAPTER 1: INTRODUCTION AND AIMS OF THE THESIS .................................. 15

CHAPTER 2: THE PROCESS OF INVESTIGATING AND PROSECUTING CHILD SEXUAL ABUSE ................................................................. 20
  Key Characteristics of Inquisitorial and Adversarial Legal Systems .... 21
  From Reporting to Trial......................................................................................... 22
  The Trial .............................................................................................................. 24
  The Role of the Police in CSA Investigations ............................................. 26
  The Role of Prosecutors in CSA Investigations ....................................... 27

CHAPTER 3: A BRIEF OVERVIEW OF THE EMPIRICAL LITERATURE ON INVESTIGATIVE INTERVIEWING ............................................. 29
  Common Components of Child Witness Interviews ................................. 30
  The NICHD and SIM Protocols and Coding Schemes........................... 34
  Question Types ................................................................................................. 40
  Adherence to Best Practice Interviewing .................................................. 45
  Interviewer Training ......................................................................................... 48
  The “Dual Use” of the Interview ................................................................. 49

CHAPTER 4: A REVIEW OF THE RECEPTION OF CHILDREN’S EVIDENCE AT TRIAL – TRADITIONAL ATTITUDES AND REFORM...... 51
  Problems Faced by Sexual Assault and Child Witnesses Prior to Reforms................................................................. 51
  Reforms to the Giving of Evidence by Vulnerable Witnesses .......... 53
CHAPTER 5: A REVIEW OF THE RESEARCH ON EVIDENTIAL QUALITY OF CHILD WITNESS INTERVIEWS .............................................. 59
  Style of Questioning .............................................................. 60
  Amount of Detail Sought ....................................................... 60
  Particularisation ................................................................. 62
  The Relationship Between Best Practice Interviewing and Case Outcomes ....................................................................... 65
  Conclusions and Justification for Current Thesis .................. 67

CHAPTER 6 – CHARACTERISTICS OF CHILD SEXUAL ABUSE TRIALS: EXPLORING DATA FROM THREE AUSTRALIAN STATES
(STUDY 1) ............................................................................................................ 68
  Method ................................................................................................................... 72
    Trial and Child Witness Interview Transcripts .......................... 72
    Procedure ........................................................................................................... 73
  Results ................................................................................................................... 81
    Case Characteristics ........................................................................ 81
    Evidence Characteristics ........................................................... 82
    Offence Characteristics ............................................................. 84
    Trial Characteristics ........................................................................ 85
    Interview Quality ............................................................................. 87
    Associations Between Case Factors and Interview Quality ........ 88
    The Relationship Between Age at Cross-Examination and Case Variables ................................................................. 90
  Discussion ............................................................................................................. 91

CHAPTER 7 - EXAMINING DISCUSSIONS ABOUT CHILD WITNESS INTERVIEWS HELD DURING TRIALS OF CHILD SEXUAL ABUSE
(STUDY 2) ............................................................................................................ 97
  Method ................................................................................................................. 100
    Trial Transcripts and Child Witness Interviews .......................... 100
    Data Management and Analysis .................................................. 100
# Results

- Usefulness of the Complainant’s Child Witness Interview as Evidence-In-Chief
- Legal Issues
- Trial Planning

# Discussion

## CHAPTER 8: CHILD WITNESS INTERVIEW QUALITY AND INCONSISTENCIES RAISED DURING CROSS-EXAMINATION: IS THERE A CONNECTION? (STUDY 3)

- Method
  - Trial and Child Witness Interview Transcripts
  - Procedure
- Results
  - Relationship Between the Child Witness Interview and Inconsistencies
  - Nature of the Inconsistencies

## CHAPTER 9: THE CONTRIBUTION OF INTERVIEW QUALITY AND INCONSISTENCIES IN PREDICTING VERDICTS IN CASES OF CHILD SEXUAL ABUSE (STUDY 4)

- Method
  - Trial and Child Witness Interview Transcripts
  - Procedure
- Results
  - Individual Contributions of Variables
  - Predicting Outcome Using Multiple Variables of Interest
### CHAPTER 10: GENERAL DISCUSSION ........................................................ 164

- The Relationship Between the Child Witness Interview and the Trial Process ....................................................................................... 165

- Relationship Between the Child Witness Interview and Cross-Examination ........................................................................................ 169

- The Relationship Between the Child Witness Interview and Trial Outcome .......................................................................................... 173

- Limitations and Future Research ................................................................................................................................. 175

- Implications for Practice and Research ......................................................................................................................... 177

- Concluding Comment ............................................................................................................................................................... 182

---

**References** ........................................................................................................... 184
List of Figures

Figure 3.1: Classification of Question Types Used by the NICHD and SIM Coding Protocols ........................................ 37

List of Tables

Table 6.1 Case Characteristics for 85 Child Sexual Abuse Complainants .................................................. 81
Table 6.2 Evidence Characteristics for 85 Child Sexual Abuse Complainants .................................................. 83
Table 6.3 Offence Characteristics for 85 Child Sexual Abuse Complainants .................................................. 85
Table 6.4 Trial Characteristics for 85 Child Sexual Abuse Complainants .................................................. 86
Table 6.5 Interview Quality Variables ........................................ 88
Table 8.1 Evidential Coding Scheme Categories .................... 128
Table 8.2 Examples of Central and Peripheral Inconsistencies ... 130
Table 8.3 Examples of Inconsistencies and Corresponding Interview Questions ........................................ 132
Table 8.4 Source of Inconsistencies Which Included the Child Witness Interview ........................................ 135
Table 8.5 Type of Inconsistency ............................................. 135
Table 8.6 Source of Own-Statement Inconsistencies ............. 136
Table 8.7 Expected Frequencies Used in Follow-Up Chi-Square Analyses .................................................. 137
Table 8.8 Content of Inconsistencies .................................... 139
Table 9.1  Binomial Regression Analyses Between Each Variable and Outcome ................................................. 153

Table 9.2  Hierarchical Binomial Regression Analyses Between Each Variable of Interest and Outcome ................. 154

Table 9.3  Correlations Between Different Interview Quality Variables .......................................................... 155

Table 9.4  Hierarchical Logistic Regression Predicting Likelihood of At Least One Guilty Verdict ............... 158
CHAPTER 1: INTRODUCTION AND AIMS OF THE THESIS

Child sexual abuse (“CSA”) is a global issue that affects millions of children worldwide. For the purposes of this thesis, CSA is defined as “the involvement of a child in sexual activity that he or she does not fully comprehend, is unable to give informed consent to, or for which the child is not developmentally prepared and cannot give consent, or that violates the laws or social taboos of society” (World Health Organization, 2003, p. 75), where a “child” is anyone under the age of 18 years. In 2005, the United Nations Children Fund (“UNICEF”) estimated that 150 million girls and 73 million boys had been victims of sexual abuse involving some form of physical contact (UNICEF, 2005). In Australia, the prevalence rate is about 73/100,000 for boys and 246/100,000 for girls under the age of 14 (Australian Institute of Criminology, 2014). CSA occurs in both developing and developed countries, across all races and ethnic groups, and amongst all economic classes (Australasian Institute of Judicial Administration, 2012; Ronken & Johnston, 2012; UNICEF, 2005). Victims of CSA often suffer significant health, behavioural, and psychological issues as a result of the abuse, including anxiety, depression, post-traumatic stress disorder, eating disorders, suicidal and self-harming behaviours, as well as HIV and other sexually transmitted infections (Australasian Institute of Judicial Administration, 2012; Papalia, Luebbers, Ogloff, Cutajar, & Mullen, 2017; World Health Organization, United Nations Development Programme, & United Nations Office on Drugs and Crime, 2014).

The sexual abuse of a child is a crime both under Australian domestic law (for example, see ss. 45-49 Crimes Act 1958 (Vic)) and under international law
(e.g., Principle 9 of the UN Convention on the Rights of the Child). However, CSA is one of the most difficult crimes to prosecute (Australasian Institute of Judicial Administration, 2012). It is also significantly under-reported and has one of the highest attrition rates of all criminal offences (Australian Law Reform Commission, 2010; Eastwood, Krift, & Grace, 2006). Although exact figures are difficult to come by, estimates suggest that only about 5-25% of CSA cases are reported (Kelly, Lovett, & Regan, 2005), between 76% and 85% of reported cases are “lost” at the police stage (Ferrante, Clare, Randall, & Boyd, 2017; Fitzgerald, 2006), and only between 9% and 19% of cases are prosecuted (Cross, Whitcomb, & De Vos, 1995; Cashmore, Taylor, Shackel, & Parkinson, 2016; Hood & Boltje, 1998). Of those few cases which reach court, a relatively high percentage (studies suggest anywhere between 41% and 90%: Cashmore et al., 2016; NSW Bureau of Crime Statistics and Research, 2017; Wundersitz, 2003) result in a conviction. These conviction rates are markedly lower than most other offences; for example, between 2012 and 2016 the conviction rate for child sexual offences in NSW was 60%, compared with 70% for assault, and 73% for robbery. Only adult sexual assault conviction rates were lower, at 50% (NSW Bureau of Crime Statistics and Research, 2017).

There are a number of factors that may contribute to this high attrition and low prosecution rate. One is the unique dynamics that exist in CSA cases. Child sexual abuse usually occurs in secret (Bradshaw & Marks, 1990; Oates, 2007; Quadara, 2014), and is typically perpetrated by someone known to the child (Bradshaw & Marks, 1990; Oates, 2007; Quadara, 2014). It is frequently preceded by months or even years of grooming behaviour, in which the perpetrator creates a relationship of trust and prepares the child for future abuse (Australasian
and often involves repeated offending over weeks, months or even years (Cossins, 2006; Quadara, 2014). Victims of CSA are frequently threatened or sworn to silence (Australasian Institute of Judicial Administration, 2012; Oates, 2007; UNICEF, 2005), which, in combination with the relationship they (or their family) have with the offender, the fear that they will not be believed, and the fear of the consequences of telling someone, usually prevents them reporting the abuse in a timely manner (Australian Law Reform Commission, 2010; Collin-Vézina, De La Sablonnière-Griffin, Palmer, & Milne, 2015; Tashjian, Goldfarb, Goodman, Quas, & Edelstein, 2016). Delayed complaint (including complaints made in adulthood) is thus common, whereas immediate reporting is rare (Goodman-Brown, Edelstein, Goodman, Jones, & Gordon, 2003; Shead, 2014).

Apart from leading to low reporting rates, these features of CSA also influence the nature of the evidence that is available to investigators and prosecutors. Most importantly, the secretive nature of CSA means that there are rarely eyewitnesses to the abuse (Australasian Institute of Judicial Administration, 2012; Cossins, 2000). Other corroborating evidence, such as medical evidence of penetration, physical injury, or the presence of semen is also usually absent (Australasian Institute of Judicial Administration, 2012; Cossins, 2000; Gothard, 1987; Oates, 2007; W. A. Walsh, Jones, Cross, & Lippert, 2010). This lack of eyewitness and medical evidence likely has a huge impact on the rate of prosecution. For example, Gallagher and Please (2000) noted that out of 1000 CSA investigations in the United Kingdom, 47% were dropped by the prosecution because of insufficient corroborative or medical evidence. It follows, therefore, that in most cases the evidence given by the complainant of the abuse they
suffered is crucial to successful prosecution. Rarely can a case advance to trial if the complainant is not able or willing to testify (Cross, De Vos, & Whitcomb, 1994; Hoyano & Keenan, 2010).

A child’s evidence takes the form of an “investigative” or “child witness interview” (Oates, 2007). This is conducted by specially trained investigators using questioning techniques that encourage the child to give a reliable and detailed account of the abuse, while minimising the risk of contamination (Hoyano & Keenan, 2010). In Australia, a video-recording of this interview can be used as a child’s evidence-in-chief at trial (see e.g., ss. 366 – 368 Criminal Procedure Act 2009 (Vic); ss. 306(2), 306U(1) – (2) Criminal Procedure Act 1986 (NSW)).

Most research to date has focused on improving the quality of the questioning in child investigative interviews to maximise the reliability of the child’s account (Lamb & Fauchier, 2001; Lamb, Hershkowitz, Orbach, & Esplin, 2008; Powell, Fisher, & Wright, 2005; Powell & Thomson, 2001). Some research has also examined the evidential quality of child witness statements from the perspectives of prosecutors (Burrows & Powell, 2013, 2014c, 2014d). However, to date, no studies have investigated how these interviews are utilised in court, and there is only limited research examining how – or whether – they influence the trials.

The aim of this thesis was to examine the relationship between child witness interview quality, CSA trial process, and trial outcome. Specifically, it examined 1) the profile of a sample of 73 Australian CSA cases, 2) whether and how legal professionals discuss child witness interviews during CSA trials, with a particular focus on the interviews’ usefulness as evidence-in-chief, and 3) the
relationship between child witness interview quality, case characteristics, cross-examination, and verdict.

Chapters 2 to 5 provide a critical review of the literature relevant to the thesis, namely: the process of investigating and prosecuting CSA (Chapter 2); best practice investigative interviewing (Chapter 3); the use of the pre-recorded investigative interview in court (Chapter 4); and the evidential quality of investigative interviews with children (Chapter 5). As the main focus of the thesis is on child complainants in Australia, this literature review focuses primarily on research surrounding child complainants from adversarial legal systems.
CHAPTER 2: THE PROCESS OF INVESTIGATING AND PROSECUTING CHILD SEXUAL ABUSE

The investigative interview of a CSA complainant takes place at the very beginning of the investigation into alleged CSA. If the case reaches trial, this same interview may then become pivotal again at trial. Between these two time points, the audio-visual or written record of the interview (along with any corroborating evidence) passes between a number of agencies and its evidential strength and the likelihood of successful prosecution is assessed. As noted in the introduction, a large percentage of cases are lost through this process, and the proportion of cases that successfully reach trial is very small. Although the focus of this thesis is on the relationship between the quality of the interview and trial, it is necessary to have some understanding of how cases progress to trial, and how the trial fits into the broader process of prosecuting CSA. This chapter provides a brief outline of the process of investigating and prosecuting CSA in Australia up to and including the trial. Given that the studies included in this thesis are set in an adversarial legal system, the literature discussed is limited to literature from countries with adversarial legal systems (except where necessary for comparison purposes). The main differences between inquisitorial and adversarial legal systems are thus briefly explained first to illustrate why this restriction is necessary. The chapter concludes with a brief examination of the roles of two key decision-makers in the CSA investigation and prosecution process: the police and the prosecutor.
Key Characteristics of Inquisitorial and Adversarial Legal Systems

Although no legal systems are exactly alike, systems in western countries can broadly be categorised as either being mostly adversarial or inquisitorial in nature (Cordon, Goodman, & Anderson, 2003). Examples of countries with adversarial systems include Australia, the UK, the USA, and Canada. Most European countries, in contrast, have inquisitorial style systems. These include the Netherlands, France, Germany, Belgium, and Norway.

Generally speaking, in inquisitorial legal systems the judge and court are active participants investigating the facts of a case (Myklebust, 2017; van Koppen & Penrod, 2003), and there is a strong preference for documentary evidence (Cordon et al., 2003). In contrast, in adversarial systems the judge has a largely passive role, impartially adjudicating the case on the basis of evidence presented to him or her by the prosecution and defence (Finn, 2009; Myklebust, 2017). There is also a strong preference for oral evidence (Cossins, 2010). In inquisitorial criminal trials, the case is decided by the judge, who has any relevant evidence at his or her disposal (van Koppen & Penrod, 2003); in contrast, in adversarial criminal trials, the case is decided by the jury, and there are strict rules of evidence regarding which evidence is admissible and may be presented to the jury (Cordon et al., 2003; see also, for example, Evidence Act 1995 (NSW); Evidence Act 2008 (Vic)). For example, there are strict rules against the admission of hearsay in adversarial systems (that is, information that has been told to the witness by another person, not something they heard or saw themselves: Victoria Law Foundation, 2015; s. 59 Evidence Act 2008 (Vic)), while such evidence can be admitted freely in inquisitorial systems (Cordon et al., 2003). Of note, a pre-recorded child witness interview is also classified as hearsay
evidence, which is why reforms needed to be enacted to allow this type of
evidence to be admitted in court.

One key difference between the systems for complainants is how the
evidence of witnesses is tested and challenged. In adversarial systems, this is
usually done through cross-examination in front of a jury (Australian Law Reform
Commission, 2010; Cossins, 2009); however, in inquisitorial systems, cross-
examination generally does not take place. Instead, judges, magistrates or other
legal representatives question a witness outside of court (e.g., in the judge's
chamber; Cordon et al., 2003; Finn, 2009). The cross-examination tactics that are
employed in adversarial systems to influence juror’s assessments of witnesses are
thus not utilised in inquisitorial trials. Further, in many continental European
countries in which inquisitorial systems are in place, the defendant and his or her
legal representative are not allowed to confront a complainant in child sexual
assault cases (Cordon et al., 2003; Myklebust, 2017). A complainant’s
involvement in the inquisitorial court process thus ends after they are questioned
on their interview pre-trial (Myklebust, 2017); in contrast, complainants are
required to be present for cross-examination (even if via CCTV) in adversarial
criminal trials. Thus, due to these significant differences between the two legal
systems, research concerning the utility of the child witness interview as
evidence-in-chief should be evaluated within the frameworks of the relevant
system.

From Reporting to Trial

A suspected CSA case enters the system when it is reported to authorities.
A report may be made by the victim or a person concerned about the welfare of
the child such as a teacher, doctor or neighbour (Australian Law Reform
Commission, 2010). As soon as possible after this initial report, trained police investigators will conduct a child witness interview that is video and audio recorded (ALRC, 2010). If the case proceeds to trial, this recorded interview usually forms all or part of the child’s evidence-in-chief. Initially, however, the interview will help police decide whether further investigation is called for, and whether a prima-facie case exists (Christensen, Sharman, & Powell, 2014). If this is decided in the affirmative, the suspect will be charged, a brief of evidence compiled, and the case referred on to the Office of the Director of Public Prosecutions (ODPP) (Australian Law Reform Commission, 2010; Lievore, 2005). The case will then be reviewed by a prosecutor assigned by the ODPP who will decide whether or not to accept the case for prosecution. If the case is accepted, an indictment will be prepared and filed, and the defendant will be required to enter a plea. If the defendant pleads guilty, he or she is convicted; if he or she pleads not guilty, the case will be decided by a jury at trial (Christensen et al., 2014).

Cases can drop out of the system at any stage of this process. They are continually reassessed, and subject to increasingly stringent evidentiary standards as they move “upwards” through the system towards trial (Australian Law Reform Commission, 2010). At the police stage, a decision to charge is based on a “prima facie” test—that is, that on the face it, there is enough evidence to suggest that the prosecution’s case would be successful (Australian Law Reform Commission, 2010; Lievore, 2005; Victoria Law Foundation, 2015). Prosecutors, on the other hand, decide whether or not to accept the case based on a “reasonable prospects of conviction test”; and a jury can only convict if it is convinced of a
suspect’s guilt “beyond reasonable doubt” (Australian Law Reform Commission, 2010; Lievore, 2005)

Before a case reaches trial, the decision-making process is subject to few laws and external oversights; that is, police and prosecutors have a considerable amount of discretion in making their decisions (Lievore, 2004; Bronitt & Stenning, 2011). As a result, much of the basis of decisions by police and prosecutors is unknown (Kelly et al., 2005; Lievore, 2005). Communication between police and prosecutors has also generally been poor (Christensen et al., 2014), and police in particular are uncertain of legal requirements of investigative interviews, and the details that are required for a successful prosecution (Davis, Hoyano, Keenan, Maitland, & Morgan, 1999; Guadagno, Powell, & Wright, 2006; Burrows, Powell, & Anglim, 2013). The disparity between the details police and prosecutors believe are required for prosecution will be discussed further below. For now, it is sufficient to note that greater inter-agency communication is needed to improve CSA case processing and prosecution (Christensen et al., 2014; Lievore, 2004), particularly in a system where decisions made at one stage of the investigation process are influenced by decisions expected at the next stage of the process (Australian Institute of Family Studies, 2010).

The Trial

The trial process that must be followed in all Australian jurisdictions is set out in legislation and bench books (e.g., Criminal Procedure Act 1986 (NSW); Criminal Procedure Act 2009 (Vic); Jury Act 1977 (NSW); Judicial College of Victoria, 2015; Judicial Commission of New South Wales, 2015). A trial usually starts with the jury being empanelled, and the judge providing an opening address
to the jury. The judge’s address is usually followed by an opening address by the prosecutor or “Crown” (who represents the state), and the defence counsel may also address the jury. Following the defence’s address, the witnesses are called to give evidence under oath (unless they are considered not competent to give evidence under oath, in which case they will give unsworn evidence). This presentation of the witness’ evidence is called “evidence-in-chief”. It is followed by “cross-examination”, in which the defence tests the witnesses’ evidence by asking them questions about their testimony. Cross-examination may be followed by “re-examination”, in which the prosecutor may ask the witness additional (or “supplementary”) questions to clarify issues that arose during cross-examination.

After all witnesses for the prosecution have been called, the accused may give evidence. However, it is up to the Crown to prove their case (that is, they have the “burden of proof”) and the accused need not give evidence. At the conclusion of a trial, the prosecutor and defence counsel provide a closing address to the jury in which they summarise their case. The judge then instructs the jury regarding issues of law, provides any relevant warnings, and the jury retires to consider its verdict. For criminal cases (such as CSA cases), a unanimous verdict (meaning an agreement between all 12 jury members) is usually necessary but in certain circumstances a majority verdict (an agreement between 11 jurors) is permissible. The verdicts for each offence are then revealed to the court and the accused, and the jury is discharged. The accused has the right to appeal any errors of law, and frequently does so; however, errors of fact (which is the domain of the jury) cannot be appealed.
The Role of the Police in CSA Investigations

The police are usually the first contact that a CSA victim has with the criminal justice system. How police respond to a report of suspected CSA is crucial for the successful passage of the case through the criminal justice system, as well as the welfare of the victim (ALRC, 2010; (Royal Commission into Institutional Responses to Child Sexual Abuse, 2017). In recognition of this last function, police in Victoria and other Australian states employ inter-agency teams to investigate CSA cases in an effort to provide better victim support and continuity for victims (Australian Law Reform Commission, 2010; Victoria Police, 2014). Along with specialist detectives, inter-agency teams always include a child protection worker and some also include other health workers (ALRC, 2010).

Although police are there to support victims, they also act as the most important gatekeepers to the criminal justice system (Bunting, 2008; Ferrante et al., 2017). It is at this stage that the highest percentage of cases are “lost” from the system (Kelly et al., 2005; Wundersitz, 2003). Given that the criminal justice system has a limited amount of resources, and it is not possible to attempt to prosecute all alleged cases of criminal conduct (Lievore, 2004), the loss of some cases is not necessarily undesirable. However, it is crucial that cases dropped by police are not “prosecutable” cases discontinued due to avoidable evidential difficulties.

In deciding whether to refer a case for prosecution, the police must be reasonably satisfied that 1) an offence occurred, 2) prosecution would be in the public interest, and 3) there is sufficient evidence to enable a prosecution (Bronitt & Stenning, 2011; Gallagher, 1999). The victim (or in the case of young children,
the victim’s parents) must also be willing to pursue a prosecution (Cashmore et al., 2016; Ferrante et al., 2017). The limited research that has been conducted into police decision-making suggests that in CSA cases, the quality and availability of evidence (including the existence of corroborating evidence) is one of the key factors that determines whether or not a case is referred on to the prosecution (Powell, Murfett, & Thomson, 2010; W. A. Walsh et al., 2010). However, the quality of the child investigative interview is rarely mentioned in this decision-making process despite it being one of the key pieces of evidence in CSA prosecutions (Powell, Murfett, et al., 2010; Cross & Whitcomb, 2017).

**The Role of Prosecutors in CSA Investigations**

Prosecutors’ main pre-trial role is deciding whether to accept a case referred by police for prosecution. Prosecutors have immense discretion in making this decision (Alderden & Ullman, 2012), and can alter charges laid by police or add additional charges (ALRC, 2010). However, their decisions must be guided by two key considerations. The first of these is the sufficiency of the evidence, and the second is the public interest (ALRC, 2010; Eastwood et al., 2006).

Generally, public interest considerations are not pivotal in CSA cases as these usually weigh in favour of prosecution (Gallagher, 1999). Again, however, the limited research that has been conducted suggests that the evidential quality of a CSA case is key in prosecutorial decision-making (Davis et al., 1999; Gallagher, 1999; Gothard, 1987). Other factors that have been found to make it more likely for a case to be prosecuted include the victim being older than five at the time of the offence, the victim being female, and reporting of the offence no later than 10 years after it occurred (Edelson, 2012; Fitzgerald, 2006). Prosecutors also take
into account the competence, credibility and reliability of witnesses, and the likely impression that a witness might have on a jury (ALRC, 2010). For children, these latter factors can be influenced significantly by the quality of their child witness interview; thus, it is crucial that best practice interviewing recommendations are adhered to. The following chapter provides an overview of the empirical literature on best practice interviewing with children.
CHAPTER 3: A BRIEF OVERVIEW OF THE EMPIRICAL LITERATURE ON INVESTIGATIVE INTERVIEWING

It is well established that interviewers can influence children’s responses through their questioning practices (Brown, Lamb, Pipe, & Orbach, 2008; Bull, 2010; G. Davies, Westcott, & Horan, 2000; Lamb et al., 2008; Oates, 2007; Wood, 2000), and that obtaining reliable, detailed and accurate accounts from child witnesses is extremely difficult (Hoyano & Keenan, 2010; Powell & Thomson, 2001). Among other factors, children’s cognitive, linguistic and developmental limitations influence their ability to understand questions and give accurate answers (Lamb, Orbach, Hershkowitz, Horowitz, & Abbott, 2007; Powell & Thomson, 2001; W. A. Walsh et al., 2010). These limitations are likely compounded in children who have been abused, as child abuse has been associated with impoverished language development and competence (Snow, Powell, & Sanger, 2012).

Nevertheless, after more than three decades of research on memory and interviewing, there is a broad international consensus that children can give accurate accounts of events so long as they are questioned appropriately (Bull, 2010; Lamb, Orbach, Hershkowitz, Esplin, & Horowitz, 2007; Powell & Snow, 2007). In fact, the main factor determining the length, accuracy and structure of children’s accounts is the type of questions they are posed (Brown & Lamb, 2015). Errors and misunderstandings in children’s accounts are reduced if the interviewer adheres to best-practice interviewing guidelines. Thus, it is the responsibility of the interviewer to ensure that the interviewee’s account is as accurate and coherent as possible (Powell & Thomson, 2001).
This chapter presents a summary of the empirical literature on best-practice interviewing of children. It first provides a brief overview of the common components in child witness interviews, and a comparison between two commonly used protocols: the National Institute of Child Health and Human Development Investigative Interview Protocol (NICHD Protocol) and the Standard Interview Method (SIM). It then evaluates the literature on three question types – open, specific, and leading – in more detail, before briefly examining the literature on adherence to best practice interviewing and interviewer training. Finally, the issue of the “dual use” of the interview is discussed.

**Common Components of Child Witness Interviews**

*Ground rules, truth and lies.* Interviews typically commence with a short introduction about the purpose of the interview, followed by an explanation of the ground rules (or conversational rules) of the interview (Benson & Powell, 2015a; Lamb et al., 2008; Sternberg et al., 1997). This explanation is important because investigative interviews are very different to the interactions that children usually experience in everyday life. For example, when children are questioned by teachers, they are usually expected to provide the best answer they can - even if they do not know the correct one (Powell & Thomson, 2001). It is also generally accepted in everyday discourse that the teacher, parent or other adult knows more about the topic of conversation than the child (Powell & Thomson, 2001). This assumption is not the case in investigative interviews, where children are the experts and it is crucial that they do not guess an answer that they do not know (Lamb et al., 2008). Common ground rules include the need to tell the truth, to only report events that the child has personally experienced, that it is ok to use
any words including sexually explicit or swear words, and to tell the interviewer if they do not understand a question, do not remember, or do not know the answer (Benson & Powell, 2015a; Brubacher, Poole, & Dickinson, 2015; Lamb et al., 2008). Interviewers should also explain to children that the interviewer does not know what happened, and the child should correct the interviewer when necessary (Brubacher et al., 2015; Lamb et al., 2008).

Establishing ground rules has been shown to reduce confusion and errors, and increase children’s resistance to suggestion (Benson & Powell, 2015a; Ceci & Bruck, 1995). Nevertheless, the best method of implementing ground rules is still an issue of contention (Brubacher et al., 2015). For instance, prosecutors have raised concerns that lengthy instructions fatigue the child and undermine their perceived competence in the eyes of a jury (Benson & Powell, 2015a). Such concerns extend to the truth-lies test, which is a competency test frequently conducted in interviews before the child is asked to promise to tell the truth (Benson & Powell, 2015a; Lyon, 2011). This test has been shown not to predict a child’s truthfulness throughout the interview and, although historically a legal requirement of child witness interviews, it is now no longer required in many commonwealth jurisdictions (Benson & Powell, 2015a; Bowden, Henning, & Plater, 2014; Lyon, 2011).

**Rapport and the practice narrative.** Rapport refers to the relationship or interpersonal connection between an interviewer and child; good rapport is crucial in investigative interviews. The better the rapport, the more likely the child is to share sensitive information, or information that he or she has been told to keep secret (Powell & Thomson, 2001; Wood, 2000). Rapport is usually established during the practice narrative, where the child is questioned about a neutral event
or favourite activity, although it is the type of interaction that is important during this stage rather than the topic of conversation (Hershkowitz, 2011; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007). During the practice narrative, children are encouraged to provide as much detail as possible in response to open-ended questions without any pressure to provide information they do not have (Benson & Powell, 2015a; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007).

This practice narrative stage is important for at least two reasons. First, it familiarises the child with the questioning style that will be used during the substantive phase of the interview and the amount of detail that is required in response to questions (Lamb, Orbach, Hershkowitz, Esplin, et al., 2007). Research has shown that children who are asked open questions during the rapport phase of an interview provide quantitatively and qualitatively better responses during the substantive phase of an interview than children who are asked specific questions during the rapport phase (Price, Roberts, & Collins, 2013; Roberts, Brubacher, Powell, & Price, 2011; Sternberg et al., 1997). Some research also suggests that a well conducted practice narrative can also reduce the length of the substantive phase of an interview, without any negative effect on the amount of information reported (Brown et al., 2013). Second, teaching children that they need not provide an answer to a question if they do not know that answer helps to reduce their susceptibility to suggestion in the substantive phase of the interview (Powell & Thomson, 2001).

Finally, to establish rapport, the context of the interview is also important. Rapport should be easier to establish in a supportive environment, which is less formal and helps the child feel less anxious and intimidated (Lamb, Orbach, Hershkowitz, Esplin, et al., 2007; Powell & Thomson, 2001). The environment
should be free from distractions, such as incoming phone calls and toys (Lamb et al., 2008). This need for reduced formality should, however, be balanced against the need for *some* formality, as it is always possible the interview will be played in court as a child’s evidence-in-chief, and prosecutors have suggested that less formal interviews may reduce the child’s credibility in the eyes of a jury (Burrows & Powell, 2014d). Additionally, although the rapport stage is important, it should not be overly long so as to fatigue the child or the jury (G. Davies et al., 2000).

**The substantive phase.** At the beginning of the substantive phase of the interview, when children are prompted to discuss the topic of concern, they should be encouraged to give an account of the event or events in question in their own words using a free narrative (Powell & Snow, 2007). A free narrative account is one in which interviewees recount an event in their own words, at their own pace, without interruption (Powell & Thomson, 2001). It links activities and experiences in such a way that a person who is unfamiliar with the story can understand precisely what has happened (Powell & Snow, 2007). A free narrative is typically started in response to an open-ended question that encourages an elaborate response such as, “Tell me what you’re here to talk to me about today” (Earhart, Danby, Brubacher, Powell, & Sharman, in press). Interviewing guidelines suggest that such broad open-ended questions are the best way to transition from the pre-substantive to the substantive phase of the interview, and raise the topic of concern (Benson & Powell, 2015a).

After the initial free narrative, the interviewer should encourage the child to continue talking using minimal non-verbal encouragers (for example, head nodding, saying “mmm”) and additional non-leading open-ended questions (Powell, 2002; Powell & Thomson, 2001). Only once open-ended questions have
been exhausted and further details are required for forensic or legal purposes should interviewers utilise more specific questions, such as “What colour were your pants?” (Hughes-Scholes & Powell, 2013; Lamb & Fauchier, 2001; Powell & Snow, 2007). These specific questions may also be required to clarify inconsistencies that have arisen during the interview (Brown et al., 2008).

Although there are no definitive guidelines on the proportion of questions that should be open-ended, some child interviewing experts have suggested that around 80% of questions should be open (Luther, Snook, Barron, & Lamb, 2015), while others have classified interviewers who used fewer than 40% open-questions as “low performers” (Powell, Guadagno, & Benson, 2016). Suggestive techniques and questions, including tag questions (e.g., “he forced you to do that, didn’t he?”), should be avoided throughout the interview, and interviewers should also avoid introducing new information themselves (Lamb, 2016; Powell, 2002).

The empirical evidence on which these best practice questioning guidelines are based will be reviewed further below. However, first it is necessary to introduce two protocols and associated question coding schemes that have been used in the empirical literature to evaluate question types and interview quality: The NICHD and SIM protocols and coding schemes. The NICHD protocol is reviewed because it is the most commonly used protocol in the empirical literature, and the SIM is discussed because it is the protocol used most commonly in Australia.

**The NICHD and SIM Protocols and Coding Schemes**

Interviewing protocols have been developed to help interviewers adhere to best practice guidelines. There is a wide variety of child interview protocols available, including the CornerHouse Forensic Interview Protocol (Anderson et al., 2010), The Stepwise Approach (Yuille, Marxsen, & Cooper, 1999), the State
of Michigan Forensic Interviewing Protocol (The State of Michigan Governor's Task Force on Child Abuse and Neglect and Department of Human Services, 2011), and the Guidance for Achieving Best Evidence in Criminal Proceedings (Home Office, 2011). These are all used in various jurisdictions in Canada, the US, and the UK.

One widely used protocol is the NICHD Protocol, developed by Lamb and colleagues (Lamb et al., 2008; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007). It is a flexible but structured guide on how to use best-practice techniques in interviews with children in a manner that promotes memory retrieval and the reporting of events (Lamb et al., 2008). To date, it has been tested and shown to be effective in improving interview quality (measured through the number of open-ended, suggestive, and risky prompts utilised) in Israel, the US, the UK, Canada, and Korea (Benia, Hauck-Filho, Dillenburg, & Stein, 2015; Hershkowitz, Horowitz, & Lamb, 2005; Lamb et al., 2008; Yi, Jo, & Lamb, 2016).

In Australia, there has been a gradual uptake of a different protocol, referred to as the SIM (see e.g., Benson & Powell, 2015a; Powell, 2014). Of particular importance in the Australian context is the fact that the SIM has been evaluated and shown to be effective for interviews with Aboriginal children (Hamilton, Brubacher, & Powell, 2016; Hamilton, Powell, & Brubacher, 2017). The phases of the SIM are like those of other best-practice protocols, and include the introduction and ground rules, episodic memory training or practice narrative, the substantive phase, the break, further questioning, closure, and a neutral topic. One unique aspect of the SIM is its guidelines regarding the length a practice narrative should last (considering the limited cognitive resources of children). The
SIM also provides detailed instructions around questioning children who experienced repeated incidents of abuse.

Researchers using the NICHD and the SIM protocols have developed coding schemes to evaluate the quality of investigative interviews (e.g., the number of open-ended questions asked). Figure 3.1 outlines the two coding schemes; it was created by examining the NICHD and SIM coding schemes, studies utilising the schemes, and correspondence with researchers who had used them to clarify any discrepancies. The left-hand “thesis coding scheme” column illustrates the three broad question categories that were utilised in the studies presented in this thesis, and which are discussed in more detail below.

Figure 3.1 shows that the two coding schemes define question-types in ways that are not completely compatible. This can make comparisons between studies using the different protocols difficult. The main difference between the NICHD and the SIM definition of open-ended questions is the requirement in the SIM scheme that prompts not dictate what information is required. In this way, it is a slightly more conservative definition of open-ended questions than the one utilised by the NICHD scheme (Benson & Powell, 2015b). Additionally, the SIM scheme categorises all “can you” type prompts according to their correct grammatical construction as yes/no prompts, including, “can you tell me…” type questions (Benson & Powell, 2015a). In contrast, the NICHD scheme typically codes questions according to their intended meaning; thus, “can you tell me…” type prompts should be coded as open-ended.
<table>
<thead>
<tr>
<th>Thesis Coding Scheme</th>
<th>NICHD Question Coding Scheme</th>
<th>SIM Question Coding Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Invitation</td>
<td>Defined as any invites that encourage an elaborate response</td>
<td>Defined as a question that encourages a narrative response but does not dictate what information is required</td>
</tr>
<tr>
<td>General Invitation</td>
<td>An invitation which requests information about what else happened, or what happened next.</td>
<td>Breadth questions encourage the child to recall what happened next, preceding, first, or last activity that occurred during the event.</td>
</tr>
<tr>
<td>Cued Invitation</td>
<td>A type of invitation which refocuses the child’s attention on details s/he mentioned and uses them as cues to prompt further free-recall of information</td>
<td>Depth questions encourage the child to provide more detail about a part of a pre-disclosed event/experience</td>
</tr>
<tr>
<td>Specific Directive</td>
<td>A cued-recall prompt that focuses the child’s attention on pre-disclosed information and requests additional information about specific categories (Wh-questions)</td>
<td>Questions that focus the child’s attention on pre-disclosed details of the alleged event (Wh-questions)</td>
</tr>
<tr>
<td>Specific</td>
<td>Option-posing</td>
<td>A question that focuses the child’s attention on aspects or details that <em>may or may not</em> have been previously mentioned, but does not imply that a particular response is requested. These questions are typically yes-no or forced choice questions.</td>
</tr>
</tbody>
</table>

**Figure 3.1: Classification of Question Types Used by the NICHD and SIM Coding Protocols.**

* see e.g., Earhart, La Rooy, Brubacher, & Lamb, 2014; Lamb et al., 2008; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007.  
  see e.g., Benson & Powell, 2015b; Powell, Guadagno, et al., 2016; Powell & Hughes-Scholes, 2009.  
  The definition of “leading” here is broader than the NICDH’s “suggestive utterances” because introducing information not previously disclosed by the child is usually considered a suggestive technique: Brown et al., 2008. High-risk specific yes/no questions have previously been classified as leading in some studies: Benson & Powell, 2015b; Powell & Hughes-Scholes, 2009.
The SIM has three categories of specific question (cued recall, forced- 
choice, and yes/no-low risk), which correspond roughly with the directive and 
option-posing questions in the NICHD scheme. The main difference is that, in the 
NICHD scheme, option-posing questions may refer to information not previously 
disclosed (as long as no particular response is suggested), in which case they 
would be categorised as yes/no-high risk (and thus leading) or suggestive in the 
SIM scheme. In fact, early studies by Lamb and colleagues even referred to 
option-posing questions as “leading” utterances (Lamb et al., 2008, p. 67). 
However, recent studies using the NICHD scheme vary as to whether they define 
option-posing questions as being about “information not previously disclosed” by 
the child. Some studies have included this phrase in their definition of option- 
posing questions (Cederborg, Alm, Lima da Silva Nises, & Lamb, 2013; Cyr & 
Lamb, 2009; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007; Lamb et al., 2009), 
while others omit it (Brown et al., 2013; Luther et al., 2015; Yi et al., 2016).

Given that differences in definitions of question types may result in 
questions being assigned to a different overall question category, these differences 
need to be kept in mind when comparing results from studies utilising different 
schemes or definitions within schemes. Direct comparisons between some studies 
may not be possible, or may need to be qualified. It is a limitation of the literature 
on interview quality that has not received a significant amount of attention (cf. 
Gagnon & Cyr, 2017; Waterman, Blades, & Spencer, 2001). Therefore, the next 
section provides an overview of what can be concluded generally about open, 
specific, and leading questions. These broad categories correspond to the “thesis 
coding scheme” categories, and the corresponding rows in the NICHD and SIM 
schemes, in Figure 3.1.
Question Types

**Open-ended questions.** The empirical evidence for the superiority of open-ended questions over specific or leading questions is strong. Both laboratory and field studies have shown that responses to open-ended questions are significantly longer, more accurate and detailed than responses to specific questions (Brown et al., 2013; G. Davies et al., 2000; Dent & Stephenson, 1979; Hershkowitz, 2001; Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007; Phillips, Oxburgh, Gavin, & Myklebust, 2012; Sternberg et al., 1996). For example, in one analysis of CSA interview transcripts, responses by 4- to 12-year-old children to open-ended questions were four times longer and had three times as many relevant details as responses by the same children to direct, leading or suggestive questions (Sternberg et al., 1996). Open-ended questions also elicit more forensically-relevant details than other types of questions (Lamb et al., 1996; Lamb et al., 2003; Phillips et al., 2012; Sternberg et al., 1996), and result in more coherent accounts (Feltis, Powell, Snow, & Hughes-Scholes, 2010) with fewer self-contradictions (Lamb & Fauchier, 2001). These last two factors are particularly important in cases where interviews are admitted as evidence-in-chief (Davis et al., 1999), to be subjected to scrutiny by both defence lawyers and jury.

Responses to open-ended questions tap into recall memory and require deeper memory processing and more elaborate memory retrieval than responses to specific questions (Brown et al., 2008; Powell, 2002; Wright & Powell, 2006). Thus, when providing narrative accounts, children usually provide information that they actually remember (Powell, Roberts, & Guadagno, 2007). In comparison, children can attempt to answer specific questions even if they do not
remember the relevant information, and can hide a lack of comprehension by guessing the answer (Powell & Thomson, 2001).

Employing open-ended questions has benefits for interviewers too. Studies have demonstrated that when interviewers use open-ended questions, it reduces the damaging impact of interviewer confirmation bias and encourages the child to adopt an active role in the interview process (Powell et al., 2005; Powell, Hughes-Scholes, & Sharman, 2012). Further, when used effectively, open-ended questions can prompt very specific details about events (Powell et al., 2007). Even very young children and children with mild intellectual disabilities can provide valuable and accurate “Who? What? When?” information in response to open-ended questions (Brown, Lewis, Lamb, & Stephens, 2012; Lamb et al., 2003).

Nevertheless, some individual differences do influence the information that children provide in response to open-ended questions. Some studies suggest that younger children (G. Davies et al., 2000; Dent & Stephenson, 1979; Gagnon & Cyr, 2017; Hutcheson, Baxter, Telfer, & Warden, 1995; Phillips et al., 2012), as well as children with intellectual disabilities (Agnew & Powell, 2004; Brown & Lamb, 2015), can struggle to provide enough detail in response to open-ended questions, necessitating follow-up questions by the interviewer. These groups of witnesses respond better to more focused open-ended prompts (e.g., cued or depth open-ended prompts) that remind them of actions or details previously mentioned (Brown & Lamb, 2015; Gagnon & Cyr, 2017; Lamb et al., 2003).

Thus, not all open-ended questions are of similar quality and usefulness with child witnesses (Powell & Guadagno, 2008). The general term open-ended question includes a wide range of questions, not all of which can be considered “good”, such as suggestive and complex open-ended questions (Powell, Garry, &
Brewer, 2013; Powell & Snow, 2007). Further, “can you” type questions, which implicitly ask the child to provide information (e.g., “Can you tell me what happened?”; “do you know what happened?”) but explicitly ask the child if they know the answer (i.e., “can you tell me” is explicitly a yes/no question asking whether the child knows the information) (Evans, Stolzenberg, Lee, & Lyon, 2014), are less useful with young children who have a very literal approach to language (Walker, 2013). Thus, to encourage the best quality responses, researchers recommend that open-ended questions are phrased using simple, non-coercive language, in a way that encourages an elaborate response, and allows the interviewee to decide which details to include (Powell & Snow, 2007).

**Specific questions.** Here, specific questions refer to cued-recall, forced choice, and yes/no questions that do not mention information that has not been previously disclosed by the child. Interviewing guidelines recommend specific questions are used sparingly because they are associated with increased error rates (Brown et al., 2013; Powell & Thomson, 2001). Specific questions are also more likely to result in inconsistent statements (Lamb & Fauchier, 2001). These inconsistent statements may occur for several reasons. First, in answering specific questions, children tend to report familiar details without thoroughly checking their source (i.e., they adopt a familiarity criterion) (Ibabe & Sporer, 2004; Powell & Snow, 2007). Second, when they do not know the answer, children tend to guess or invent a response due to social pressures, rather than say “I don’t know” (Brown et al., 2008; Powell & Snow, 2007). Finally, more focused specific questions tap into recognition memory, which requires less effort but is also less accurate (when accuracy is defined as the net difference between the number of
correct and incorrect details reported) than recall memory (Dent & Stephenson, 1979; Hutcheson et al., 1995; Ibabe & Sporer, 2004; Lamb & Fauchier, 2001).

Despite these general findings, not all specific questions are equal. Cued recall, or directive, questions are generally preferable to forced choice or yes/no (option-posing) questions. While they produce fewer details than open-ended questions, cued recall prompts still elicit more details than forced choice or yes/no questions (Andrews & Lamb, 2016; Brown et al., 2013; Lamb et al., 2008). The NICHD scheme does not categorise directive questions as either open or closed; instead, it considers them to be on a continuum from open (e.g., “what happened?”) to closed (e.g., “what colour was your dress?”) without clearly defined boundaries (Gagnon & Cyr, 2017; Lyon, 2014). Given that this category (directive questions) thus encompasses questions with very different effects on accuracy, error rates, and amount of information obtained, it is not a useful classification for this thesis. Directive questions are therefore categorised depending on whether specific information was requested, in line with the SIM scheme. For example, “What colour was your dress?” is considered to be a specific (cued-recall) question, while “What happened?” is considered to be an open-ended prompt.

Forced choice and yes/no questions are the least productive specific questions, and should be avoided (Luther et al., 2015). Both types of questions force children to choose their answer from the options provided by the interviewer (Lyon, 2014). They elicit the least amount of detail (Cederborg, Orbach, Sternberg, & Lamb, 2000; Korkman, Santtila, & Sandnabba, 2006), encourage guessing (Luther et al., 2015), and are associated with lower accuracy and more errors and inconsistencies than any other question type (Lamb &
Children and other vulnerable witnesses are particularly susceptible to the negative consequences of these questions (Agnew & Powell, 2004; Powell, Guadagno, et al., 2016; Waterman et al., 2001). The danger these questions pose was well demonstrated in two studies by Waterman and her colleagues, in which children were asked to provide answers to nonsensical yes/no questions (Waterman et al., 2000, 2001). In both studies, most children (aged 6 to 8 and 5 to 9, respectively) attempted to answer the unanswerable yes/no questions, rather than admit that they did not know the answer.

**Leading questions.** Leading questions in this thesis refer to suggestive questions, as well as high-risk yes/no questions and option-posing questions that introduce information that children have not previously mentioned. This broader definition is utilised because introducing information not previously disclosed is generally considered a suggestive technique (Brown et al., 2008). Other examples of suggestive techniques include asking children the same question multiple times, pressuring children to give a particular response, and instructing children to pretend or imagine what might have happened (Brown et al., 2008). These other suggested techniques, if present, were not coded as leading questions in the current thesis.

Children (particularly those under 6 years) are more susceptible to suggestion than adults (Ceci & Bruck, 1995; Hoyano & Keenan, 2010; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007). Suggestibility refers to “the degree to which the encoding, storage, retrieval, and reporting of events can be influenced by a range of internal and external factors” (Ceci & Bruck, 1995, p. 44). Children’s susceptibility to suggestion, and the possibility of contamination of
their evidence through bad interviewing techniques, has been a key concern for
courts in the context of children’s testimony (Hoyano & Keenan, 2010). This
concern is particularly important, as one suggestive interview is enough to
contaminate children’s recollections of events, and children can report false
answers given in response to initial suggestive questions in subsequent non-
suggestive interviews (Brown & Lamb, 2015; Ceci, Kulkofsky, Klemfuss,
Sweeney, & Bruck, 2007). Suggestive questioning can also lead to false
allegations, reduce children’s credibility, lead to inconsistent statements, and
result in unsuccessful prosecutions (Bull, 2010; Lamb & Fauchier, 2001; Wood,
2000). Therefore, suggestive questions should never be used in investigative
interviews.

**Adherence to Best Practice Interviewing**

The previous section illustrated the importance of interviewers employing
certain question types, and avoiding others, in accordance with best-practice
guidelines. Despite this importance, adherence to best practice questioning by
interviewers in the field is generally poor (Lamb, 2016; Luther et al., 2015;
Powell, Cavezza, Hughes-Scholes, & Stoove, 2010; Powell, Murfett, et al., 2010;
Wolfman, Brown, & Jose, 2016). Interviewers tend to ask many specific
questions and too few open questions, particularly when the child does not
disclose abuse or does not disclose specific details (Hughes-Scholes & Powell,
2013). Further, interviewers tend to regard factors external to themselves, such as
the child or the setting, as more important for a successful interview than their
own questioning skills (Lamb, Orbach, Hershkowitz, Esplin, et al., 2007; Powell,
Murfett, et al., 2010; Wright, Powell, & Ridge, 2007).
A good example is a study by Wright and colleagues, in which police officers’ perceptions of CSA interview quality and success were examined by asking 75 officers to evaluate their own performance in mock interviews with a 5-to 7-year-old child (Wright et al., 2007). Results demonstrated that the wide range of interviewer performance (including open-ended question usage, which ranged from 6% to 79%) was not associated with corresponding variability in how interviewers believed they performed. Further, 92% of the interviewers equated success in the interview with whether the child talked about the event. Open-ended questions were only referred to by 8% of officers in their evaluations. The researchers concluded that, consistent with prior studies, police interviewers undervalued open-ended questions and considered the success of the interview to be dependent on the child.

The proportion of different question types utilised by interviewers in field interviews varies greatly, likely due to the differences in definitions of question types employed, the training received, and the justice systems (adversarial or inquisitorial) in the countries in which the research was conducted. It is thus difficult to compare findings across studies. This difficulty is true even for studies purporting to use the same protocol. Definitions differ, for example, in studies employing the NICHD scheme on whether option-posing questions must focus on information not previously mentioned (cf. Cederborg et al., 2013; Luther et al., 2015), and whether suggestive questions are only those in which the interviewer “strongly communicates” the information that is required, or whether simply introducing information not previously disclosed by the child is enough (cf. Gagnon & Cyr, 2017; Luther et al., 2015; Price & Roberts, 2011). Whether such differences are only due to lack of information being provided about the coding
scheme in the study, or whether questions were actually coded differently, is not clear.

Well-established question categories are sometimes also split in new ways. For example, some studies have introduced new categories of invitations (Gagnon & Cyr, 2017; Price & Roberts, 2011), and directive questions (Gagnon & Cyr, 2017; Price & Roberts, 2011; Thoresen, Lønnum, Melinder, & Manussen, 2009). While the rationales for the question categories employed in studies are usually sound, it nevertheless results in a situation in which it is hard to determine the “usual” rates of open, closed, and leading questions used by interviewers in the field.

The most consistently defined question type, and the one that differentiates best between effective and less effective child witness interviewers, is the open-ended question (Powell et al., 2005). While a recent study quotes the rate of open-ended questions in field studies as approximately 25% (Powell, Guadagno, et al., 2016), the studies cited by the author in support of this figure are approximately one to two decades old. More recent research conducted over the last five years in countries with adversarial legal systems suggests that the rates of open-ended questions in field interviews ranges from 4% to 57% for experienced interviewers (that is, interviewers who have received some form of training) (Benson & Powell, 2015b; Gagnon & Cyr, 2017; Luther et al., 2015; Price & Roberts, 2011; Wolfman et al., 2016). Studies from countries with inquisitorial systems have not been included in this figure because the function of investigative interviews differs between adversarial and inquisitorial systems (as discussed in Chapter 2), and a much wider range of professionals (e.g., judges,
lawyers, psychologists) conduct these interviews in inquisitorial countries (Cederborg et al., 2000; Johnson et al., 2015; Thoresen et al., 2009).

**Interviewer Training**

In an attempt to improve interviewer performance, a significant amount of research has been conducted to examine training programs for investigative interviewers. This research has demonstrated that short, intensive training courses may increase interviewers’ knowledge and help to eliminate major interviewing errors such as leading questions, but have little impact on less-obvious errors and long-term interviewing skills (Jones, Cross, Walsh, & Simone, 2005; Lamb, 2016; Lamb, Sternberg, Orbach, Esplin, & Mitchell, 2002; Powell et al., 2005). To improve interviewing skills in the long term, training programs should be distributed over several days with significant breaks in-between (Powell et al., 2005); should include opportunities to practice interviewing skills (Hughes-Scholes & Powell, 2013); should include concrete examples of questions (Hughes-Scholes & Powell, 2013; Lamb, Sternberg, Orbach, Hershkowitz, et al., 2002); and most importantly, must include continual supervision and individual feedback (Benson & Powell, 2015b; Cyr, Dion, McDuff, & Trotier-Sylvain, 2012; Jones et al., 2005; Lamb, 2016; Orbach et al., 2000; Powell, 2002). Training programs should also employ a structured interview protocol to help trainees adhere to best practice interviewing, such as the SIM or NICHD protocols (Lamb, Orbach, Hershkowitz, Esplin, et al., 2007; Lamb et al., 2009).

One recent study has illustrated that training can be effectively administered using web-based activities. The study, which evaluated a training model combining evidence-based, web-based activities with continual (face-to-face or skype) supervision, showed that significant improvements in usage of
open-ended questions by interviewers can be achieved without the need for traditional classroom style workshops (Benson & Powell, 2015b). Post-training interviews were also significantly shorter on average (41.30 mins) than pre-training interviews (65.49 mins). Most importantly, these improvements in interview quality were maintained in follow-up interviews conducted three to six months after training ceased, and were achieved without a decrease in the amount of key evidential information sought.

The “Dual Use” of the Interview

Most research discussed so far has focused on the use of the investigative interview pre-trial. However, the interview is not only used for investigative purposes; if a CSA case proceeds to trial, the interview will likely constitute the main piece of evidence for the prosecution. Combining the need to obtain information from a witness for both evidentiary and investigative purposes into one interview is useful in preserving the witness’ evidence, as information is obtained at an early stage in the process (Freckelton & Selby, 2013). However, this “dual use” of the investigative interview also causes tensions, as different information is required for investigative and evidentiary purposes (Freckelton & Selby, 2013; Powell, Wright, & Hughes-Scholes, 2011). While the emphasis during investigation is on getting as much information about an event as possible (including forensic details), the emphasis during evidence-in-chief is to establish that an offence has occurred by eliciting only relevant and admissible information (Westera, Kebbell, & Milne, 2013b). Additionally, the type of questioning employed during evidence-in-chief is very different to the questioning used in a child witness interview. Prosecutors tend to ask short, specific questions in an attempt to control the witness’ account, while police (ideally) give more control to
the witness and ask open-ended questions in an attempt to encourage a narrative
account (Powell et al., 2013; Westera et al., 2013b).

These tensions between the different purposes of the investigative
interview cannot be completely avoided or overcome (Powell et al., 2011; Wilson
& Davies, 1999). Attempting to reconcile them would place unrealistic demands
on interviewers (Davis et al., 1999). However, given the dual use of the interview,
both child testimony experts and legal professionals should be involved in
developing training programs for child witness interviews, and should engage in
greater inter-agency collaboration (Burrows, Powell, & Anglim, 2013; Powell et
al., 2011). Lamb (2016) suggested that legal professionals should also be trained
to recognise poor practice in interviewing so that they can apply pressure to force
improvements in interviewing practice. Legal professionals’ current involvement
with investigative interviews, and the benefits and limitations of using a pre-
recorded investigative interview as evidence-in-chief, will be discussed next.
CHAPTER 4: A REVIEW OF THE RECEPTION OF CHILDREN’S EVIDENCE AT TRIAL – TRADITIONAL ATTITUDES AND REFORM

This chapter provides an overview of the barriers that CSA complainants face when their sexual abuse claims are prosecuted, and the reforms that have been enacted in recent years to facilitate complainants’ involvement in trials. Again, this chapter is limited to literature from countries with adversarial legal systems. It begins by briefly summarising how sexual assault and child complainants have traditionally been regarded by the courts, and the problems they faced giving evidence, to establish why reforms were necessary. It then describes the different ways in which complainants may now give evidence in these cases, before examining in detail the benefits and limitations of the reform central to this thesis: the pre-recording of child’s investigative interviews to be used as their evidence-in-chief at trial.

Problems Faced by Sexual Assault and Child Witnesses Prior to Reforms

Evidence from sexual assault victims and from children has historically been regarded as unreliable in the British - and thus also Australian - legal system (Australian Law Reform Commission, 1997; G. Davies, 1994). This assumed unreliability is evident in traditional common law rules of evidence and procedure such as the competency requirement (a test to determine whether a child is capable of functioning as a reliable witness); the corroboration rule (that a complainant’s evidence must be corroborated because it is “dangerous” to convict on the uncorroborated evidence of a child or sexual assault victim); and judicial cautions (which require judges to warn juries that children and sexual assault complainants are an unreliable class of witness and their evidence must be treated
with caution) (Australian Law Reform Commission, NSW Law Reform Commission, & Victorian Law Reform Commission, 2005; Cossins, 2006; G. Davies, 1994). It is also evident in the undue focus placed on the complainant’s credibility in these cases (Cossins, 2006). The mandatory warnings about the unreliability of classes of victims and their evidence have now been abolished (Australian Law Reform Commission, 2010), but judges still have the discretion to give warnings “necessary to avoid … a miscarriage of justice” (*Longman v. The Queen* (1989), p 86).

In addition, the adversarial legal system is a poor fit for sexual assault cases (Australian Law Reform Commission et al., 2005; Shead, 2014). In the adversarial legal system, a criminal trial is not an inquiry into the truth; rather, it is a forum in which two competing parties “battle” to present their evidence in front of an impartial judge and decision-maker: the jury (Cossins, 2006, 2010; Shead, 2014). It is a forum in which the best arguments and evidence – not necessarily the truth – prevails (Cossins, 2010). The focus in an adversarial trial is very much on oral evidence and the testing of this evidence by oral means (i.e., cross-examination), in a way that ensures the accused has a fair trial (Cossins, 2010). The traditional aim of the process has not been to protect victims or children from further abuse (Cossins, 2006). In fact, comments by defence lawyers in a 1998 Queensland study such as, “if in the process of destroying the evidence, it is necessary to destroy the child, then so be it” and, “it would be considered cowardly not to go for the jugular when cross-examining a child” demonstrate that victims have been regarded by some lawyers merely as a means to advance their case or destroy their opponents’ (Eastwood, Patton, & Stacey, 1998, p. 3). CSA trials were thus described as a “cruel and horrible” process
resulting in great stress and anxiety for vulnerable witnesses, and with a great potential for re-victimisation (Eastwood et al., 2006, p. 85; Eastwood & Patton, 2002; Hamlyn, Phelps, Turtle, & Sattar, 2004).

Reforms to the Giving of Evidence by Vulnerable Witnesses

Recognition of difficulties faced by CSA complainants and other vulnerable witnesses have led to the introduction of procedural reforms that allow these complainants to give evidence in alternative ways (Friedman & Jones, 2005). These “alternative measures” include giving evidence via CCTV, giving evidence behind a screen, using pre-recorded investigative interviews as evidence-in-chief, and pre-recording the entirety of witnesses’ testimony (including cross- and re-examination) at a preliminary hearing.¹ The remainder of this chapter focuses on the use of the pre-recorded child witness interview as evidence-in-chief.

The use of the child witness interview as evidence-in-chief. All Australian jurisdictions now allow children’s pre-recorded child witness interviews to be admitted as part or whole of their evidence-in-chief (see e.g., Criminal Procedure Act 1986 (NSW) ss. 306R, 306U; Criminal Procedure Act 2009 (Vic) ss. 366-368). The record of the interview may be edited (with a court order), and prosecutors may also ask additional (“supplementary”) questions of the complainant at trial if certain necessary evidentiary details are absent from the child witness interview (Australasian Institute of Judicial Administration, 2012; Freckelton & Selby, 2013). However, in all jurisdictions apart from the Northern Territory, this special measure is not applicable to adult witnesses (ALRC, 2010).

¹ For a detailed review of special measures, please see Appendix A of Powell et al., 2016.
It is therefore not available to complainants of historical CSA, who must continue to give evidence live in court or using other alternative measures such as CCTV.

**Benefits.** There are numerous benefits of using the child’s investigative interview as their evidence-in-chief. One of the most significant benefits is the effect that this special measure has on the stress and trauma experienced by the child (Corns, 2001). It is well documented that the traditional trial process is highly distressing for child witnesses (Australasian Institute of Judicial Administration, 2012; Henderson, 2011; Shead, 2014; W. A. Walsh et al., 2010). In one study, 8% of cases were dropped by the prosecution because of concerns that the child could not handle the trial (Gallagher, 1999). In another study, two-thirds of legal professionals interviewed stated that they would not want their own child to go through the trial process because of the trauma the child would suffer (Eastwood et al., 2006). However, experimental studies have demonstrated that when children pre-record their testimony, they are significantly less nervous (Landström & Granhag, 2010) and stressed (Eastwood et al., 2006) than when they testify live. This finding has been replicated in a field study in which judges, barristers, police officers and social workers rated children testifying on video as significantly less anxious than children testifying live in court (G. Davies, Wilson, Mitchell, & Milsom, 1995).

The reasons why victims find pre-recording their evidence helpful was demonstrated in a large and comprehensive research project undertaken in the UK. This research investigated the effectiveness and use of alternative measures for vulnerable victims (Burton, Evans, & Sanders, 2006; Hamlyn et al., 2004). It involved interviews with 552 vulnerable and intimidated witnesses before, and 569 vulnerable and intimidated witnesses after, the introduction of the *Youth*
Justice and Criminal Evidence Act 1999 (which aimed to improve the process of giving testimony by vulnerable and intimidated witnesses). It also involved national surveys of the police, crown prosecution service (CPS), witness service, and the crown courts, and interviews with practitioners from each of these four agencies. The researchers also examined the decision-making processes in 191 cases, and observed 26 contested cases in court. The findings indicated that video-recorded evidence-in-chief was the most used alternative measure (at 42%). The vulnerable and intimidated witnesses found using the pre-recorded interview helpful because it meant they did not have to appear in court (43%), it was easier to say things (22%), they were less scared (13%), it helped them to remember (12%), and they were in a more comfortable environment (9%).

Additional benefits of using pre-recorded interviews as evidence-in-chief include witnesses not having to repeat their story multiple times (which may be damaging both to the child’s wellbeing and the accuracy of their account: Ghetti, Alexander, & Goodman, 2002; La Rooy, Lamb, & Pipe, 2009); that they can use the recording to refresh their memories before trial; that they do not have to wait months or years to give evidence; and that their evidence is captured close to the time of the offence (Corns, 2001; Hamlyn et al., 2004). The latter is particularly important for child witnesses as children’s memories of details erode more quickly than adults’ (Henderson, 2011; Peterson, 2011), and pre-recording their evidence early thus results in more detailed, complete and reliable evidence (Burrows & Powell, 2014d; Burton, Evans, & Sanders, 2006). Recording their evidence early also reduces the chances of it being contaminated by post-event information (Corns, 2001).
There are also procedural advantages of using pre-recorded interviews as a child’s evidence. First, because it can be edited before trial, any inadmissible or irrelevant evidence can be edited out before it is presented to the jury, preventing possible mistrials (AIJA, 2012; Corns, 2001). Second, the availability of the child’s evidence prior to trial allows the prosecution and defence to better prepare their cases (Burrows & Powell, 2014d; Cashmore & Trimboli, 2005). These advantages do not appear to come at any material cost to the defendant, with empirical studies suggesting that the use of this special measure does not significantly impact the jury’s perception of the defendant (Eaton, Ball, & O'Callaghan, 2001; Swim, Borgida, & McCoy, 1993) or conviction rates (Cashmore & Trimboli, 2005; G. Davies et al., 1995; Hamlyn et al., 2004).

**Limitations.** Despite the benefits that complainants seem to experience through the use of pre-recorded interviews as evidence-in-chief, there are also significant limitations to this special measure. For instance, juries appear to perceive children giving evidence on video in a less positive manner than children testifying live. In two studies, Landström and Granhag (2007, 2010) found that children (aged 10-11) testifying on video are perceived as providing less confident and convincing statements than children testifying live. Studies have also shown that children testifying via pre-recorded statement are less persuasive, are regarded as less honest, and evoke less empathy and emotion than children testifying live (Burton et al., 2006; E. Davies & Hanna, 2013; Goodman et al., 2006; Landström, Granhag, & Hartwig, 2005; Westera, Kebbell, & Milne, 2013a). However, video-recorded statements do not appear to have any actual impact on jurors’ ability to determine children’s veracity or the accuracy of their
statement (Goodman et al., 2006; Landström et al., 2005; 2007), which is poor irrespective of evidence presentation mode.

These findings that children seen on video are perceived in a less positive manner than children seen live may be due to problems observing body language and other visual cues on a screen (Burton et al., 2006; Doherty-Sneddon et al., 1997), or the perceived distance between child, judge and jury when recorded evidence is employed (Cashmore & Trimboli, 2005; Pipe & Henaghan, 1996). They may also be due to reduced juror engagement with the witness, or reduced clarity of the evidence due to audio-visual problems with the recording (Burrows & Powell, 2014d). These perception issues are problematic, particularly in CSA cases where the lack of corroborating evidence means that juries use a witness’ behaviour and responses in assessing the veracity of their claims (Freckelton & Selby, 2013).

A further significant disadvantage to using pre-recorded evidence for children is that these complainants then face cross-examination “cold” (Burrows & Powell, 2014d; Burton et al., 2006; G. Davies et al., 1995; Tinsley & McDonald, 2011). When complainants give their evidence-in-chief live in court or via CCTV, they have the chance to get used to being questioned by a sympathetic party first (i.e., the prosecution) before they are faced with cross-examination. In contrast, where their evidence-in-chief is recorded months or even years prior, their first experience of the trial is usually cross-examination (the exception being where supplementary questions are first asked by the prosecutor). They are thus less prepared for cross-examination, and this reduced preparedness may have a significant impact on children’s ability to withstand the
“hostile” questioning style of cross-examination and may lead to their credibility being undermined (Burton et al., 2006).

The final disadvantage of using the pre-recorded child witness interview as evidence-in-chief is the poor evidential quality of many child witness interviews (Burrows & Powell, 2014b; Powell, Wright, & Clark, 2010; J. M. Read & Powell, 2011). The impact of child witness interview quality on the conduct and outcome of a CSA trial is the key focus of this thesis. Limited research has examined how the child witness interview influences a trial. However, the literature that exists to date is described in the following chapter.
CHAPTER 5: A REVIEW OF THE RESEARCH ON EVIDENTIAL QUALITY OF CHILD WITNESS INTERVIEWS

Over the last two decades, procedural reforms and an increased awareness of CSA has (in most jurisdictions for which data is available) resulted in an increase in the number of people charged and prosecuted for CSA (Eastwood et al., 2006; Cashmore et al., 2016). However, this increase in prosecution has, overall, not resulted in a corresponding increase in conviction rates (Eastwood et al., 2006; Lievore, 2005; Cashmore et al., 2016). One possible reason for this disparity between charging and conviction rates is the continuing evidentiary difficulties faced in these cases, including the evidentiary quality of child witness interviews (Burrows & Powell, 2015). Investigative interview protocols have been developed mainly by developmental psychologists with the aim of maximising the accuracy of a child’s account (Benson & Powell, 2015a). Legal professionals have generally not been involved in their development, and limited research exists on the evidentiary quality of child witness interviews. This chapter examines the key evidentiary issues in child witness interviews that have been identified by the studies conducted to date, most of which were conducted with Australian prosecutors. It also describes some of the recommendations from prosecutors about how these issues could be addressed. The chapter concludes with a discussion of what is currently known about the relationship between child witness interview quality and case outcomes, finishing with justifications for the current Thesis.
**Style of Questioning**

As mentioned previously, interviewers have tremendous difficulty adhering to best-practice questioning. The available research suggests that this deviation from best practice has significant negative consequences for child complainants at trial. For example, a thematic analysis of 36 phone interviews conducted before and after a CSA trial with 19 prosecutors from one Australian jurisdiction revealed that lengthy interviews with elaborate details were considered highly problematic, as defence barristers can cross-examine a child about any issue or detail that is raised in the interview (Burrows & Powell, 2014b). Prosecutors believed that the more minute or irrelevant details that are raised, the more likely it is a child will provide an inconsistent or inaccurate response. These inconsistencies or errors can then be used by the defence to raise doubt about the child’s reliability and credibility. As a result, prosecutors viewed both broad open-ended questions that elicited lengthy and irrelevant details, as well as specific and closed questions that are likely to lead to inaccuracies, as problematic. Overall, child witness interviews were seen as most useful when they were concise, clear, and relevant, and the evidence elicited in a narrative format.

**Amount of Detail Sought**

Closely related to the style of questioning is the amount of detail that is sought by police through their questioning tactics. Interviewers must be extremely careful about what detail is sought and followed up, as too much detail may undermine the evidential quality of the child’s account, while not enough detail on the nature and circumstances of the offence can result in charges being dropped (Burrows & Powell, 2015). Interviews and focus group studies with
Australian prosecutors suggest that interviewers tend to engage in the former, and seek minute details using very specific questions where these details are not required for prosecution and are likely to be inaccurate (Burrows & Powell, 2013, 2014a, 2014b, 2014d, 2015). As well as leading to problems in cross-examination (as described above), continuous questioning about minute and irrelevant details is fatiguing for the child and the jury. A fatigued child is likely to become uncooperative, distracted and irritable and perceived as less reliable in the eyes of a jury (Burrows & Powell, 2014b), while a fatigued jury is likely to be less attentive to the child’s testimony.

Overall, prosecutors recommended that interviewers focus their questioning around the elements of an offence (that is, the aspects of an offence contained in legislation that must be proven for the offence to be established), and seek only enough detail to prove a charge “beyond reasonable doubt” (Burrows & Powell, 2013, 2014c). Interviewers should also clarify inconsistencies and ambiguities in a child’s account, while being careful not to cross-examine the child (Burrows & Powell, 2014d). Australian prosecutors have identified certain areas that are specifically problematic in child witness interviews, and have provided specific recommendations on how to improve questioning around these topics.

One of these problematic topics is questioning around alleged sexual acts. A focus group study involving nine prosecutors from every Australian State and Territory (except one small jurisdiction) revealed that specific questioning to obtain a “frame-by-frame” description of sexual acts is extremely unhelpful from the prosecution perspective (Burrows & Powell, 2015). Rather, interviewers should focus on elicitating information that helps to determine the nature of the
alleged acts. In deciding whether follow up questions are necessary, interviewers should consider whether the nature of the act is sufficiently clear from the child’s account (e.g., would a juror understand which sexual act is meant by the child?), the developmental age of the witness (e.g., does the child have the ability to provide further information?), and whether there is corroborative evidence available to support the child’s account. Corroborative evidence, such as the presence of semen, would lessen the need for follow-up questions regarding the sexual act.

A second problematic area is the clarification of terms for genitalia. Interviewers are often overzealous in their attempt to get the child to use a specific anatomical term. However, from the prosecution perspective, the terms for genitalia used by children when describing abuse are usually sufficiently clear (Burrows & Powell, 2014a). In fact, the naïve way children use genitalia terms usually acts to enhance the credibility of their account, rather than detract from it. The only instances in which terms for genitalia should be followed up are those in which a layperson would likely not understand what is meant.

**Particularisation**

A final key problem area in child witness interviews from the evidential perspective is questioning around particulars. Particulars refer to the how, when, and where of an offence. In order for an offender to be charged, particulars must be established but they do not need to be proven to the same degree as the elements of an offence (Burrows & Powell, 2014d; Shead, 2014; S v. The Queen (1989)). The availability of particulars is necessary for procedural fairness, as defendants must have enough information about the alleged offence to be able to mount a defence. However, particularisation is a complex area of law and there is
considerable confusion and uncertainty both across and within disciplines regarding specifically what information is required (Guadagno, Powell, & Wright, 2006). This uncertainty about what is required is evident in police investigative interviews.

Studies involving interviews with legal professionals suggest that child witness interviewers place too much emphasis on particulars (Burrows & Powell, 2014d; Guadagno et al., 2006). Whereas legal professionals (including prosecutors, defence counsel and a judge) considered highly specific details about location, date and time unnecessary for particularisation, it seems that police professionals believed such specific details are crucial for particularisation to be achieved (Guadagno et al., 2006). Further, while police officers tended to believe that “the more the better” (in terms of details and offences), legal professionals considered too many details damaging to a successful prosecution. This is because these professionals perceived that asking too many specific questions in an attempt to elicit specific details increases errors and inconsistencies, which then undermines children’s accounts and damages their credibility (Burrows & Powell, 2014d; Guadagno et al., 2006).

There are two circumstances in which particularisation is especially an issue. The first is in historic cases of abuse, that is, adult complainants reporting childhood sexual abuse. Because particulars require the complainant to specify when an offence occurred, it can be difficult to sufficiently particularise an alleged offence if a complainant can specify what conduct occurred, but due to the passage of time is unable to specify exactly when it took place. It is very unlikely, for example, that a complainant could say that an offence occurred “on my 12th birthday”; it is more likely for a complainant to state something like
“when I was in 4th class and Mrs Dyers was my teacher [an offence] happened” (Shead, 2014, p. 61). This difficulty in particularising historic abuse can frequently be overcome by charging “between dates”; that is, nominating two dates between which an offence is alleged to have occurred (Shead, 2014). Much more difficult to address are the particularisation issues that arise in cases of repeated abuse.

CSA frequently involves repeated incidents of abuse over long periods of time that are similar in nature (Guadagno et al., 2006). In such cases, children often fail to maintain distinct memories of each separate occurrence of abuse and establishing the necessary particulars can therefore be difficult (Powell et al., 2007). One tool that may help children identify specific instances of abuse is the use of labels. Labels for abuse incidents can either be temporal (“the last time”), locational (“the time at the swimming pool”) abuse related (“the time he put his hand down my pants”) or situational (“the time mum was at the movies”) (Brubacher, Malloy, Lamb, & Roberts, 2013). Labels are useful because they help clarify which instance of abuse is being described, help keep the complainant focused, and emphasise the need for details specific to one event (Powell et al., 2007). Best practice recommendations are for interviewers to encourage children to create their own labels (Powell et al., 2007), but this recommendation is not always followed.

For example, one study that examined CSA interviews of 81 girls and 16 boys aged 5 to 13 who alleged repeated instances of sexual abuse found that in 66% of interviews in which children provided at least one label, this label was either ignored or replaced by the interviewer (Brubacher et al., 2013). Further, this study found that interviewers used significantly more temporal labels than
children, and the more interviewers ignored children’s labels or replaced them with their own, the fewer event details were reported by children. Given that contextual cues are more effective in facilitating recall than temporal cues, and labels are more effective if generated by the children themselves (Powell et al., 2007), this finding is not surprising. It is important, therefore, for interviewers and other professionals questioning child complainants about repeated events (e.g., prosecutors, defence counsel) to use labels for abuse instances that the child provided, and to use these consistently.

**The Relationship Between Best Practice Interviewing and Case Outcomes**

When discussing the relationship between interview quality and outcome, psychological researchers have suggested that employing good quality interviews in court should increase the rates of “successful prosecutions” of CSA offenses (Burrows & Powell, 2015; Pipe, Orbach, Lamb, Abbott, & Stewart, 2013). In these studies, the term “successful prosecution” is used to refer to a guilty verdict, without any qualifications. This is understandable given that research has demonstrated that CSA offences have low prosecution and conviction rates in comparison to other offences, and one of the reasons posited for this is insufficient or poor quality evidence (see previous chapters). However, it must be acknowledged that employing good quality evidence in CSA trials may also lead to acquittals – if the person charged is innocent. Thus, in this thesis, “successful prosecution” is used to refer to a just outcome, where defendants are found guilty who ought to be, and cases do not result in “not guilty” verdicts on the basis of insufficient or poor evidence.

Only one study from an adversarial legal system has, to date, examined what influence (if any) interview quality has on CSA outcome. Pipe and
colleagues (2013) compared a sample of 350 CSA cases from the US in which investigative interviews were conducted with 3-14 year old victims before the introduction of the NICHD protocol (the “pre-protocol cases”) with 410 CSA cases in which investigative interviews were conducted after the introduction of the NICHD protocol (the “protocol cases”). The researchers found that although pre-protocol cases were no more likely than protocol cases to produce one or more guilty pleas, police laid more charges in protocol cases than pre-protocol cases. Thus, overall, protocol cases were associated with more guilty pleas than non-protocol cases. Additionally, at trial, protocol cases were more likely to lead to guilty verdicts than pre-protocol cases. These findings suggest that interview quality may influence case processing and outcome.

However, the study suffers from a number of significant limitations. These include the fact that only 23 cases were decided at trial, limiting the generalisability of the finding regarding trial outcomes; the fact that corroborative evidence was not controlled for; and the fact that the actual quality of the interviews was not evaluated in the study. Given that there are significant difficulties with the application of best-practice interviewing skills learned during training into the field (Lamb, 2016), it does not necessarily follow that interviews in the post-protocol condition were better in quality than interviews in the pre-protocol condition. Such a conclusion could only be drawn if at least a subset of interviews in each condition had been coded for interview quality. Thus, further research is needed to examine the relationship between interview quality and case outcomes.
Conclusions and Justification for Current Thesis

The research on the evidential quality of child witness interviews, and how this quality relates to trial processes and outcome, is still in its infancy. Most research to date has been conducted through interviews and focus groups with prosecutors and other legal professionals. This research has revealed that the prosecutors are concerned about a number of aspects of interview quality, including the amount and type of detail that is being sought by interviewers. These prosecutors have also suggested that interview quality is affecting their chances of a successful prosecution. However, only one study has to date examined whether interview quality influences CSA case outcomes, and this study has significant limitations. Further research is thus needed to establish whether child witness interview quality is related to trial processes and outcomes, and more precisely, the aspects of child witness interviews that are most problematic at trial. This thesis aims to address this gap in the literature. The studies presented in the next four chapters expand on the research presented in a larger project conducted for the Royal Commission into Institutional Responses to Child Sexual Abuse (‘RC’) (Powell, Westera, Goodman-Delahunty, & Pichler, 2016). The larger project investigated the use of alternative measures in child sexual abuse cases more generally; the studies presented in this thesis examine the relationship between the child witness interview and trial processes in more detail, moving beyond the findings presented in the RC report.
CHAPTER 6 – CHARACTERISTICS OF CHILD SEXUAL ABUSE TRIALS: EXPLORING DATA FROM THREE AUSTRALIAN STATES (STUDY 1)

This chapter presents the first study of this thesis. The study examined the type of CSA cases being prosecuted in our courts, and more specifically, the sample of CSA cases that will form the basis of Studies 2 to 4. Although CSA trials are highly confidential, and obtaining detailed trial data is extremely difficult (McConachy, 2002), there are two broad bodies of literature that provide some insight into the characteristics of CSA cases in our criminal justice system. One body focuses on attrition; that is, identifying the cases that leave the criminal justice system, at what points, and for what reasons. The other body focuses on factors related to likelihood of prosecution and trial outcome. Together, these bodies of research suggest that there are certain characteristics that make it more or less likely for cases to reach trial and result in conviction. These characteristics include the presence of corroborative evidence, age and gender of the complainant, the delay between the abusive event(s) and trial, and offence type and seriousness. Each of these characteristics are reviewed below in order of the amount of research evidence available.

One consistent finding in the literature is that cases are more likely to have charges laid and be accepted for prosecution if they have corroborative evidence (Davis et al., 1999; Gallagher, 1999; Powell, 2008; Cross & Whitcomb, 2017). Evidential considerations are key in these cases (Duron, 2018); insufficient evidence is the stated reason for not proceeding with prosecution in around 40%
of CSA cases (Christensen et al., 2014). This finding is not restricted to older research, despite the fact that corroboration of children’s evidence is no longer a legal requirement, and the perception by some Australian prosecutors that “oath against oath” cases are now more likely to be accepted for prosecution (Success Works, 2011). Examples of the type of corroborative evidence that have been linked with an increased likelihood of prosecution include confessions (Bradshaw & Marks, 1990; Cross et al., 1994; W. A. Walsh et al., 2010), eyewitness evidence (Cross et al., 1994), physical evidence (Cross et al., 1994), and multiple victims (K. D. Brewer, Rowe, & Brewer, 1997; W. A. Walsh et al., 2010).

Conversely, the support for medical evidence is mixed: Some studies have found that its presence increases prosecution and conviction rates (Bradshaw & Marks, 1990), others have found no relationship (Cross et al., 1994), while still others found that medical evidence only increased prosecution rates in a specific subset of cases (K. D. Brewer et al., 1997). These inconsistent findings may be due to different legal requirements in the jurisdictions examined, or differences in the study samples (for example, Brewer and colleagues (2007) did not include “unsubstantiated” cases in their sample).

Age of complainant, gender of the complainant, and reporting delay have also been linked to the likelihood of charging and prosecution. Research suggests that young children (Bunting, 2008; Leach, Powell, & Anglim, 2016; Stroud, Martens, & Barker, 2000) and, in some studies, teenagers (Bunting, 2008; Leach et al., 2016) are less likely to have their cases referred to prosecution than children aged 7-12, while cases are more likely to be accepted for prosecution if complainants are older children (K. D. Brewer et al., 1997; Cross et al., 1994; Fitzgerald, 2006). Male complainants appear less likely to have their cases
charged and prosecuted than female complainants (Edelson, 2012; Fitzgerald, 2006; Stroud et al., 2000). In terms of delay, the majority of the literature suggests that cases are less likely to reach trial the longer the delay between abuse and report (Bradshaw & Marks, 1990; K. D. Brewer et al., 1997; Bunting, 2008; Fitzgerald, 2006), although delay may differentially affect different age groups of complainants (Bunting, 2014).

In terms of offence type and seriousness, and the complainant’s relationship to the offender, research has shown mixed results. Seriousness of abuse and aggravating factors appear to increase the likelihood of prosecution (Cross et al., 1994; Fitzgerald, 2006; Parkinson et al., 2002), although this relationship is not supported by all studies (K. D. Brewer et al., 1997; Stroud et al., 2000). The influence of the complainant’s relationship to the offender is hard to evaluate, as studies have not used the same variables to examine this factor and results are inconsistent. Some studies suggest that cases are less likely to result in a charge or be prosecuted if the offender is closely related to the complainant (K. D. Brewer et al., 1997; Cross et al., 1994; Stroud et al., 2000) or known to them (Fitzgerald, 2006), and others suggest that knowing the offender decreases the likelihood of cases being forwarded for prosecution (Bunting, 2008).

One factor that has not been examined in the literature to date is the effect of the evidential quality of child witness interviews on charging and prosecution decisions (Powell, 2008). This gap in the research is despite the fact that child witness interviews form the central piece of evidence in most CSA prosecutions (Powell, Murfett, et al., 2010; Cross & Whitcomb, 2017). Only one US study has suggested that interviews that were conducted after the implementation of a best practice interview protocol resulted in more charges being laid, and consequently
more guilty verdicts (Pipe et al., 2013). However, this study did not examine the
evidential quality of the interviews in their sample, nor the adherence to the
interview protocol by interviewers, despite the well-known problem of low
adherence to interview protocols in the field (Lamb, Orbach, Hershkowitz, Esplin,
et al., 2007). A recent study suggests that evidential quality of child witness
interviews can be assessed by examining five factors: interviewer questions,
interviewer behaviour, adherence to the interview protocol, interview length, and
investigative questions (Benson & Powell, 2015b). This method of assessing
interview quality is broader than traditional methods of evaluation, which have
focused purely on question types, and have overlooked prosecutors’ concerns
about evidential issues with interviews (Benson & Powell, 2015b; Burrows et al.,
2013).

In terms of CSA trial characteristics, conviction rates in CSA cases in
Australia have been shown to range from 10 to 17% of cases reported to police, or
44 to 62% of cases accepted for prosecution (Australian Law Reform
Commission, 2010; NSW Bureau of Crime Statistics and Research, 2017;
Cashmore & Trimboli, 2005; Cashmore et al., 2016; Fitzgerald, 2006;). Other
trial characteristics, such as length of cross-examination, have not been examined
in these studies. Cross-examination can have a large impact on complainant
credibility (Caruso, 2012), which is critical in most CSA prosecutions (Australian
Law Reform Commission, 2010; Davis et al., 1999; Powell, 2008). There is some
anecdotal evidence that despite reforms requiring judges to stop repetitive,
harassing and other forms of inappropriate questioning (see e.g., s. 41 Evidence
Act 1995 (NSW), s. 41 Evidence Act 2008 (VIC)), judges are reluctant to do so,
and complainants often endure very long and unfocussed cross-examinations
(Shead, 2014). However, little quantitative data is available in support of such assertions. Studies providing an in-depth examination of the characteristics of CSA cases that have reached trial are also absent.

Given that the ALRC has concluded that the collection and evaluation of comprehensive statistics on sexual assault cases and outcomes is critical to identifying and responding to problems in how the criminal justice system deals with sexual assaults (Australian Law Reform Commission, 2010; Australian Law Reform Commission & NSW Law Reform Commission, 2010), and the County and Children’s Courts of Victoria have declared data collection on such cases to be an “urgent priority” for effective policy development (Magistrates’ Court and the Children’s Court of Victoria, 2010), research is needed to address this gap. Thus, the current study aims to provide a detailed and up-to-date description of the CSA cases currently being prosecuted in Australian courts. It also aims to examine the quality of complainants’ child witness interviews, and whether interview quality is related to other case characteristics. Cross-examination and other trial variables such as outcome will be explored more fully in Studies 3 and 4, however, some basic variables are presented here to provide a comprehensive overview of these cases. Due to the exploratory nature of the study, no specific hypotheses were made.

Method

Trial and Child Witness Interview Transcripts

Relevant ethical approval to access CSA trial transcripts and child witness interviews was obtained. Trial transcripts for 96 CSA trials with 120 complainants and corresponding child witness interviews were provided by police organisations and the ODPP in three Australian states. The trials were selected
randomly by the ODPP and included 40 child (aged 6–12 years), 40 adolescent (aged 13–17 years) and 40 adult (aged over 18 years) complainants from each of the three jurisdictions. Of these, 35 complainants were excluded on the basis that they did not give evidence using a child witness interview. Except for two teenagers aged 16 and 17, who gave evidence via CCTV link, the rest of these excluded complainants were adults. The final sample comprised 85 complainants (40 child, 38 adolescent, and 7 adult) across 73 trials held between 2011 and 2015.

Due to ODPP resourcing issues and differing legal requirements across jurisdictions and courts, complete trial transcripts were not obtained for all trials in the final sample. For example, in one jurisdiction it is the policy not to order a daily transcript where the trial is shorter than five days. In contrast, longer trials would usually have a transcript ordered because there is more information to keep on top of. Circuit courts would not usually get a daily transcript, while complete transcripts are always ordered where there is a conviction appeal. Judicial discretion also allows the judge to order certain crucial parts of the trial to be transcribed to aid in his or her summing up of the trial. As a result, data for some variables was more complete than for others.

**Procedure**

**Trial transcripts.** Trial transcripts were read and information was extracted concerning the case, evidence, offence, and trial characteristics. All variables are explained in more detail below.

**Case Characteristics.** Gender of the complainant and the defendant were extracted and categorised as male or female. Although two complainants were
transgender, their gender was categorised using their biological sex given the sexual nature of the offences.

Complainant’s age at the first and last offence was calculated using date of birth and the dates of the first and last offences specified in the charge. Where the charge was particularised as “between dates”, the first date of the first charge, and the last date of the last charge, were used. Complainant’s age at first and last child witness interview or statement (in cases where the complainant underwent both an interview and statement) was used as recorded in the child witness interviews or statements. Complainant’s age at cross-examination was calculated from the complainant’s date of birth and the date of cross-examination. The date of the complainant’s cross-examination was chosen as a time-point instead of the date of trial as some complainants pre-recorded their entire testimony and were not present at trial. Thus, the time complainants were cross-examined best represents their last involvement in the prosecution process.

Age of defendant at trial was recorded if it was given at the trial. The defendant's relationship to the complainant at the time of the offence was categorised as one of the following: parent, stepparent or defacto of parent, guardian, uncle/aunt, grandparent (step included), other family (e.g., cousin), family friend, acquaintance, teacher/childcare worker/carer, religious worker, or stranger.

Time (in months) between the first and last child witness interview or statement, and time (in months) between the last offence and report were recorded. The latter variable was used as a measure of reporting delay, where “report” was defined as the date a report was made to police or a child protection agency. Where this date was unavailable, the date of the first child witness
interview was utilised. Finally, time (in months) between the complainant’s report and the cross-examination was used as a measure of delay in legal proceedings.

**Evidence characteristics.** Whether physical evidence was present was recorded, and the type of physical evidence noted as one of the following: DNA, sexually transmitted infection (STI), cuts and bruises, blood present in underwear, torn or affected hymen, vaginal/anal rash, irritation, indecent photos of the complainant/s, computer evidence (porn), and other.

The number of eyewitnesses was recorded for each complainant, but this measure did not include co-complainants as these were recorded in a separate variable (number of complainants in the trial). The number of non-complainant victims of the defendant who gave evidence at trial, and the presence of prosecution expert witnesses, were also noted. Expert witnesses were classed as one of doctor, clinical psychologist, research psychologist, forensic specialist, computer crimes expert, or multiple experts. An overall measure of the total number of prosecution witnesses (including expert witnesses, non-complainant victims, eyewitnesses, and any other prosecution witnesses such as family members) was also included.

Evidence variables for the defence included whether or not the accused gave evidence at the trial, and whether the defence called an expert witness. The type of expert defence witness was recorded in the same manner as for the prosecution. The total number of witnesses for the defence, including expert witnesses, was also included. Finally, whether the accused argued a defence or partial defence to the charges made by each complainant was recorded.

**Offence Characteristics.** The number of charges against the defendant for alleged offences against each complainant, as well as the total number of offences
in each trial, were recorded. The time between date of first and last offence, as specified in the charges, was noted in months.

Given the variety of offences available in the different jurisdictions included in this study, all alleged offences against each complainant were classed as being either a non-contact offence (e.g., grooming, exposure, masturbation), a contact offence (e.g., indecency, touching, kissing, forcing complainant to touch/kiss defendant), a non-penile penetrative offence (e.g., penetration with fingers or tongue), or a penile penetrative offence. Note that in the jurisdictions in this study, “sexual penetration” includes penetration to any extent of the external female genitalia, not just the vagina (Boxall, 2014).

Each of these offence categories was then split into whether consent was in issue or not. Although sexual offences against children generally do not require lack of consent to be proven, sometimes the prosecution chooses a charge that applies to all age groups (e.g., rape) and which requires lack of consent to be proven. One possible reason to prosecute “rape” as opposed to “sexual penetration of a child”, for example, might be the higher maximum penalties attached to the rape offence in some jurisdictions (For a review of contemporary Australian sexual offence and child sexual abuse legislation, see Boxall, 2014).

Finally, whether the offence was aggravated or attempted, and whether the offence was a repeated offence, were recorded.

**Trial Characteristics.** The number of child witness interviews undergone by the complainant was recorded. This included all interviews, with the exception of disclosure interviews,² irrespective of whether they were used as evidence-in-

---

² In some jurisdictions, in cases where the child has not orally disclosed the abuse to any person, an “initial disclosure interview” is conducted before a visually recorded forensic interview takes
chief. Whether these interviews were edited prior to trial, whether the complainant’s cross-examination was recorded prior to trial at a special hearing, and whether supplementary questions were asked by the prosecutor of the complainant at trial, were also noted.

The number of defence lawyers’ questions posed to the complainant during cross-examination were tallied as a measure of cross-examination length. Trial length in days was calculated from the date the jury was empanelled to the date of the verdict. Pre-trial deliberations between legal professionals were not included in this measure.

Finally, the verdict at trial and, where applicable, retrial, was recorded. As individual verdicts for every charge were not always available, verdicts were recorded for each complainant and for the trial as follows: not guilty of all charges, guilty of some charges/ not guilty of others, guilty of all charges, hung jury on all charges, hung jury on some charges/ guilty on others, hung jury on some charges/ not guilty on others, hung jury on some charges / guilty on others, or mistrial.

**Child witness interview transcripts.** Interviews were coded for four measures of interview quality: 1) interviewer questions, 2) interviewer behaviour, 3) interview length, and 4) investigative questions. These measures were recently published in an evaluation of an interviewer training program (Benson & Powell, 2015b), and reflect developmental, memory, and legal perspectives of what constitutes a “good” interview. Each measure is explained in more detail below. Adherence to the SIM interview protocol, another measure reported in this study, place. These initial interviews are conducted purely to elicit a disclosure, and are stopped immediately after any disclosure, or if the interviewer concludes no abuse occurred.
was not included here because interviewers had not been trained using this protocol.

**Interviewer questions.** Questions were coded as open-ended, specific, or leading. Categorising interview questions into these three categories is the standard measure of interview quality employed by most interviewing evaluations (Benson & Powell, 2015b; Cederborg et al., 2013; Cyr et al., 2012; Lamb et al., 2009; Sternberg, Lamb, Orbach, Esplin, & Mitchell, 2001). Open-ended questions were defined as any questions that encourage an elaborate response, but do not dictate what specific information is required (e.g., “What happened then?”). Specific questions were defined as questions that focus the child’s attention on previously disclosed details or aspects of the event and specify what precise information should be reported. These included cued recall questions (i.e., who, what, when, where, why, or how questions), “can-you” questions (e.g., “Can you remember what happened next?”), forced-choice questions (e.g., “Was it his left hand or his right hand?”), and “low-risk” yes/no questions (i.e., yes/no questions that ask the child for additional information about a particular category, but do not contain specific details, e.g., “Was anything said?”). Leading questions were defined as questions that presume or include a specific detail that was not previously mentioned by the child. “High-risk” yes/no questions were included in this category (i.e., yes/no questions that include at least one specific detail that the child had not previously mentioned, but do not suggest a desired answer, e.g., “Did he have a beard?”). The total number of questions asked in each interview was tallied, and the proportion of open-ended, specific, and leading questions calculated. For complainants who had multiple interviews, an overall measure of
the total number of questions, and the average proportion of open, specific and
leading questions across all interviews was also calculated.

**Interviewer behaviour.** Best practice interviewing behaviour was assessed
using a checklist of 14 interviewer behaviours created by Benson and Powell
(2015b). These evidence-based behaviours are recommended on the basis that
they encourage children to provide their best possible statements (American
Professional Society on the Abuse of Children, 2012; Lamb, 2016; Powell &
Snow, 2007). One point was awarded for each of the following behaviours
exhibited by interviewers: launching a narrative with an appropriate open-ended
invitation, correctly implementing ground rules, using a variety of minimal
encouragers (e.g., head nodding, “mmm”), sticking to one occurrence at a time,
exhausting each narrative, identifying the child’s use of generic language,
allowing the child to talk without interruption, asking a range (different types) of
open-ended questions, using simple language, avoiding the use of pronouns,
sticking to past tense, using developmentally appropriate language, avoiding
complex concepts, and avoiding “Can you” questions. Note that only one point
was awarded for each behaviour, irrespective of how many times the behaviour
was demonstrated in the interview. A score of the proportion of interviewer
behaviours exhibited out of those that could be exhibited in the interview was
calculated. For complainants who had multiple interviews, an overall measure of
the average proportion of interviewer behaviours exhibited across all interviews
was also calculated.

**Interview length.** The total length of interview in minutes, not including
breaks, was calculated. To account for more time being required when multiple
abusive incidents are discussed, a second measure of interview length per
occurrence was calculated. This involved dividing the total length of the interview by the number of incidents discussed. Two composite measures of overall interview length and overall interview length by occurrence were also calculated for complainants with multiple interviews.

Investigative questions. The evidential information interviewers sought to elicit was assessed using a list of evidential categories established in two prior studies in consultation with prosecutors and detectives (Benson & Powell, 2015b; Burrows et al., 2013). These categories were the identity of the offender, the approximate time of the offense, the location, the offense type, possible witnesses, and possible physical evidence. If the interviewer attempted to elicit a response about one of these categories during the interview, one point was awarded. For each attempt, it was also recorded whether the interviewers’ questioning method was appropriate (i.e., reflected developmentally appropriate language and concepts and refrained from questioning about irrelevant, minute details). The proportion of evidence categories sought, and the proportion of evidence categories sought appropriately, was calculated for each interview. For complainants with multiple interviews, the average proportion across their interviews was also calculated.

Reliability. All transcripts were coded by one person who had previously obtained interrater reliability with a second coder. Interrater reliability between these coders was analysed using the Kappa statistic, and Landis and Koch’s (1977) classifications of 0.0 to 0.2 as a slight agreement, 0.21 to 0.40 as a fair agreement, 0.41 to 0.60 as a moderate agreement, 0.61 to 0.80 as a substantial agreement, and 0.81 to 1.00 as an almost perfect agreement were used for interpretation. The coders had a high interrater agreement for interviewer
questions (κ = 0.96). With regard to interviewer behaviours, the coders had substantial agreement for the use of simple language and an appropriate open-ended invitation (Kappas = 0.80), and moderate agreement for the use of “Can you” question (κ = 0.42). Coders had complete agreement between for the 12 remaining interviewer behaviours (Kappas = 1.00). Finally, the coders also had substantial reliability for investigative questions (κ = 0.77).

**Results**

**Case Characteristics**

Case characteristics for the 85 CSA complainants from 73 trials are shown in Table 6.1. As can be seen, the majority of complainants were female and the majority of defendants were male. Nearly all defendants were known to the complainant, and slightly less than half (42.36%) were part of the complainant’s extended family (parent, stepparent, uncle/aunt, grandparent, or other family). The average reporting delay was around three and a half years, and it took another one and a half years on average for complainants to be cross-examined.

**Table 6.1**

*Case Characteristics for 85 Child Sexual Abuse Complainants*

<table>
<thead>
<tr>
<th>Categorical variables</th>
<th>n</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex of Complainant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>22</td>
<td>22.35</td>
</tr>
<tr>
<td>Female</td>
<td>63</td>
<td>77.65</td>
</tr>
<tr>
<td>Sex of Defendant a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>70</td>
<td>97.26</td>
</tr>
<tr>
<td>Female</td>
<td>3</td>
<td>2.74</td>
</tr>
</tbody>
</table>
Relationship to Complainant

<table>
<thead>
<tr>
<th>Relationship to Complainant</th>
<th>85</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent</td>
<td>8.24</td>
</tr>
<tr>
<td>Stepparent or de facto of parent</td>
<td>3.53</td>
</tr>
<tr>
<td>Guardian</td>
<td>8.24</td>
</tr>
<tr>
<td>Uncle/aunt</td>
<td>9.41</td>
</tr>
<tr>
<td>Grandparent (step included)</td>
<td>17.65</td>
</tr>
<tr>
<td>Other family</td>
<td>3.53</td>
</tr>
<tr>
<td>Family Friend</td>
<td>34.12</td>
</tr>
<tr>
<td>Acquaintance</td>
<td>11.76</td>
</tr>
<tr>
<td>Teacher/childcare worker/carer</td>
<td>1.18</td>
</tr>
<tr>
<td>Religious worker</td>
<td>2.35</td>
</tr>
<tr>
<td>Stranger</td>
<td>3.53</td>
</tr>
</tbody>
</table>

**Continuous variables**

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complainant age at first offence</td>
<td>82</td>
<td>9.99</td>
<td>3.31</td>
<td>2.95 – 16.72</td>
</tr>
<tr>
<td>Complainant age at last offence</td>
<td>82</td>
<td>11.24</td>
<td>3.04</td>
<td>5.74 – 16.80</td>
</tr>
<tr>
<td>Complainant age at first CWI</td>
<td>84</td>
<td>11.73</td>
<td>3.13</td>
<td>5.77 – 17.18</td>
</tr>
<tr>
<td>Complainant age at last CWI/statement</td>
<td>84</td>
<td>12.14</td>
<td>3.11</td>
<td>5.77 – 17.56</td>
</tr>
<tr>
<td>Complainant age at Cross</td>
<td>85</td>
<td>13.56</td>
<td>3.35</td>
<td>6.87 – 18.78</td>
</tr>
<tr>
<td>Defendant's age at trial a</td>
<td>44</td>
<td>45.48</td>
<td>13.34</td>
<td>22.00 – 75.00</td>
</tr>
<tr>
<td>Months between CWIs/statement</td>
<td>28</td>
<td>14.04</td>
<td>18.37</td>
<td>0.00 – 114.00</td>
</tr>
<tr>
<td>Months between last offence and report</td>
<td>77</td>
<td>41.87</td>
<td>216.85</td>
<td>0.00 – 1,365.00</td>
</tr>
<tr>
<td>Months between report and Cross</td>
<td>83</td>
<td>19.67</td>
<td>13.97</td>
<td>3.00 – 114.00</td>
</tr>
</tbody>
</table>

*Note: CWI = Child Witness Interview. Cross = Cross-examination.

a Variable calculated by trial (N = 73).
b Excludes complainants with only one CWI.

d Evidence Characteristics

Evidence characteristics for the 85 CSA complainants are shown in Table 6.2. The majority of cases did not have any corroborating physical evidence, but of those that did, the most common type was DNA evidence. The prosecution called an expert witness in around one quarter of cases, with the most common type of witness being a doctor. Eyewitnesses and non-complainant victims were
very rare, although trials had on average more than one complainant and about five prosecution witnesses in total.

The defence rarely called expert witnesses, and on average had less than one witness per case. The accused chose to give evidence in around half the cases, but only pleaded a defence or partial defence in a small minority of cases.

Table 6.2

<table>
<thead>
<tr>
<th>Categorical variables</th>
<th>n</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Evidence</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>DNA Present</td>
<td>7.06</td>
<td></td>
</tr>
<tr>
<td>STI</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Cuts and Bruises</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>Blood in underwear</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Torn or Affected Hymen</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>Vaginal/anal rash</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Irritation</td>
<td>2.35</td>
<td></td>
</tr>
<tr>
<td>Indecent Photos of the Complainant</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>Computer Evidence (Porn)</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>85.88</td>
<td></td>
</tr>
<tr>
<td>Prosecution Expert Witness</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>Doctor</td>
<td>9.52</td>
<td></td>
</tr>
<tr>
<td>Forensic specialist</td>
<td>5.95</td>
<td></td>
</tr>
<tr>
<td>Clinical psychologist</td>
<td>3.57</td>
<td></td>
</tr>
<tr>
<td>Research psychologist</td>
<td>1.19</td>
<td></td>
</tr>
<tr>
<td>Computer Crimes Expert</td>
<td>2.38</td>
<td></td>
</tr>
<tr>
<td>Multiple Experts</td>
<td>4.76</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>72.62</td>
<td></td>
</tr>
</tbody>
</table>
Defence Expert Witness * 72

None 97.22
Doctor 1.39
Clinical Psychologist 0.00
Research psychologist 1.39
Forensic Specialist 0.00
Computer Crimes Expert 0.00
Multiple Experts 0.00

Accused Testimony * 72

No 54.17
Yes 45.83

Defence or Partial defence 84

No 96.43
Yes 3.57

Number of Eyewitnesses 85

None 81
One 4

Continuous variables

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Complainants at Trial *</td>
<td>73</td>
<td>1.29</td>
<td>0.61</td>
<td>1.00 – 4.00</td>
</tr>
<tr>
<td>Number of Non-Complainant Victims *</td>
<td>73</td>
<td>0.07</td>
<td>0.38</td>
<td>0.00 – 3.00</td>
</tr>
<tr>
<td>Number of Prosecution Witnesses *</td>
<td>72</td>
<td>4.90</td>
<td>2.73</td>
<td>0.00 – 14.00</td>
</tr>
<tr>
<td>Number of Defence Witnesses *</td>
<td>72</td>
<td>0.81</td>
<td>1.33</td>
<td>0.00 – 5.00</td>
</tr>
</tbody>
</table>

Note: CWI = Child Witness Interview. Percentages for evidence present may not add up to 100% as cases can have multiple types of evidence present.

a Variable calculated by trial (N = 73).

Offence Characteristics

Offence characteristics for all 85 complainants are shown in Table 6.3.

Contact offences where consent was not in issue were the most common type of offences prosecuted, followed by non-penile penetrative offences where consent was not in issue. Non-contact offences and penile penetrative offences were the next most frequent, and were prosecuted in approximately a quarter of cases.
Aggravated and attempted offences were charged infrequently, and half of the charges in the sample involved a repeated offence. There were about three offences per complainant and five per trial, with an average of 14 months between offences.

Table 6.3
Offence Characteristics for 85 Child Sexual Abuse Complainants

<table>
<thead>
<tr>
<th>Categorical variables</th>
<th>Percent (%)</th>
<th>n</th>
<th>No</th>
<th>Yes</th>
<th>Attempt</th>
<th>Aggravation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-contact offence</td>
<td></td>
<td>85</td>
<td>75.29</td>
<td>24.71</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Contact offence, consent in issue</td>
<td></td>
<td>85</td>
<td>98.82</td>
<td>0.00</td>
<td>0.00</td>
<td>1.18</td>
</tr>
<tr>
<td>Contact offence, consent no issue</td>
<td></td>
<td>85</td>
<td>30.59</td>
<td>69.41</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Non-penile penetrative, consent in issue</td>
<td></td>
<td>85</td>
<td>95.29</td>
<td>2.35</td>
<td>1.18</td>
<td>1.18</td>
</tr>
<tr>
<td>Non-penile penetrative, consent no issue</td>
<td></td>
<td>85</td>
<td>68.24</td>
<td>30.59</td>
<td>0.00</td>
<td>1.18</td>
</tr>
<tr>
<td>Penile penetrative, consent in issue</td>
<td></td>
<td>85</td>
<td>95.29</td>
<td>4.71</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Penile penetrative, consent no issue</td>
<td></td>
<td>85</td>
<td>78.82</td>
<td>20.00</td>
<td>1.18</td>
<td>0.00</td>
</tr>
<tr>
<td>Repeated offence</td>
<td></td>
<td>84</td>
<td>50.00</td>
<td>50.00</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Continuous variables</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of offences per complainant</td>
<td>81</td>
<td>3.43</td>
<td>2.75</td>
<td>1 – 13</td>
</tr>
<tr>
<td>Number of offences per trial a</td>
<td>71</td>
<td>4.76</td>
<td>4.56</td>
<td>1 – 20</td>
</tr>
<tr>
<td>Months between offences b</td>
<td>62</td>
<td>14.42</td>
<td>25.13</td>
<td>0 – 119</td>
</tr>
</tbody>
</table>

Note: CWI = Child Witness Interview.

a Variable calculated by trial (N = 73).
bExcludes complainants with only one offence against them.

Trial Characteristics

Trial characteristics are shown in Table 6.4. Interviews were edited and supplementary questions asked in the majority of cases. Complainants had between one and four child witness interviews each. Cross-examination was pre-recorded for over half the complainants, during which complainants were asked almost 500 questions on average by defence lawyers. Trials were about a week
long, and just over half resulted in a not guilty verdict on all charges. Forty percent of trials resulted in a guilty verdict on at least one charge, while about one in 20 trials resulted in a mistrial. Of the four cases that were retried, two resulted in acquittal, and one in a verdict of guilty on some charges.

**Table 6.4**

*Trial Characteristics for 85 Child Sexual Abuse Complainants*

<table>
<thead>
<tr>
<th>Categorical variables</th>
<th>n</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was the interview edited?</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td>18.92</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
<td>81.08</td>
</tr>
<tr>
<td>Were supplementary questions asked?</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td>3.61</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
<td>96.39</td>
</tr>
<tr>
<td>Was the cross-examination pre-recorded?</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td>57.65</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
<td>42.35</td>
</tr>
<tr>
<td>Verdict (per complainant)</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Guilty of all charges</td>
<td></td>
<td>28.57</td>
</tr>
<tr>
<td>Guilty of some charges/ not guilty of others</td>
<td></td>
<td>11.69</td>
</tr>
<tr>
<td>Not guilty of all charges</td>
<td></td>
<td>50.65</td>
</tr>
<tr>
<td>Hung jury on all charges</td>
<td></td>
<td>3.90</td>
</tr>
<tr>
<td>Mistrial</td>
<td></td>
<td>5.19</td>
</tr>
<tr>
<td>Verdict (per trial)</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Guilty of all charges</td>
<td></td>
<td>26.76</td>
</tr>
<tr>
<td>Guilty of some charges/ not guilty of others</td>
<td></td>
<td>12.68</td>
</tr>
<tr>
<td>Not guilty of all charges</td>
<td></td>
<td>52.11</td>
</tr>
<tr>
<td>Hung jury on all charges</td>
<td></td>
<td>2.82</td>
</tr>
<tr>
<td>Mistrial</td>
<td></td>
<td>5.63</td>
</tr>
</tbody>
</table>
Retrial verdict (per complainant) 5
Guilty of all charges 0.00
Guilty of some charges/ not guilty of others 20.00
Not guilty of all charges 60.00
Mistrial 20.00

Retrial verdict (per trial) a 4
Guilty of all charges 0.00
Guilty of some charges/ not guilty of others 25.00
Not guilty of all charges 50.00
Mistrial 25.00

<table>
<thead>
<tr>
<th>Continuous variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Number of child witness interviews</td>
</tr>
<tr>
<td>Length of cross-examination (questions)</td>
</tr>
<tr>
<td>Length of trial (days)</td>
</tr>
</tbody>
</table>

a Variable calculated by trial (N = 73).

**Interview Quality**

Variables representing child witness interview quality are shown in Table 6.5. Complainants were asked over 250 questions on average, and the vast majority of these were specific. Only slightly more open than leading questions were asked, and approximately one third of desirable interviewer behaviours were exhibited. Interviews lasted about an hour on average, and slightly over half an hour was spent discussing each abusive occasion. However, a number of interview transcripts did not include interview times (start, break or finish times), so interview length could not be calculated for these. Finally, the majority of important investigative questions were asked (80%), but only half of these were asked appropriately.
Table 6.5

*Interview Quality variables*

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of questions</td>
<td>74</td>
<td>256.30</td>
<td>201.09</td>
<td>38.00 – 1003.00</td>
</tr>
<tr>
<td>Proportion of Open Questions</td>
<td>74</td>
<td>.13</td>
<td>.08</td>
<td>.03 – .38</td>
</tr>
<tr>
<td>Proportion of Specific Questions</td>
<td>74</td>
<td>.78</td>
<td>.10</td>
<td>.50 – .92</td>
</tr>
<tr>
<td>Proportion of Leading Questions</td>
<td>74</td>
<td>.09</td>
<td>.07</td>
<td>.00 – .42</td>
</tr>
<tr>
<td>Proportion of Int Behav exhibited</td>
<td>74</td>
<td>.39</td>
<td>.11</td>
<td>.17 – .70</td>
</tr>
<tr>
<td>Interview Length (mins)</td>
<td>60</td>
<td>57.63</td>
<td>35.70</td>
<td>19.00 – 200.00</td>
</tr>
<tr>
<td>Length per Occasion (mins)</td>
<td>60</td>
<td>36.56</td>
<td>16.48</td>
<td>13.00 – 82.00</td>
</tr>
<tr>
<td>Proportion of InQ asked</td>
<td>74</td>
<td>.80</td>
<td>.20</td>
<td>0.00 – 1.00</td>
</tr>
<tr>
<td>Proportion of InQ asked appropriatelly</td>
<td>74</td>
<td>.44</td>
<td>.29</td>
<td>0.00 – 1.00</td>
</tr>
</tbody>
</table>

*Note:* Int Behav = Interviewer Behaviours; InQ = Investigative Questions.

**Associations Between Case Factors and Interview Quality**

Analyses were conducted to examine associations between four case factors and interview quality. Four interview variables were selected to reflect interview quality: the proportion of open questions, the proportion of interviewer behaviours exhibited, the interview length per occasion, and the proportion of investigative questions asked appropriately. These variables were chosen because they best represented different aspects of interview quality, without being highly correlated.³

The four case factors examined were average age at child witness interviews (mean of the complainant’s ages at each of their interviews), type of case (worst offence), whether or not there were repeated offences, and whether or

³ The largest correlation was between open-ended questions and length per occasion, $r_s = .342, p = .008$. The size of this effect was small, $r^2_s = .12$. 
not there was corroborative evidence present. Corroborative evidence was
deemed to be present if there was either physical evidence, co-complainants, or
eyewitnesses present, or when previous sexual convictions had been mentioned.
The type of case variable was created by collapsing the offence variables (apart
from repeated offence) into three variables: penile penetrative, non-penile
penetrative, and non-penetrative (which included both contact and non-contact
offences). The case was then classified according to the worst type of offence that
was at issue for each complainant, with penile penetrative was classed as worst,
then non-penile penetrative, then non-penetrative.

As the assumptions for a MANOVA were not met, separate univariate
tests were conducted for each interview variable and Bonferroni corrections (α =
.0125) were applied to control for Type 1 error rates. For any variables that did
not met the assumption of normality, non-parametric analyses were performed. At
this statistical level, none of the tests were significant. The proportion of open-
ended questions was not significantly related to the average age at child witness
interviews, $r_s = .018$, $p = .88$; type of case $H(2) = 1.04$, $p = .60$; whether the
offence was a repeated offence ($Mdn = .11$) or not ($Mdn = .12$), $U = 639.00$, $p =
.77$; nor whether there was corroborative evidence present ($Mdn = .09$) or not
($Mdn = .13$), $U = 544.50$, $p = .14$.

The proportion of interviewer behaviours exhibited was not related to
average age at child witness interviews, $r = .098$, $p = .41$, the type of case,

---

4 Both the proportion of open questions and length per occasion variables were non-normal despite
having a large sample size > 30, with significant Kolmogorov-Smirnov tests at $p = .018$ and $p =
.006$; skewed distributions at Zskew = 5.17 and Zskew = 3.42; and the proportion of open
questions also being significantly kurtosed, Zkurtosis = 3.97. Further, not all dependent variables
were independent, with average age at CWI being significantly different according to the type of
case $F(2, 81) = 9.02$, $p < .001$, $η^2 = .18$. The assumption for independence of covariance matrices
was also violated, with Box’s test being significant at $p < .001$. 
$H(2) = 3.06, p = .22$, whether the offence was a repeated offence ($M = .38, SD = .10$) or not ($M = .39, SD = .12$), $t(71) = 0.55, p = .58$; nor whether there was corroborative evidence present ($M = .40, SD = .12$) or not ($M = .37, SD = .10$), $t(72) = 1.19, p = .24$.

The length of the interview per occasion was also not significantly related to average age at child witness interviews, $r = -.004, p = .98$; the type of case, $H(2) = 5.30, p = .07$; whether the offence was a repeated offence ($Mdn = 30.44$) or not ($Mdn = 37.00$), $U = 269.50, p = .15$; or whether there was corroborative evidence present ($Mdn = 32.50$) or not ($Mdn = 33.33$), $U = 424.00, p = .71$.

Finally, the proportion of investigative questions asked appropriately was not significantly related to average age at child witness interview, $r = .045, p = .71$; type of case, $F(2, 71) = 1.92, p = .15$; whether the offence was a repeated offence ($M = .39, SD = .26$) or not ($M = .49, SD = .32$), $t(71) = 1.56, p = .12$; nor whether corroborative evidence was present ($M = .37, SD = .26$) or not ($M = .51, SD = .31$), $t(72) = 2.12, p = .037$.

The Relationship Between Age at Cross-Examination and Case Variables

To examine whether the age of a complainant at trial (cross-examination) was related to the type of case that was being prosecuted, three further statistical tests were conducted. A Bonferroni correction ($\alpha = .017$) was applied to control for Type 1 error rates. An ANOVA revealed that the type of case (worst offence) was significantly related to age at cross-examination, $F(2, 82) = 11.08, p < .001, \eta^2 = .21$. Hochberg’s GT2 showed that older children were significantly more likely to be the victims of both penile ($M = 15.42$ years, $SD = 2.79$ years) and non-penile ($M = 14.65$ years, $SD = 3.28$ years) penetrative offences than younger children ($M = 12.00$ years, $SD = 2.95$ years), both $ps < .01$. 

Age of a complainant at cross-examination was not significantly related to whether the offence was a repeated offence ($M = 14.32$ years, $SD = 3.19$) or not ($M = 12.74$ years, $SD = 3.36$), $t(82) = 2.21$, $p = .03$. It was also not significantly related to whether there was corroborative evidence present ($M = 13.29$, $SD = 3.30$) or not ($M = 13.88$, $SD = 3.42$), $t(83) = 0.80$, $p = .43$.

**Discussion**

This study provided a detailed description of 73 CSA trials held in three Australian jurisdictions between 2011 and 2015. The most important findings for each of the four groups of variables examined – case, evidence, offence, and trial characteristics – are reviewed below, followed by a discussion of the results regarding interview quality.

First, in terms of case characteristics, the majority of complainants were female, ranging in age from six to 18 years, and the majority of defendants were male. Nearly all defendants were known to the complainant, and slightly less than half were part of the complainant’s extended family. Although no information is available regarding the preceding stage in the prosecution process, and the characteristics of the cases rejected for prosecution, this last finding most likely does not support prior research suggesting that cases are less likely to be prosecuted if the offender is known or related to the complainant (Bunting, 2008; Fitzgerald, 2006; Stroud et al., 2000).

Further, when evidence characteristics were examined, results showed that most cases in the current sample did not have any corroborating physical evidence, and of those that did, the most common type was DNA evidence. This finding supports the perception by some Australian prosecutors that “oath against oath” cases are now more likely to be prosecuted (Success Works, 2011).
Whether or not corroborative evidence was present was not significantly related to the age of the complainant at cross-examination, suggesting that even young children’s “oath against oath” cases are being prosecuted, and that corroborative evidence is not more likely in younger compared to older children’s cases when it comes to the likelihood of their cases being prosecuted.

For offence characteristics, offence type and seriousness have previously been linked to an increased likelihood of prosecution (Cross et al., 1994; Fitzgerald, 2006; Parkinson et al., 2002), but in the current sample, contact offences (followed by non-penile penetrative offences) were actually the most common type of offence prosecuted. Although not directly inconsistent with past research (as it is not known what type of cases were rejected for prosecution), the current finding indicates that sexual offences against children need not be of the most serious kind to be accepted for prosecution. Given that the type of case, in terms of the worst offence prosecuted, was related to the age of the complainant at cross-examination, the finding that a significant number of contact offences are being prosecuted is particularly important in terms of access to justice for younger complainants.

Results for trial characteristics showed that supplementary questions were asked in nearly all cases, indicating that calls for legislative reforms to allow for supplementary questioning in CSA trials is not warranted (Burrows & Powell, 2014c). However, the current findings support anecdotal evidence suggesting that complainants often undergo long cross-examinations despite reforms that aimed to stop inappropriate and repetitive questioning (Shead, 2014), with complainants being asked almost 500 questions on average by defence lawyers. Finally, just over half of the cases in the current sample resulted in a not guilty verdict on all
charges, which is consistent with past studies examining conviction rates (Australian Law Reform Commission, 2010; Cashmore & Trimboli, 2005; Fitzgerald, 2006; Wundersitz, 2003).

One unique contribution of the current study is its examination of the quality of child witness interviews of complainants whose cases had reached trial. One of the best measures of interview quality is the proportion of open-ended question used (Powell et al., 2005). Although there are no definitive guidelines on the percentage of questions that should be open-ended, some child interviewing experts have suggested that around 80% of questions should be open (Luther et al., 2015), while others have classified interviewers who used fewer than 40% open-questions as “low performers” (Powell, Guadagno, et al., 2016). Studies conducted over the last five years in countries with adversarial legal systems have found rates of open-ended questions in field-interviews for experienced interviewers (that is, interviewers who have received some form of training) ranging from as low as 4% to as high as 57% (Benson & Powell, 2015b; Gagnon & Cyr, 2017; Luther et al., 2015; Price & Roberts, 2011; Wolfman et al., 2016). In the current study, 13% of questions in interviews were open-ended, which suggests that the interviews in this sample were poor quality on average.

Other measures of interview quality backed up this finding. A high percentage of questions in interviews were specific (78%), and nearly as many leading questions were asked (9%) as open-ended questions. This finding contrasts best practice interviewing guidelines, which recommend that specific questions should be used sparingly (Brown et al., 2013; Powell & Thomson, 2001), and leading questions completely avoided (Lamb, 2016; Powell, 2002). Additionally, interviews in the current sample were relatively long, lasting for an
average of one hour (or 36 minutes per occasion discussed). In comparison, one recent study found that interviewers who had undergone a comprehensive training program significantly shortened their interviews, taking only approximately 40 minutes on average (or 23 minutes per occasion) without any loss in important evidential information sought (Benson & Powell, 2015b). Again, the length of interviews in the current sample is inconsistent with interviewing guidelines and prosecutors’ recommendations, both of which suggest that interviews should be kept short (American Professional Society on the Abuse of Children, 2012; Burrows & Powell, 2014c). Finally, although investigative questions were asked in the majority of interviews, only half of these were asked appropriately (cf. 92% asked appropriately in post-training interviews; Benson & Powell, 2015), and preferred interview behaviours were only exhibited in slightly over one third of cases (cf. 81% exhibited in post-training interviews; Benson & Powell, 2015).

The finding that these field interviews were poor on average is consistent with past research (Lamb, 2016; Luther et al., 2015; Powell, Cavezza, Hughes-Scholes, & Stoove, 2010; Powell, Murfett, et al., 2010; Wolfman, Brown, & Jose, 2016), and is likely due to inadequate training programs (Benson & Powell, 2015b).

An examination of the association between certain interview variables (the proportion of open questions, the proportion of interviewer behaviours exhibited, the interview length per occasion, and the proportion of investigative questions asked) and case variables (average age at child witness interviews, worst offence prosecuted, whether or not there were repeated offences, and whether or not there was corroborative evidence present) showed that none of these variables were significantly related. There are two possible explanations for the lack of association.
First, it is possible that prosecutors do not take interview quality (as measured in this study) into account when deciding whether or not to prosecute a case. If it were taken into account, a relationship between the presence of corroborative evidence and interview quality, for example, might be expected, so that cases with no corroborative evidence were associated with better interviews. A similar relationship might be expected for younger children and cases with less severe charges. Each of these factors (no corroborative evidence, young complainants, less severe offense) have been found to be associated with a lessened likelihood of prosecution (Bunting, 2008; Christensen et al., 2014; Davis et al., 1999; Fitzgerald, 2006; Leach et al., 2016), meaning that a good child witness statement (in the form of the interview) logically might be expected when these factors are present to improve the strength of the case. This supposition is supported by the finding in one past study that the implementation of a best practice interview protocol results in more guilty verdicts (Pipe et al., 2013).

It is also possible, however, that the lack of a significant association between interview and case variables in the current study is due to methodological limitations. One key limitation is the fact that most interviews in the current sample were of poor quality with a limited range in quality; open-ended questions, for example, were non-normally distributed, with the majority of interviews having a low proportion of open-ended questions. Thus, there may not have been adequate variability in the sample to detect any relationships.

In conclusion, the current study presented a comprehensive description of the sample of CSA cases that form the basis of Studies 2-4 in this thesis, and provided statistics that may be used to inform policy development and criminal justice responses in the area of child sexual abuse. Specifically, the study showed
that the quality of complainants’ child witness interviews is generally poor, and does not appear to be a factor taken into account when deciding whether to prosecute a case. However, due to methodological limitations, further research is needed to support this finding. In particular, trial transcripts should be studied and discussions about child witness interviews examined to determine whether, and how, the quality of interviews is discussed by legal professionals during the conduct of a trial.
Video-recorded child witness interviews were introduced in a number of jurisdictions including Australia and the United Kingdom as part of a host of reforms designed to improve vulnerable complainants’ experiences at trial and the quality of their evidence. Rather than having to give their evidence live in court in front of the alleged offender, complainants in CSA cases can now use their pre-recorded interviews as their evidence-in-chief. Research has shown this measure has numerous benefits, including reducing the amount of stress and trauma that complainants experience in court (Corns, 2001; Eastwood & Patton, 2002; Hamlyn et al., 2004; Wilson & Davies, 1999), preserving evidence by recording the complainants’ accounts closer to the time of the incidents (Corns, 2001; E. Davies & Hanna, 2013; Hanna, Davies, Henderson, Crothers, & Rotherham, 2010), and reducing the need for complainants to tell their story numerous times (Ghetti et al., 2002; La Rooy et al., 2009). Pre-recorded evidence also allows prosecutors and defence to better prepare their cases (Burrows & Powell, 2014d; Cashmore & Trimboli, 2005), and enables the court to edit out inadmissible evidence and thus reduce the chance of a mistrial (Australasian Institute of Judicial Administration, 2012; Corns, 2001; E. Davies & Hanna, 2013; Hanna et al., 2010).

Despite these many advantages to pre-recorded child witness interviews, issues remain with their use in court. First and foremost, the quality of the interview (and evidence) is dependent on the skills of the interviewer (E. Davies & Hanna, 2013; Powell & Thomson, 2001), and studies have shown that the
quality of interviews is often poor. For example, legal professionals (including both prosecutors and defence lawyers) have expressed concerns that interviews are frequently too long and include irrelevant and excessive detail (Burrows & Powell, 2013; Burrows et al., 2013; Cashmore & Trimboli, 2005; McConachy, 2002). Such interviews are fatiguing for a jury to watch, and difficult for a child to be cross-examined on (Burrows & Powell, 2014b; E. Davies & Hanna, 2013).

Interviewers also typically employ poor questioning techniques, asking low proportions of open-ended questions (e.g., “What happened when…?”) and high proportions of specific questions (e.g., “What was he wearing?”, “when did that happen?”), which is contrary to best-practice guidelines (G. Davies et al., 2000; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007; Powell & Guadagno, 2008). Memory research has demonstrated that open-ended questions elicit more accurate information than closed questions (Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007; Powell et al., 2007), and a high proportion of open-ended questions is one of the key markers of a good interview (Benson & Powell, 2015b; Powell, 2005). However, concerns about the question types employed in interviews have chiefly emanated from psychological researchers; indeed, one interviewing expert has recently queried why legal professionals have not criticised the generally poor quality of investigative interviews, and demanded a change in line with best-practice interviewing principles (Lamb, 2016). It is possible that, as Lamb concluded, these professionals cannot recognise a “good quality” interview, and do not know how much information children can provide when questioned appropriately. However, it is also possible that what is considered problematic from a psychological perspective is not the same as what is problematic from a legal perspective, and the issues that are raised in the
interviewing literature are not those of most concern to legal professionals. One way to examine which explanation has the most support is to examine the discussions held between legal professionals about interviews in court.

One final disadvantage of the use of pre-recorded interviews in court is the high rate of technological problems, such as equipment failure (Plotnikoff & Woolfson, 2009) and poor sound and visual quality of interviews (Burton et al., 2006; Cashmore & Trimboli, 2005; Criminal Justice Joint Inspection, 2009). Given that jurors consider demeanour and expression in assessing the veracity of complainants’ accounts (Fisher, 2015; Porter & ten Brinke, 2008), poor quality recordings are likely to affect their judgments (Burrows & Powell, 2014d).

To date, most of the problems that have been identified with child witness interviews have come from examinations of interviews themselves or from discussions with prosecutors and other stakeholders (Burrows & Powell, 2014d; Hill & Davies, 2013; Powell et al., 2011). Some observational studies have also noted technological issues with interviews during court proceedings (Burton et al., 2006; Cashmore & Trimboli, 2005). No research has examined whether these problems have a direct influence on the way in which CSA trials occur, nor examined the discussions between legal professionals at trial to determine how these professionals view the interviews they are using in their cases. The aim of the current study was therefore to examine whether and how legal professionals discuss child witness interviews during CSA trials, with a particular focus on the interviews’ usefulness as evidence-in-chief. It is important to identify whether such discussions occur as they highlight the long-reaching effect that child witness interviews can have. Given that there are a number of opportunities for issues to be addressed prior to trial (e.g., before charges are laid, when deciding to
prosecute, and during pre-trial hearings), any issues that are discussed in court should represent those that are the most important at trial.

**Method**

**Trial Transcripts and Child Witness Interviews**

This study utilised the same sample of 85 complainants described in Study 1. There were 19 (22.4%) male and 66 (77.7%) female complainants, ranging in age from 6.87 to 18.78 years at trial or pre-recording ($M = 13.56$, $SD = 3.34$). Complainants had an average of 1.4 child witness interviews (range 1-4), which were comparable in quality to interviews conducted by untrained interviewers (Lamb et al., 2008).

**Data Management and Analysis**

Trial transcripts were read and any discussions between legal professionals concerning the child witness interview recorded. The data were then subjected to open coding (Strauss & Corbin, 1990), which involved a line-by-line analysis of the discussions (i.e., reduction) and identification of topics or issues raised therein. Discussions concerning similar topics were grouped together. The data were then re-examined for statements that supported the identified topics and issues. Thus, the identification of core topics helped to reduce the large volume of data into meaningful and discrete units of analysis (Miles & Huberman, 1984).

**Results**

Prosecutors, defence counsel and judges discussed the child witness interviews for 81 of the 85 complainants (95.3%). These discussions were held in three different contexts. First, interviews were discussed in terms of their usefulness as evidence-in-chief. Second, interviews were mentioned during legal discussions about judicial directions and the admissibility of evidence. Third,
interviews were referred to in the context of planning and organising the trial. As the first context was the primary focus of the study, the second and third contexts are only briefly described below.

**Usefulness of the Complainant’s Child Witness Interview as Evidence-In-Chief**

Discussions between legal professionals regarding the usefulness of the child witness interview as evidence-in-chief were held in 51 of the 85 complainants’ cases (60.0%), and focused on three broad topics: interview procedure, technological issues, and the structure of the interview. All three topics were discussed for one complainant, two topics were discussed for 17 complainants, and one topic was discussed for 33 complainants. The following section examines these topics in more detail, in order of frequency of their occurrence.

**Interview procedure.** Interview procedure was the first broad theme that arose from trial discussions about the usefulness of the interview as evidence-in-chief. This theme arose for 29 of the 85 complainants (34.1%). Subthemes within this topic included interviewer questioning, and the behaviour and actions of the interviewer and complainant.

Discussions about interviewer questioning included mistakes that were made by the interviewers, such as putting the wrong version of events to children during the interview, and mixing up details across incidents of abuse. Such mistakes often caused confusion for children during cross-examination, and cross-examinations were usually stopped to clarify children’s responses or to highlight interviewers’ errors. In one case, for example, the interviewer moved so rapidly between incidents of abuse that the child ended up stating that the accused did not
tell her to “suck his rude part”, even though this act had in fact constituted a charge. This resulted in the defence arguing that the charge should be dropped. In another case, the interviewer put an incorrect version of events to the child, which the child failed to correct, resulting in the defence attacking his reliability during cross-examination. After the cross-examination was stopped, the judge suggested that interviewer was more at fault than the complainant, but the defence did not agree.

**Judge:** When the police have put something to him that’s wrong, he hasn’t picked them up, so you say that reflects adversely on his reliability of recollection.

**Defence:** When it’s put to him twice in quite a short time frame…then that is of some concern.

**Judge:** Well, it’s a criticism of the child witness interviewer, more than a criticism of the witness, it seems.

**Defence:** But it is also a criticism of the witness. The prosecution is putting up what is said in that interview and what’s accepted by this witness in that interview as being the evidence for the prosecution, and in those circumstances if he does say something or he’s unwilling to correct the police officer, then that is of concern.

(Complainant 20)

Another way in which interviewer questioning was mentioned during trials was whether certain elements of the offence (e.g., penetration) had been established, and whether sufficient particularisation or detail of an offence had been obtained. In at least one case, it was suggested that interviewers had used leading questions to obtain this information, which could not be edited out of the evidence.

**Prosecutor:** There are portions of the pre-recording where the interviewing officer almost seems to be trying to do a repair job, for
want of a better expression, which includes leading statements. But
the converse of that is that the defence can address the jury in that
regard and tell them what a leading statement is so that there’s no
real point in trying to edit that out [the officer’s leading statement]
because it’s in. (Complainant 102)

Judges also made comments about interviewers’ behaviour and the effect
of this behaviour on the child. For example, one judge mentioned that an
interviewer interrupted a child’s train of thought and moved to an irrelevant topic
just after the child disclosed an act of abuse.

**Judge:** If I can just direct your attention again to page 4 of that
second interview; she says ‘…he did with his rude part to me. He
told me to suck it first. I told him no. He forced me to.’ Of course,
this happens with these interviewers who want to continually
interrupt the train of thought of a child, and so we have this question,
‘Q. Tell us what your mother said. A. Well, [s]he said he mightn't go
to jail because we don’t have enough evidence to prove it.’
(Complainant 42)

As well as discussing the behaviour of interviewers during interviews,
legal professionals discussed the behaviour of complainants. Children’s actions
and demeanour during an interview were often used to judge the veracity of their
claims. In fact, when addressing the jury at the end of the trial judges usually gave
the instruction to regard the child’s demeanour and behaviour carefully to help the
jury come to its decision. In a number of cases, however, children’s behaviour
was used by the defence as a ground for arguing that a reliability warning should
be given. Although warnings about children being unreliable witnesses as a class
have been outlawed, reliability warnings may still be given about particular child
witnesses if deemed necessary on the facts of the case. Some of the circumstances
in which the defence have argued for a reliability warning in this sample of cases
included the manner in which children disclosed during the interview (giggling and laughing), and children’s behaviour during an interview break (playing with the interview equipment in the absence of the interviewer).

**Defence:** …the grounds [for the reliability warning] that I rely on are this…in the interview, in particular, [the complainant] smiles, she laughs, she giggles at different points, which might be seen as inconsistent with the substance of what she's disclosing, and also, ultimately, that the disclosure—or the event—is said to have happened when she's asleep, or sleepy or just been woken…

**Judge:** Yes, well I'm not persuaded by…the argument about her demeanour… But in relation to the matters of inconsistency, I think that that's probably a matter that would go to highlight that the jury need to be cautious, and I'm happy to give them a warning in relation to exercising caution when viewing the interview, in the context of reliability and credibility. (Complainant 55)

**Judge:** …when the police officer went off to consult and came back to find, as we saw, a mischievous child playing with the equipment…You've got a child here who is technically savvy, on any view, and has no compunction about playing with equipment.

**Prosecutor:** One ought not to attribute to the child that aspect in relation to his savvy-ness; the fact that he tinkers with a unit which is sitting beside him and which he's seen the police officer playing with.

**Judge:** No, it's not just that. That he would do that in a police station and also at his own home…

**Prosecutor:** That, Your Honour, doesn't necessarily lead to a fair and reasonable inference that he's the sort of person who will invent a story of this nature. (Complainant 58)

**Technological and transcription issues.** Technological issues were discussed for 21 of the 85 complainants (24.7%). Consequences of technological
issues included delays in court proceedings and difficulties in understanding children’s testimony. Although these difficulties in understanding children’s testimony could often be remedied by providing transcripts of the interviews to the judge and jurors, these transcripts often contained errors, which further exacerbated the problem.

Judge: There are a number of errors in the interview transcript that I have noted and I probably haven’t got them all, but at p 11, question 121, ‘So tell me more about Uncle; what do you guys do together?’ rather than ‘what do guys do together?’ The word ‘you’ to my hearing was missing [from the interview] and that’s a significant difference. (Complainant 84)

The structure of the child witness interview. The structure of the child witness interview was another broad theme that arose from trial discussions about the usefulness of the interview as evidence-in-chief. It was discussed for 20 of the 85 complainants (23.5%). Legal professionals were concerned about the lack of clarity in the interviews, which at times resulted in confusion for complainants and legal professionals during both cross-examination and discussions held between legal professionals in the absence of the jury (voir dires). Interviews were described as “very poorly structured” and “jumping from place to place”. They were also criticised for not following any chronological order, not clearly relating to the charges on the indictment, and being overly long.

Trial transcripts showed that in some cases, the lack of structure in children’s interviews led to objections and arguments over whether the necessary evidence had been gathered from the interview. At other times, poorly structured interviews provided an opportunity for the defence to confuse children during cross-examinations. The prosecutor in the following case illustrated the problem:
Prosecutor: Your Honour, the toing and froing in the questioning of the complainant in the particular interview is confusing in one sense, but, more importantly, to cross-examine based on a series of questions which in themselves are confusing—and to refer to those questions—adds unnecessary complication. And what I'd be submitting is, that for the purpose of cross-examination it's fundamentally important that the complainant understand an event and be able to give evidence about an event rather than crisscrossing the transcript by way of questions put by a police officer seeking to ascertain facts. It's fundamentally important, Your Honour, otherwise the witness becomes confused and it's patently obvious that that's what's happening. (Complainant 58)

The trial transcripts also revealed that long and convoluted interviews created difficulties in some cases because children could not remember all of their evidence; thus, they could not be effectively cross-examined. In the following case, for example, the defence lawyer attempted to cross-examine the complainant on whether or not she put her clothes on a speaker, but she could not remember this aspect of her interview.

Judge: I think the other problem is when she can’t remember now - it’s a long interview. To say, ‘Do you generally remember the concept of participating in an interview?’ ‘Yes, I do, but the specifics of it, no, I don’t’ - and to be quite honest, even for an adult, if you said, ‘Is there a reference in there to putting clothes on a speaker?’ we’d all have to probably turn all of the pages to satisfy ourselves to make such a concession. I don’t think she can remember what she did or didn’t say, and that’s the problem…. (Complainant 68)

Long interviews were also seen as problematic because they resulted in child and juror fatigue. For example, in the following case, a long interview led to
one judge halting the complainant’s evidence-in-chief because a juror was falling asleep.

*Prosecutor:* [in opening address] You heard from his Honour that [complainant] participated in an interview where she gave her account of what happened. Now, that interview is rather lengthy and we’ll spend most of today, if not all of today, watching that interview. It goes for about four and a half, five hours. …

*Judge:* [To the jury during complainant’s evidence] Ladies and gentlemen, as I said, this is more difficult than watching someone in the witness box so I think what we’ll do is that we’ll take an early lunch and we’ll resume at two o'clock [the jury retired]. [To Counsel] Just by way of explanation, one of the ladies was nodding off. (Complainant 25)

Finally, legal professionals discussed the length of interviews in the context of trial organisation. These discussions arose when deciding when interviews should be played, and whether the complainant had had the opportunity to refresh his or her memory. Apart from interview length, factors taken into consideration when deciding when to play the interviews included the number of complainant interviews, when complainants would be cross-examined, and whether the jury would need a break while watching the evidence. When discussing the need to refresh complainants’ memories of their interviews, legal professionals differed over what they deemed an acceptable time frame between the complainants watching their interviews and having to give evidence in court. In some cases, complainants watched their interview on the day of the trial; in others, they watched their interview a week or more earlier. There was no uniform view as to what time frame was ideal; some prosecutors expressed concern that
too much time had passed since the complainant had watched his or her interview, while others deemed similar time frames (e.g., a week) completely acceptable.

**Legal Issues**

Child witness interviews were frequently discussed in the context of legal issues, such as the admissibility of interview topics and the necessity of judicial directions. In terms of directions, the majority of discussions concerned directions about evidence of prior offences or lesser misconduct of the accused (“uncharged acts evidence”), inconsistent statements, and reliability warnings.

The admissibility of topics raised in child witness interviews was usually discussed in the context of an application by the defence to edit out questions or sections of the interview. In most cases, edits were agreed between prosecution and defence, and judges were simply informed of these. The majority of agreed edits were made on the basis that they lacked relevance, and judges usually informed the jury of this fact. However, sometimes prosecutors and defence counsel did not reach an agreement and judges were required to make a ruling on what should be edited from the interview. Discussions of which parts of the interview to be edited included evidence of the accused’s motive, tendency to act or think in a particular manner, relationship to the complainant, and competency testing conducted by interviewers.

In a number of cases, the defence made an application to cross-examine the complainant on older or other child witness interviews, and a ruling had to be made as to whether these were admissible. The grounds on which defence argued that such interviews were relevant included the need to demonstrate the complainant’s sexual abuse history (which might give rise to a reasonable doubt
about the offender’s identity), and to demonstrate that the complainant had a
history of making complaints.

**Defence:** Your Honour, the defence makes an application in relation
to two prior complaints and the use of evidence from those when
cross-examining the complainant in the matter…The defence
position is that this particular young lady has a fascination about men
and that there are increasing claims once she gets somebody’s
attention. What evidence do we say supports that? Well, in relation
to the February complaint, we have the fact that there’s a first
interview about an occasion and then she comes back for a second
bite of the cherry with a second interview. Now, in that second
interview she does exactly the same as she does on this occasion.
Right at the end of that interview, she then starts saying, ‘It wasn’t
just the individual times that I’ve talked to you about. It was lots of
times. We started having sex regularly several times a week’.
There’s this inflation of what she says on each occasion.
(Complainant 15)

**Trial Planning**

The final context in which child witness interviews were discussed in
court was trial planning. Much of the discussion here centred on jury issues.
Juries were provided with transcripts of the complainant’s interview in nearly
every case; however, some judges preferred to provide these before, and others
after, the recording of the interview had been played in court. A more contentious
issue was whether or not juries could watch the complainant’s interview a second
time, and, if so, whether they could take the recording into the jury room.
Whether or not this was allowed appeared to hinge on the circumstances of the
case, as well as the inclinations of the judge in the trial.

Discussions about child witness interviews in the context of trial planning
revealed that some judges lacked knowledge or experience with the process of
using interviews as evidence-in-chief. Some also lacked knowledge about the
procedure for vulnerable child witnesses in CSA cases more generally.

**Judge:** So what happens now? When you had these special hearings
the judge used to speak to young complainants about horses and
wigs and silly things like that; do we still do that? (Complainant 108)

**Judge:** Can I say at the outset that it has been a long time since
we've done a trial with a [child witness interview] and a pre-trial—a
special hearing—and the like, so I'll just need to be guided carefully
through the various procedural requirements, Mr [Prosecutor], if you
don't mind, and Mr [Defence].

**Defence:** Mr [Prosecutor] will be on top of it, Your Honour.
(Complainant 87)

The final way in which child witness interviews were included in
courtroom discussions about trial planning concerned a number of cases in which
child witness interviews had been lost. The loss of these interviews caused delays
in trials and concerns about the security of this highly sensitive material. In at
least two trials there were issues with the use of a postal service to transport
edited interviews, with one judge expressing concern about the practice.

**Judge:** Anyway, so the edited interview is coming back by…

**Prosecutor:** Express Post.

**Judge:** Express Post. Oh goodness, you're making me nervous now.
**Prosecutor:** I know, but in [assistant’s] long experience, it's never
failed, he tells me. It can even be tracked by the internet.

**Defence:** Almost never fails.

**Judge:** …I'm a bit interested to hear about interviews going through
Express Post... I wonder whether that's what should be happening. I
know it's practical but they are what they are. They're pretty
significant things to be entrusting to post. (Complainant 110)
Discussion

The aim of the current study was to examine whether and how legal professionals discuss child witness interviews during CSA trials, and to determine whether problems highlighted in interview and observational studies are echoed by these professionals at trial. Results showed that judges, prosecutors and defence lawyers discussed child witness interviews in three different contexts: usefulness as evidence-in-chief, during legal discussions, and while planning the trial.

Discussions about the usefulness of the interview as evidence-in-chief were explicitly raised in the majority of cases and concerned the structure of the interview, interview procedure, and technological issues. The issues raised suggested that child witness interviews had an impact on complainant memory, credibility, and reliability. They also caused delays to proceedings, prolonged trials and created confusion for complainants and legal professionals alike. These disadvantages of using pre-recorded interviews as children’s evidence in chief are consistent with past research (Benson & Powell, 2015a; Burrows & Powell, 2014b, 2014d). However, results of the current study offer further insight into the downstream effects of child witness interviews at trial.

For example, in terms of the first broad theme – interview procedure – research has shown that prosecutors are concerned with the reduced sense of formality in children’s interviews (Burrows & Powell, 2014d). Prosecutors noted that many children in interviews were dressed in clothes that would be inappropriate in court and displayed oppositional behaviour which prosecutors feared would negatively impact the jury’s assessment of the child’s credibility. Although the current study could not measure jury assessments of credibility,
results showed that reduced formality in the form of giggling or playing with recording equipment could in fact lead to the defence arguing for a reliability warning to be given regarding that child’s evidence. Prosecutors in past studies were also concerned about interviewers interrupting a child’s disclosure of abuse by asking irrelevant questions about contextual details (Burrows & Powell, 2014b, 2014d), concerns which were again echoed by legal professionals in the present study. It is essential for interviewers to keep in mind that anything that is recorded during the interview - even during breaks in the interview - may be used in court by the defence to undermine the complainant’s account.

The second broad theme that emerged about using children’s child witness interviews as evidence-in-chief was technological problems that occurred when the interviews were played in court. Technological problems are not new, nor are they restricted to Australia (Burton et al., 2006; Cashmore & Trimboli, 2006; Criminal Justice Joint Inspection, 2009; Eastwood & Patton, 2002; Plotnikoff & Woolfson, 2009; Powell et al., 2016). However, the fact that issues with technology emerged as a theme in the current sample of cases suggests that problems documented in older studies, such as poor sound and image quality of video recordings, have not yet been overcome. Given that poor audio and visual quality of pre-recorded interviews may reduce the clarity of the evidence presented (Burrows & Powell, 2014d), rectifying these issues should be a priority.

The third broad theme – interview structure – was also consistent with past research. For example, prosecutors have consistently stated that child witness interviews are too long (Burrows & Powell, 2013, 2014b, 2014c; Burrows et al., 2013). The results of the current study support their views as the length of child witness interviews emerged as one of the contexts in which child witness
interviews were discussed in court. Lengthy interviews can be problematic because they may deplete children’s cognitive resources and provide more opportunities for errors and inconsistencies between statements (Benson & Powell, 2015a; Burrows & Powell, 2014c, 2014d). Lengthy interviews may also fatigue jurors when played in court, as documented in the current study through a judge’s discussion of a juror falling asleep during a trial.

The results also showed that legal professionals discussed the lack of clarity in child witness interviews, particularly because they did not follow any chronological order. This finding adds weight to prosecutors’ concerns that, to be useful, child witness interviews should be concise, relevant, and clear, with the evidence elicited in a narrative format (Burrows & Powell, 2014b). Such narrative accounts should be more persuasive and credible to a jury because they allow children to tell their story in their own words. The results of the current study also demonstrated the confusion that can occur when children are cross-examined about what they did or did not say in the child witness interview. When the original interview was confusing, complainant’s confusion in court during the cross-examination was exacerbated, which may have affected their credibility in the eyes of the jury.

The other two main contexts in which child witness interviews were discussed centred on legal issues and trial planning. The legal issues included discussions about judicial directions and admissibility of certain types of evidence. Discussions were also held about the admissibility of older child witness interviews. Trial planning included discussions about jurors, such as when they should receive the transcript of the children’s interviews and the number of times they could watch the video-recorded interviews. Although these
contexts were not the focus of the current study, they are important to examine in their own right and should be a topic for future research.

Taken together, the results of the current study illustrate some of the downstream effects of child witness interviews and, in particular, the issues that can arise when playing these interviews as children’s evidence-in-chief in court. The results reinforce past concerns about the quality of the interviews and technology issues that have emerged through examinations of child witness interviews and through focus groups with prosecutors (Burrows & Powell, 2014d; Hill & Davies, 2013; Powell et al., 2011). However, results also demonstrate that the issues that are raised by legal professionals do not directly align with those mainly raised by interviewing professionals. That is, there were no discussions about question types or other aspects of best-practice interviewing (e.g., establishing rapport, ground rules) which form part of the evidence-based recommendations for interviewers (Lamb, 2016). This suggests that the aspects of interviews that are considered “problematic” by interviewing experts are not necessarily those that are considered most problematic at trial.

Nevertheless, there are at least three strategies that may assist with improving the quality of the interviews and technology, and improve their utility as complainants’ evidence-in-chief. First, the quality of the interviews may be improved through interviewers’ use of non-leading open-ended questions to encourage children to give coherent narrative accounts of the alleged abuse (Malloy, Johnson, & Goodman, 2013; Newlin et al., 2015; Powell & Cauchi, 2013). Although legal professionals did not specifically mention question types in their discussions, the increased utilisation of non-leading open-ended questions may increase the coherence of children’s accounts, which was an issue raised in
court. Specific questions should then be used sparingly to follow up on forensically important information that was not mentioned in the narrative account (Benson & Powell, 2015a, 2015b; Cederborg et al., 2013; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007; Powell et al., 2005), and interviewers should not interrupt a child’s narrative account (Powell & Thomson, 2001).

Second, the quality of interviews may also be improved through increased interdisciplinary communication and collaboration between legal professionals, police, and other professionals involved in child witness interviewing (McConachy, 2002; Powell, 2008; Victorian Law Reform Commission, 2004). Research has suggested that part of the reason for “overzealous” questioning by interviewers is a misunderstanding of what is legally required for a successful prosecution, and a lack of feedback from legal professionals about the child witness interviews (Burrows et al., 2013, p. 269; Hoyano & Keenan, 2010; McConachy, 2002). Third, with regards to technological problems, the broader literature suggests that such issues are frequently due to insufficient testing and planning (K. Walsh, 2014). Although the current study was not able to evaluate whether these factors contributed to technological problems experienced in court, it is possible that improved testing of equipment and familiarity with playing electronic interviews could reduce the incidence of technological problems with interviews at trial.

In conclusion, the findings from the current study are largely consistent with previous research concerning the utility of using child witness interviews as children’s evidence-in-chief in CSA trials. However, the current results also highlight the potential downstream effects that child witness interviews can have in these trials. Future research is now necessary to quantitatively examine
whether, or to what extent, interview quality influences both the conduct and outcome of CSA trials.
CHAPTER 8: CHILD WITNESS INTERVIEW QUALITY AND INCONSISTENCIES RAISED DURING CROSS-EXAMINATION: IS THERE A CONNECTION? (STUDY 3)

In adversarial legal systems, such as those that exist in all Australian jurisdictions, cross-examination is a central part of the trial (Bowden et al., 2014). Under adversarial theory, cross-examination is regarded as both the best method of testing the truthfulness and accuracy of a witness’ account, as well as a chance to “break down” the opposing lawyer’s case (Cossins, 2009; Henderson, 2008, p. 282). An adversarial criminal trial is not an inquiry into the truth; rather, it is a forum in which two competing parties “battle” to present their evidence in front of an impartial judge and decision-maker (the jury) (Cossins, 2006, 2010; Shead, 2014). As such, the underlying purpose of questioning in cross-examination is radically different to the purpose of investigative interviewing, as the latter is conducted on the basis of social science research with the aim of producing the most reliable and accurate testimony (Henderson, 2008).

Given this combative nature of cross-examination, it is unsurprising that cross-examination is one of the most traumatic aspects of a trial for complainants (Cossins, 2009; Eastwood & Patton, 2002). Numerous studies have shown that despite reforms to the way vulnerable witnesses can give evidence (“alternative measures”; e.g., via pre-recorded child witness interview; giving evidence via CCTV), cross-examination remains a source of great stress and anxiety for complainants, with a great potential for re-victimisation (Eastwood et al., 2006; Eastwood & Patton, 2002; Hamlyn et al., 2004). In one large UK survey of vulnerable witnesses, for example, the majority (71%) of 569 victims and
witnesses interviewed after the introduction of special measures stated they were upset by cross-examination, the same proportion as before the implementation of reforms (Hamlyn et al., 2004).

In CSA trials, where the complainant’s evidence is usually crucial to a successful prosecution (Cross et al., 1994; Hoyano & Keenan, 2010), cross-examinations can be particularly traumatic for complainants. This is partly because in such “word against word” cases, usually the defence lawyers’ only way of defending their client is by discrediting complainants and their testimony (Bowden et al., 2014; Zydervelt, Zajac, Kaladelfos, & Westera, 2017). One tactic used in cross-examination to discredit the complainant is to highlight any inconsistencies in the complainant’s account (E. Davies & Seymour, 1998; Hanna, Davies, Crothers, & Henderson, 2012; Mauet & McCrimmon, 2013; Zydervelt et al., 2017). Inconsistencies may arise between two different statements given by the same witness, between the witness’ testimony and another witness, or between the witness’ testimony and external evidence (N. Brewer, Potter, Fisher, Bond, & Luszcz, 1999). If a complainant brings up new information at a later point in time, this may also be treated as an inconsistency (N. Brewer et al., 1999; Burrows & Powell, 2014b; Mauet & McCrimmon, 2013). Inconsistencies about specific aspects of complainants’ evidence are then used to cast doubt on the entirety of their testimony (E. Davies, Henderson, & Hanna, 2010; Mauet & McCrimmon, 2013).

The tactic of using inconsistencies during cross-examination to discredit the complainant is based on the belief that true memories are complete in every detail, and inconsistencies in evidence mean the complainant is being untruthful (Cashmore & Trimboli, 2006; Cossins, 2009; Hanna et al., 2012). This belief is
evident in cross-examination manuals, which explain that “unreliability flows” from inconsistent statements (Selby, 2009, p. 169), and suggest that “in your closing argument, you can argue that the witness’s inconsistent statements impugn his or her credibility” (Mauet & McCrimmon, 2013, p. 221). It is further evident from research investigating beliefs regarding inconsistency and accuracy. For example, an informal survey of judges and lawyers found that these professionals strongly believe that inconsistency means inaccuracy, and that witnesses who make inconsistent statements should be regarded as less credible (Fisher & Cutler, 1995). Similar results have been found in experimental studies with mock jurors (undergraduate students), with mock jurors regarding inconsistent statements as strong markers of inaccurate evidence (Berman, Narby, & Cutler, 1995; N. Brewer et al., 1999; Connolly, Price, Lavoie, & Gordon, 2008).

Although inconsistencies may indicate deceit, the general proposition that inconsistencies regarding certain details indicates unreliability of the whole account is not supported by human memory research (N. Brewer et al., 1999). Experimental studies have shown that some inconsistency in testimony is normal, and that consistency of an account is only weakly, if at all, related to the overall accuracy of the account (N. Brewer et al., 1999; E. Davies & Seymour, 1998; Fisher, Brewer, & Mitchell, 2009; Oeberst, 2012; Smeets, Candel, & Merckelbach, 2004). Core or central information tends to be better remembered than peripheral information (Brown et al., 2008; Toth & Valentino, 2008); thus, inconsistencies in the recall of peripheral details (such as sequences, dates, specific locations or times of events) are a common and normal occurrence (Connolly et al., 2008; E. Davies & Seymour, 1998; Fisher et al., 2009). The
accuracy of reminiscent details (that is, details remembered in a second account that were missing in the first) has also been shown to be quite high (Krix, Sauerland, Lorei, & Rispens, 2015), further contradicting the notion that a witness who recounts new information in a second interview should be assigned lower credibility.

Nevertheless, defence lawyers take advantage of the likelihood that a witness will contradict him- or herself when recalling peripheral details, and specifically conduct their questioning in the hope of eliciting inconsistencies (Cossins, 2009; E. Davies & Seymour, 1998; Eastwood & Patton, 2002; Hanna et al., 2012; Kebbell, O’Kelly, & Gilchrist, 2007). In the words of one Australian lawyer and academic, “it is [an] effective tactic to diminish a child’s testimony to engage in subtle questioning. … One can do that by focusing on peripheral events [and o]ne can focus on minor inconsistencies of the child’s testimony” (Standing Committee on Law and Justice, 2002, p. 64). Focusing on inconsistencies and peripheral details is a very common tactic; studies that have examined trial transcripts from the United Kingdom and New Zealand have reported that defence lawyers questioned complainants on peripheral details in between 60% to 85% of cases (E. Davies & Seymour, 1998; G. Davies et al., 1995; Hanna et al., 2012). The tactic is also very effective; a series of studies by Zajac and colleagues demonstrated that, regardless of the accuracy of the initial response, a large percentage (79-85%) of both younger (5-6 year old) and older (9-10 year old) children changed at least one of their earlier responses during cross-examination (Zajac, Gross, & Hayne, 2003; Zajac & Hayne, 2006). These findings appear to be due specifically to cross-examination style questioning, not merely to the effects of repeated interviewing (Fogliati & Bussey, 2014).
The tactic of focusing on minor inconsistencies, and cross-examination questioning in general, has been subjected to significant critique. For example, Cossins (2009) argued that rather than uncovering dishonesty and true errors in testimony, cross-examination questioning and tactics actually produce inaccurate testimony. Researchers from the social sciences also frequently critique cross-examination questioning for failing to abide by best-practice interviewing guidelines (Andrews & Lamb, 2016; Righarts, O'Neill, & Zajac, 2013; Zajac, O'Neill, & Hayne, 2012). These guidelines suggest that children should be questioned using non-leading open-ended questions (e.g., “Tell me what you remember…”), specific questions should only be used once open-ended questions have been exhausted, and leading questions should be avoided altogether (Lamb & Fauchier, 2001; Powell, 2002; Powell & Snow, 2007).

In two studies analysing children’s self-contradictions and the question types employed by lawyers in Scottish and Californian CSA trials, Andrews and colleagues (2015, 2016) reported that lawyers virtually never used open-ended prompts, and defence lawyers used a significant number of suggestive questions that resulted in more self-contradictions than any other question types. The authors concluded that legal practitioners require training to learn how to avoid suggestive questions, and how to “safely” pose questions to children. This conclusion, however, ignores the adversarial nature of trials explained above, and the professional duty of defence lawyers to advance their client’s case (Bowden et al., 2014). It is unlikely that, given the current combative nature of adversarial trials, education will lead to defence lawyers questioning children in accordance with best practice interviewing guidelines. As described above, eliciting inconsistent testimony is a deliberate tactic. In fact, as argued by Henderson
(2008), “education may even backfire as it enables lawyers to design more sophisticated and discrete methods of controlling witnesses.” Thus, given the current system, it may be more advantageous to examine whether investigative interviewers can amend their questioning techniques to reduce the number of inconsistencies raised during the cross-examination of complainants.

Investigative interviewers are trained how to question children in accordance with best-practice guidelines; however, adherence to best-practice interviewing in the field is generally poor (Lamb, 2016; Luther et al., 2015; Powell, Cavezza, et al., 2010; Powell, Murfett, et al., 2010; Wolfman et al., 2016). Given that open-ended questions have been shown to elicit more accurate responses than specific questions (Dent & Stephenson, 1979; Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007; Powell & Thomson, 2001; Sternberg et al., 1996), it would follow that answers to open-ended questions should be less likely to be the subject of inconsistencies during cross-examination. A field study examining self-contradictions made by seven children in investigative interviews of alleged sexual abuse supports this supposition (Lamb & Fauchier, 2001). The study found that open-ended questions were significantly less likely to lead to self-contradictions than specific or leading questions. In fact, the authors reported that open-ended questions did not elicit a single self-contradiction.

Another area of concern regarding the effect of questioning in investigative interviews on trials, which has been highlighted in a series of interview and focus group studies with Australian prosecutors, is the number of minute details elicited by interviewers and the types of topics (such as sexual acts and genitalia) excessively followed up (Burrows & Powell, 2013, 2014a, 2014b, 2014c, 2014d, 2015). Although some academics have argued that maximising the
number of details collected in interviews is beneficial (E. Davies & Seymour, 1998; Hershkowitz, 2001; Westera et al., 2013b), this view is not shared by prosecutors. These professionals argue that interviewers must be extremely careful about what detail is sought and followed up, as complainants can be cross-examined on anything that was said during an interview (Burrows & Powell, 2015). As a result, even errors in irrelevant or minute details can be used to undermine the reliability and credibility of the complainant’s entire account (Burrows & Powell, 2014b).

The research on question types and qualitative studies with prosecutors suggests that questioning techniques employed by investigative interviewers might influence how children are cross-examined. However, no field study has yet examined whether any such links exist. The aims of the current study were therefore to (1) determine whether questioning in investigative interviews is related to cross-examination inconsistencies in actual cases of alleged CSA, and (2) explore the nature of the inconsistencies raised.

For the first aim, two competing hypotheses were made, given the different theories on the purpose of cross-examination. Based on the theory that the purpose of cross-examination is primarily to discredit the witness, and raising inconsistencies is a tactic employed to do so, it was hypothesised that open-questions should lead to more inconsistencies than specific or leading questions. This association should occur because answers to open-questions have been shown to be significantly longer and more detailed, and thus should provide more information for defence lawyers to target for inconsistencies. Based on an alternate theory that the purpose of cross-examination is to uncover untruthful or inaccurate information, however, it was hypothesised that open-ended questions
should be associated with fewer inconsistencies. This association should occur because, according to the psychological literature on question types and accuracy, open questions should promote more accurate evidence, which in turn, should lead to fewer inconsistencies.

For the second aim, the importance (central/peripheral) and content (topic) of the inconsistencies were examined. It was hypothesised that there should be significantly more peripheral than central inconsistencies; however, no hypotheses were made about the content of the inconsistencies and only exploratory analyses were conducted.

**Method**

**Trial and Child Witness Interview Transcripts**

The cross-examination transcripts for the original sample of 120 complainants described in Study 1 were coded for a number of cross-examination tactics including inconsistencies (Powell et al., 2016). Complainants met the inclusion criteria for the current study if they either used their child witness interview as their evidence-in-chief, or were cross-examined on a child witness interview they had previously given. Thirteen of these 86\(^5\) eligible complainants were missing all of their child witness interviews and were thus excluded from the study. This resulted in a final sample of 73 complainants (15 male; 58 female), ranging in age 7 to 19 years (\(M = 13.34; \ SD = 3.35\)), of whom nearly all (\(n = 72\)) used their child witness interview as evidence-in-chief. Of these complainants, a second child witness interview was missing for 3 complainants but

---

\(^5\) This included the complete sample discussed in study 1, plus one additional complainant who had turned 19 and not used her CWI as evidence-in-chief, but was cross-examined on a CWI they had previously given.
Procedure

Inconsistencies coding scheme. Inconsistencies were identified in complainants’ cross-examinations using a cross-examination tactics coding scheme. This coding involved identifying different types of tactics, including tactics questioning the reliability, credibility, and plausibility of complainants’ evidence, as well as identifying inconsistencies in complainants’ accounts. The scheme was originally developed for adult rape complainants (Zydervelt et al., 2017) and was modified for the use in CSA cases using a thematic analysis (for the full coding scheme, and a description of all tactics, see Powell, Westera, et al., 2016).

To capture the complexity of cross-examination, *lines of questioning* rather than individual questions were coded. Thus, if the defence spent three questions highlighting an inconsistency in the complainant’s account, and then five questions asking the complainant why they did not call out for help, this would be coded as two tactics. Further, one line of questioning could be coded as containing multiple tactics, for example, an inconsistency and a query as to the complainant’s reliability. Lastly, the tactics were inferred from the defence lawyer’s questions, rather than the complainant’s answers. Thus, if the defence stated that the complainant had made two inconsistent statements but the complainant disputed this, it was nonetheless coded as the tactic of raising an inconsistency. This decision was made because the defence lawyers’ tactics, rather than the complainant’s responses, were under investigation.
For the current study, only lines of questioning that included inconsistencies (from here on referred to as “inconsistencies”) were examined. The inconsistencies were further coded as to their type: additions (information that was not raised in statement 1 but raised in statement 2), omissions (information that was raised in statement 1 but not raised in statement 2), and contradictions (information that was raised in two statements, but differed). The source of the inconsistency was also identified, and recorded as being with the complainant’s own statement, a witness, other evidence, or the accused. Where multiple sources were inconsistent, the first two sources that were inconsistent were used in the analyses for the current study.

Inconsistencies originating from within the complainant’s own statement were further identified as being from (1) within one child witness interview, (2) between two child witness interviews, (3) between a child witness interview and the complainant’s live evidence-in-chief (usually in the form of additional questions asked by the prosecutor), (4) between a child witness interview and cross-examination, (5) between a child witness interview and a police statement (some complainants had both), (6) between three or more own statements, or (7) between the child witness interview and a statement by the complainant to a “trusted person”. This final category was created because a statement to a professional (such as a police officer) was deemed much more credible than a statement to, for example, a friend. Thus, inconsistencies to such “trusted people” (which included doctors, police officers, counsellors, and legal professionals) were deemed to be inconsistencies-with-self, as well as inconsistencies with a witness, to capture this important distinction.
Once the source of the inconsistencies was identified, inconsistencies that did not involve a question in an interview (for example, a complainant said something in cross-examination that she did not say in her interview but was highlighted by the defence) were excluded. This was due to the focus of the current study on the relationship between inconsistencies and complainants’ child witness interviews. The remaining inconsistencies were then coded for (1) the content of the evidence relating to the inconsistency, and (2) the significance of the inconsistency.

**Evidential coding.** The content of the evidence relating to the inconsistency was coded using categories established in earlier research in consultation with prosecutors and detectives (Benson & Powell, 2015b; Burrows et al., 2013). Apart from “other”, which was included here to capture content not falling into any of the other categories (cf. Study 1), these evidential categories reflect information required for the successful investigation and prosecution of CSA cases. Each category, along with an explanation of the code and an example inconsistency for each, are shown in Table 8.1. Quotations presented here and in the rest of the study were corrected for ease of reading; no other changes were made.
<table>
<thead>
<tr>
<th>Evidential Code</th>
<th>Explanation</th>
<th>Example Inconsistency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offence</td>
<td>Content relating to the offence; including descriptions of the act and body positions</td>
<td>Q: He didn’t grab your breast, did he?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A: No, but it was under my top, under my bra.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q: Because you did use the word “grab the boob” in the interview with the officers, but he didn’t grab it, did he?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A: He had his – he didn’t like grab it, but he had his hand there.</td>
</tr>
<tr>
<td>Time</td>
<td>Content relating to the timing of the offence, including the precise timing of the offence and the frequency of offending</td>
<td>Q: [Complainant], you say that it only happened when [accused's son] wasn’t there. That’s correct, isn’t it?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A: Yes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q: And you only went there when [accused's son] wasn’t there on two or three occasions. Is that a yes?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A: Yes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q: But you say it happened 20 or 30 times?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A: Yes.</td>
</tr>
<tr>
<td>Location</td>
<td>Content relating to the location of the offending, both specific (e.g., which bed) and general (e.g., which house).</td>
<td>Q: This isn't a matter of, I suggest to you, of adding more detail. This is telling a different story, because you've said that your mum came in after her shower and you were on her bed [after the offence]. Not your bed. Her bed?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A: Yeah.</td>
</tr>
<tr>
<td>Witnesses</td>
<td>Content relating to the presence or whereabouts of any potential witnesses</td>
<td>Q. Now, when [accused] touched you did you tell your parents straight away?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A. Yes.</td>
</tr>
<tr>
<td>Physical Evidence</td>
<td>Content relating to the presence of forensic evidence, such as possible DNA or other physical evidence</td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Q. Do you remember telling the interviewer that you forgot to tell your parents?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. I don’t know.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q. What were you wearing in bed?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. My nightgown.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q. Your nightie?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Yes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q. The police officer asked you what clothes you had on that night and you said, &quot;Dressing gown&quot; and then he asked you to describe that to him and you said it was blue.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Yes.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Offender Identification</th>
<th>Content relating to determining the identity of the offender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q. How did you know he was coming into your room?</td>
<td></td>
</tr>
<tr>
<td>A. Because I saw his face.</td>
<td></td>
</tr>
<tr>
<td>Q. By he you're referring to [accused]?</td>
<td></td>
</tr>
<tr>
<td>A. Yes.</td>
<td></td>
</tr>
<tr>
<td>Q. You just happened to notice [accused] walking into your room. Is that right?</td>
<td></td>
</tr>
<tr>
<td>A. Yes.</td>
<td></td>
</tr>
<tr>
<td>Q. It wasn't because you heard the door creaking like you claimed in the first interview you gave to the police?</td>
<td></td>
</tr>
<tr>
<td>A. No, I did hear the door creaking.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other</th>
<th>Content concerning miscellaneous topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q. I’d suggest that you told nothing to the Senior Constable about the situation where your stepfather came in and was searching under your pillow or feeling under your pillow for the I-Pod, you didn’t tell him anything about that at all do you remember or not--</td>
<td></td>
</tr>
<tr>
<td>A. Okay. No I don’t remember.</td>
<td></td>
</tr>
</tbody>
</table>
Significance of inconsistency. The usefulness of each inconsistency in proving the charged offence/s was coded in light of the facts of each particular case. Inconsistencies were coded as “central” if they related to content that could do one or more of the following: (1) prove an element of the offence (e.g., mens rea or actus reus of the offence), (2) directly corroborate the offence (e.g., a witness to the offence or absence of offending), (3) provide a defence for the accused, or (4) provide evidence of an initial complaint (that is, the first person to whom the complainant disclosed the offence). Inconsistencies that were not “central” were coded as “peripheral”. In some instances, an inconsistency could be either central or peripheral depending on the facts of a case. For example, the complainant’s exact age at the time of the offence would be coded as central if it were relevant to the age element of the offence, but otherwise would be coded as peripheral. Examples of central and peripheral inconsistencies are shown in Table 8.2.

Table 8.2
Examples of Central and Peripheral Inconsistencies

<table>
<thead>
<tr>
<th>Category</th>
<th>Content Topic</th>
<th>Example from Cross-examination transcript</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central</td>
<td>If the accused touched his penis</td>
<td>Q. Do you remember watching the video that you said when the interviewer asked you, she asked you this question, &quot;Do you remember the first time he touched your peanut?&quot; and you answered &quot;No, he don't touch me,&quot; do you remember that happening in the video? A. I don't know.</td>
</tr>
<tr>
<td></td>
<td>Whether accused penetrated her vagina</td>
<td>Q: “Did he touch you on the inside or the outside of your rude parts?” - “The outside”? A: Yes.</td>
</tr>
</tbody>
</table>
Q: And that's correct? That's what you told the police?
A: Yes.

Q: And your memory now is different to that, is that right?
A: Yes.

| Peripheral Colour of the car | Q. Now, you said to me, when I asked you about the colour of the car, you indicated that the car was white?
A: Yes.
Q: And if you go to page 94 of that second interview, at question 30 where it says:
“So what window could you see out of?”
Your answer: “I didn’t see out of the window. It was a red car.” |

| The type of bathers she was wearing | Q. Okay. So which version is accurate? Were you wearing a shirt and board shorts or were you wearing a tankini and bikini bottoms?
A: The tankini. |

Trial characteristics. As in Study 1, the number of defence lawyers’ questions posed to the complainant during cross-examination were tallied as a measure of cross-examination length.

**Child witness interview coding.** Child witness interview questions were coded as open, specific, or leading, as described in Study 1. Next, for each inconsistency for which the source was a child witness interview, the question that led to the inconsistency was identified and its type (open, specific, or leading) recorded. Where an inconsistency was between two interviews or between two questions within an interview, the questions that were inconsistent were both coded. However, there were not enough questions from second child witness
interviews in each of the three question code categories to conduct any analyses (only 9 open and 2 leading, compared to 79 specific).

In cases where multiple questions in one or more interviews were inconsistent with cross-examination or another source, and the defence did not specify the exact question they were referring to, the first question that was inconsistent was coded. For example, if the complainant stated in cross-examination that she was wearing pants at the time of the offence, but she told the interviewer in response to two separate questions that she was wearing a skirt, the first question in the interview to which the complainant replied that she was wearing a skirt was used. Finally, when interviewers asked the complainant multiple questions without allowing the complainant to answer, the question directly preceding the complainant’s answer was coded. Table 8.3 provides some examples of how interview questions were linked with inconsistencies and coded.

Table 8.3
Examples of Inconsistencies and Corresponding Interview Questions

<table>
<thead>
<tr>
<th>Inconsistency in Cross-examination Transcript</th>
<th>Interview Transcript</th>
<th>Question type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q: Whose bedroom did he come into?</td>
<td>Q: What sort of bed was it? A: It was my brother and his girlfriend's bed.</td>
<td>Specific</td>
</tr>
<tr>
<td>A: It was my mum’s.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q: Wasn’t it in fact that you told the police in your interviews, in one of your interviews that in fact he didn’t put his hand down the front of you? Do you recall that? A: I can’t remember saying that.</td>
<td>Q: What did he do? A: He put his hand up the leg of my shorts.</td>
<td>Open</td>
</tr>
<tr>
<td></td>
<td>Q: Okay. If you can look at your interview … on page 71?</td>
<td></td>
</tr>
</tbody>
</table>
A: Yes.

Q: “Did [accused] take off your shorts?” “Nup”. That’s question 49. Question 50: “What did he do?” “He put his hand up the leg of my shorts”? A: Yes.

Q: Well, that’s different to what you’re saying today, isn’t it? A: Yes.

Q: And in fact it’s the opposite direction of what you’re saying today, isn’t it? A: Yes.

Q: So can you really remember which way his hand went, whether it went up your leg to your rude part, or down the front to your rude part? A: No.

Q: Before [accused] did this thing to you, you were actually asleep? A: No, I wasn’t.

Q: Do you think this happened, [complainant], or did [accused] lick your private parts? Can you remember that it actually happened? … A: I just feel I think I fell asleep and then I woke up and then I felt something and then I woke up and then I saw him doing it.

Finally, the total length of the child witness interview and the total number of questions asked of the complainant in each child witness interview were recorded, as were the total number of abusive incidents discussed in the interview.
For complainants who had multiple interviews, composite measures of total length, total number of questions, and total number of incidents across all interviews were created.

Reliability. Twenty percent of the transcripts were independently coded by two researchers for cross-examination tactics, inconsistency significance, and type. They agreed on 80% of the tactics used, and interrater reliability for significance was substantial with 93% agreement ($\kappa = 0.74$) and 94% agreement for type ($\kappa = 0.75$). All disagreements were resolved by discussion, and the two researchers coded half of the remaining transcripts each.

Evidential content of the inconsistencies was coded by one researcher, and another researcher independently coded 20% of the inconsistencies for content producing a Cohen’s Kappa coefficient of $\kappa = 0.88$. As in Study 1, question types in child witness interview transcripts were all coded by one person who had previously obtained interrater reliability with a second coder. These coders had a very high interrater agreement ($\kappa = 0.96$).

**Results**

The cross-examinations of the 73 complainants included a total of 1087 inconsistencies between the child witness interview and at least one other source. For four complainants, no inconsistencies arose from their child witness interview. The frequency and percentage of all inconsistencies are shown for each source in Table 8.4. As can be seen, most child witness interview inconsistencies arose from the complainant’s own statement, followed by the statement of a witness. Nearly all inconsistencies were contradictions, as shown in Table 8.5.
Table 8.4

Source of Inconsistencies Which Included the Child Witness Interview

<table>
<thead>
<tr>
<th>Source</th>
<th>Frequency</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own Statement</td>
<td>652</td>
<td>60.00</td>
</tr>
<tr>
<td>Witness</td>
<td>383</td>
<td>35.00</td>
</tr>
<tr>
<td>Other evidence</td>
<td>57</td>
<td>5.20</td>
</tr>
<tr>
<td>Accused</td>
<td>120</td>
<td>11.00</td>
</tr>
</tbody>
</table>

Note: Percentages do not add up to 100 as inconsistencies could have more than one source.

Table 8.5

Type of Inconsistency

<table>
<thead>
<tr>
<th>Type of Inconsistency</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contradiction</td>
<td>883</td>
<td>81.20</td>
</tr>
<tr>
<td>Omission</td>
<td>80</td>
<td>7.40</td>
</tr>
<tr>
<td>Addition</td>
<td>113</td>
<td>10.40</td>
</tr>
<tr>
<td>Omission &amp; Addition</td>
<td>1</td>
<td>.10</td>
</tr>
<tr>
<td>Omission &amp; Contradiction</td>
<td>3</td>
<td>.30</td>
</tr>
<tr>
<td>Addition &amp; Contradiction</td>
<td>7</td>
<td>.60</td>
</tr>
</tbody>
</table>

The breakdown of the sources of inconsistencies within the complainant’s own statement is shown in Table 8.6. This demonstrates that the clear majority of own-statement inconsistencies arose between the child witness interview and cross-examination.
Table 8.6

Source of Own-Statement Inconsistencies

<table>
<thead>
<tr>
<th>Source</th>
<th>Frequency</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within CWI</td>
<td>49</td>
<td>7.5</td>
</tr>
<tr>
<td>Between 2 CWIs</td>
<td>100</td>
<td>15.3</td>
</tr>
<tr>
<td>Between CWI &amp; EIC</td>
<td>42</td>
<td>6.4</td>
</tr>
<tr>
<td>Between CWI &amp; Cross</td>
<td>393</td>
<td>60.3</td>
</tr>
<tr>
<td>Between CWI &amp; statement</td>
<td>15</td>
<td>2.3</td>
</tr>
<tr>
<td>Between 3 or more own statements</td>
<td>37</td>
<td>5.7</td>
</tr>
<tr>
<td>Between CWI and statement to trusted person</td>
<td>16</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Note: Percentages do not add up to 100 as inconsistencies could have more than one source.
“CWI” = Child Witness Interview; “Cross” = cross-examination; “EIC” = Evidence-in-chief (additional questioning by prosecutor).

Relationship Between the Child Witness Interview and Inconsistencies

To determine whether certain question types were more likely to lead to inconsistencies during cross-examination than others, a chi-square goodness of fit test was conducted. Given that child witness interviews did not include an equal proportion of open, specific, or leading questions, the mean percentage of each question type from the sample of interviews was entered as the expected frequency (12.70% open; 78.27% specific, and 8.98% leading). The chi-square test revealed that these three question types were not equally as likely to lead to inconsistencies, $\chi^2(2) = 178.56, p < .001$, Cramer’s $V = .29$. Standardised residuals indicated that open-ended questions were significantly more likely ($z = 12.48, p < .001$), specific questions significantly less likely ($z = -4.64, p < .001$), and leading questions no more likely ($z = -1.15, p = .13$), to lead to inconsistencies than expected.
Follow-up chi-square tests were conducted to determine whether these three question types significantly differed from each other in leading to inconsistencies. Expected frequencies were recalculated for each test, with the question type not being examined excluded from the percentage calculation (see Table 8.7). Results showed that open-ended questions were more likely than both specific and leading questions to lead to inconsistencies, $\chi^2(1) = 175.13, p < .001$, Cramer’s $V = .54$ and $\chi^2(1) = 50.38, p < .001$, Cramer’s $V = .36$, respectively. However, specific questions were no more likely than leading questions to result in inconsistencies, $\chi^2(1) = 0.19, p = .66$.

<table>
<thead>
<tr>
<th>Table 8.7</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expected Frequencies Used in Follow-Up Chi-Square Analyses</strong></td>
</tr>
<tr>
<td>Question type comparison</td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td>Open vs Specific</td>
</tr>
<tr>
<td>Open vs Leading</td>
</tr>
<tr>
<td>Specific vs Leading</td>
</tr>
</tbody>
</table>

To further examine the relationship between child witness interviews and inconsistencies, correlational analyses were conducted between the number of child witness interview inconsistencies, the overall length of the child witness interview, and the total number of questions contained in the child witness interview. To control for differences in length and number of questions due to the number of abusive incidents discussed in interviews, three new variables were computed (inconsistencies/number of incidents, length/number of incidents, and total number of questions/number of incidents). Only cases for which the number of incidents, questions, and length were available were included in the analyses.
(n = 57 complainants). Spearman correlations\(^6\) showed that when number of incidents was controlled for, the number of inconsistencies in cross-examination that arose from the child witness interview were significantly associated with the overall length of the interview ($r_s = .41, p < .01, R^2 = .16$), but not the total number of questions asked ($r_s = .23, p = .09$).

Finally, correlational analyses were conducted to determine whether the number of child witness interview inconsistencies that arose was also related to cross-examination length. Spearman’s correlations\(^7\) showed that as the number of child witness interview inconsistencies increased, so did the length of cross-examination ($r_s = .70, p < .001, R^2 = .49$).

**Nature of the Inconsistencies**

Of the 1087 inconsistencies, a total of 807 (74.2%) were about peripheral details and 280 (25.8%) were about central details. A chi-square test showed that these frequencies were significantly different than what would be expected by chance (50%), $\chi^2(1) = 255.50, p < .001$, Cramer’s $V = 0.48$.

The content of inconsistencies, in terms of important evidential categories identified by prosecutors and detectives, is shown in Table 8.8. About one-third of inconsistencies centred on the offence, and over one-fifth concerned topics that were not covered by the evidential coding scheme.

---

\(^6\) Nonparametric correlations were used as the assumption of normality was violated for all three variables. The Kolmogorov-Smirnov test was significant for all three variables, all $ps < 0.01$. In addition, the total number of questions/incidents and number of inconsistencies/incidents variables were also significantly skewed (Zskew = 6.16 and Zskew = 5.34, respectively) and kurtosed (Zkurtosis = 6.71 and Zkurtosis = 5.55, respectively).

\(^7\) Again, the assumption of normality was violated for both variables. The Kolmogorov-Smirnov test was significant at $p < 0.01$. In addition, the both variables were also significantly skewed (Cross-examination length Zskew = 10.94 and CWI Inconsistencies Zskew = 11.93) and kurtosed (Zkurtosis = 25.74 and Zkurtosis = 5.41, respectively).
Table 8.8

*Content of Inconsistencies*

<table>
<thead>
<tr>
<th>Evidential Code</th>
<th>Frequency</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offence</td>
<td>360</td>
<td>33.12</td>
</tr>
<tr>
<td>Timing</td>
<td>168</td>
<td>15.46</td>
</tr>
<tr>
<td>Location</td>
<td>88</td>
<td>8.10</td>
</tr>
<tr>
<td>Witness</td>
<td>194</td>
<td>17.85</td>
</tr>
<tr>
<td>Physical evidence</td>
<td>45</td>
<td>4.14</td>
</tr>
<tr>
<td>Offender identification</td>
<td>4</td>
<td>0.37</td>
</tr>
<tr>
<td>Other</td>
<td>232</td>
<td>21.34</td>
</tr>
</tbody>
</table>

**Discussion**

The present study examined whether questioning in investigative interviews is related to cross-examination inconsistencies in cases of alleged CSA, and explored the nature of these inconsistencies in terms of their topic and importance. Two competing hypotheses were made regarding the relationship between questioning and inconsistencies, given the different theories on the purpose of cross-examination. The first hypothesis – that open-questions would lead to more inconsistencies than specific or leading questions – was supported. The alternate hypothesis, that open-ended questions would result in fewer inconsistencies, was thus not supported. Regarding the prevalence of central and peripheral inconsistencies, results showed that there were more inconsistencies about peripheral details, which supports the third hypothesis.

The finding that open-ended questions resulted in more inconsistencies in cross-examination than either specific or leading questions is contrary to what would be expected from the memory and interviewing literature. Experts from
these areas have long assumed that “the more complete and accurate the initial investigative interview, the more complete and accurate the child's story during examination-in-chief [which], in turn, makes the child's account in cross-examination less susceptible to distortion” (Powell, 2005, p. 137; see also, Lamb & Fauchier, 2001). As a result, memory and interviewing research with children has generally focused on how to obtain the most accurate and detailed statements possible (Lamb, 2016; Powell & Snow, 2007). Given that both laboratory and field studies found that open-ended questions result in the longest, most detailed and most accurate answers (Brown et al., 2013; G. Davies et al., 2000; Dent & Stephenson, 1979; Hershkowitz, 2001; Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007; Phillips et al., 2012; Sternberg et al., 1996), open-ended questions are thus recommended above more specific questions in interviewing guidelines.

The limited number of studies that have examined inconsistencies associated with question types in children’s statements appear to support the superiority of open over leading and closed questions in reducing inconsistencies (Andrews et al., 2015; Lamb & Fauchier, 2001; Orbach & Lamb, 2001). For example, Lamb and Fauchier (2001) examined a sample of 24 interviews with 7 children, and found that open-ended questions never elicited self-contradictions, and were significantly less likely to elicit information that was later contradicted. However, all of these studies focused solely on self-contradictions. None examined inconsistencies raised during cross-examinations, which not only include self-contradictions, but also contradictions with other sources or witnesses, as well as omissions from, and additions to, statements. Further, these studies failed to consider the adversarial nature of trials, in which not just true inconsistencies can discredit a witness, but any inconsistencies argued by the
defence – whether they are true, accepted by the complainant, or not. As such, interviewing literature has assumed that an accurate statement can protect a complainant from being undermined during cross-examination, which is not supported by the current results (Henderson, 2008).

Psychological researchers and police have often also equated “good” interviews with interviews that maximised the number of details collected (E. Davies & Seymour, 1998; Gagnon & Cyr, 2017; Lamb, 2016; Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007; Westera et al., 2013b). In fact, the number of details elicited has frequently been used as a means of evaluating whether an interview has been successful (Benson & Powell, 2015b; Brown & Lamb, 2015; Brown et al., 2013; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007; Luther et al., 2015; Yi et al., 2016). However, legal professionals have cast doubt on this supposition. For example, several focus group studies with prosecutors and interviews with legal professionals (including prosecutors, defence counsel and a judge) have suggested that too much detail in investigative interviews is damaging to a successful prosecution, as defence lawyers can cross-examine a complainant on anything said in an interview (Burrows & Powell, 2013, 2014a, 2014d, 2015; Guadagno et al., 2006). The findings of the current study support this proposition.

It is likely that open-ended questions resulted in more inconsistencies in the present study because answers to these questions were longer and provided more information – including peripheral information – which was then used by the defence to create or highlight inconsistencies during cross-examination. This explanation is supported by the finding that the number of child witness interview inconsistencies was highly correlated with the length of the interview (once
The number of abusive incidents was controlled for, but the total number of questions asked was not. The difference between these correlations is important because it suggests that it is not purely the number of questions asked, but length of the answer, which is associated with an increased number of inconsistencies during cross-examination. Given the strong empirical base showing open-ended questions result in more accurate answers than specific or leading questions (Cole & Loftus, 1987; Dent, 1986; Dent & Stephenson, 1979; Hutcheson et al., 1995; Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007), these findings also support assertions from some legal academics that cross-examination is focused more on undermining the credibility of the complainant than uncovering inaccurate testimony (Cossins, 2009; Henderson, 2008).

The second finding from the current study, that significantly more inconsistencies involved peripheral than central details, further backs this notion that highlighting inconsistencies is a deliberate tactic employed to discredit the complainant. It supports past research, which has consistently shown that most cross-examination inconsistencies focus on peripheral details (E. Davies & Seymour, 1998; G. Davies et al., 1995; Hanna et al., 2012), and academics who have suggested that focusing on peripheral details is a deliberate cross-examination technique (E. Davies & Seymour, 1998; Hanna et al., 2012; Kebbell et al., 2007).

The implications of these first two findings for investigative interviewers are significant. First, they highlight the importance of keeping interviews as concise as possible, and focussed around central details. Central details are less likely to be forgotten (Brown et al., 2008; Toth & Valentino, 2008), and short interviews give the defence less information in which to find, or with which to
create, inconsistencies. This, in turn, should also lead to complainants experiencing shorter cross-examinations, given the finding that the number of inconsistencies was significantly related to length of cross-examination. Translating these recommendations into practice may necessitate some form of prosecutor instruction to clarify what information is required from an evidential perspective (Burrows et al., 2013).

Second, the long-held maxim – that open-ended questions are least likely to cause the complainant difficulties in court – is not supported by the current findings. Simply increasing the proportion of open-ended questions employed by interviewers is – in isolation – unlikely to “improve[e] the suitability of the police interviews as evidence” as has been suggested in prior research (Westera, Powell, & Milne, 2017, p. 264). Reducing the number of inconsistencies raised during cross-examination, while still maintaining the integrity (in terms of accuracy and reliability) of the evidence, will most likely involve the adoption of more focused open-ended questions. Prior research with prosecutors has already suggested that broad open-ended questions that elicit lengthy and elaborate details are problematic (Burrows & Powell, 2014b), while some studies examining the utility of question types have found that certain groups of witnesses (such as young children and children with intellectual disabilities) respond better to more focused open-ended questions (Brown, Lewis, Lamb, & Stephens, 2012; Gagnon & Cyr, 2017; Lamb et al., 2003). Encouraging more focused open-ended questions to be used with all children would thus uphold key aspects of the interviewing literature, and may benefit children in court.

The final finding of this study concerned the content of inconsistencies in terms of evidential categories required for the successful investigation and
prosecution of CSA cases (Benson & Powell, 2015b; Burrows et al., 2013).

Burrows’ study had found that prosecutors, and interviewers who had received brief prosecutor instruction, perceived information about timing, offender identity and witnesses as less necessary to follow up in interviews than interviewers who had received no instruction. Current findings suggest, however, that these categories are not those that are most problematic in terms of leading to excessive inconsistencies during cross-examination. Most inconsistencies in the current study concerned the offence, followed by “other”. The three categories identified as “unnecessary” in the Burrows et al (2013) study accounted for just one third of inconsistencies in the current study.

It is a limitation of the current study that the content of inconsistencies was coded using categories established by prior research, rather than through a content analysis. As one fifth of the inconsistencies concerned topics not covered by the coding scheme, the conclusions that can be drawn from the current findings regarding the topics of inconsistencies are somewhat limited. Given that prosecutors have frequently critiqued investigative interviews for including (from an evidential perspective) too much irrelevant information (Burrows & Powell, 2013, 2014a, 2014b, 2014d, 2015), further research is necessary to determine – more precisely – what type of information is most likely to be the subject of inconsistencies during cross-examination. However, current findings suggest that a large proportion of inconsistencies revolve around content that is not necessary from an evidential perspective. Thus, if this content could be identified in future research, interviewers could then adjust their questioning to reduce the amount of information available to the defence to undermine complainants during cross-examination.
In conclusion, the current study showed that significantly more inconsistencies arose from open-ended questions during the child witness interview than either specific or leading questions. Most of these inconsistencies concerned the offence and involved peripheral details. The findings suggest that long interviews containing too many details are problematic, and interviewers should aim to ask focused rather than broad open-ended questions. Further research is now necessary to determine whether, and if so how, the quantity and type of child witness interview inconsistencies raised during the cross-examination of complainants influences the outcomes of their trials.
CHAPTER 9: THE CONTRIBUTION OF INTERVIEW QUALITY AND INCONSISTENCIES IN PREDICTING VERDICTS IN CASES OF CHILD SEXUAL ABUSE (STUDY 4)

This chapter presents the fourth and final study in this thesis. Study 4 examined the relationship between interview quality (using variables introduced in Study 1), inconsistencies (Study 3) and trial outcome. Anecdotal evidence, mainly from interviews with Australian prosecutors, suggests that interview quality is a source of concern for legal professionals because it affects prosecutions. More specifically, prosecutors have expressed concerns that interviews are frequently too long and include too many minute and irrelevant details, which are then used by the defence to undermine complainants during cross-examination (Burrows & Powell, 2013, 2014a, 2014b, 2014d, 2015). This concern was supported by the results of Study 3, where it was shown that open-ended questions (which produce the longest and most detailed responses) were significantly more likely than other question types to lead to inconsistencies during cross-examination. However, whether these inconsistencies, and the quality of interviews (as defined in the interviewing literature), affects verdicts in CSA trials is largely unknown. This gap in the research exists despite the fact that child witness interviews form the central piece of evidence in most CSA prosecutions (Powell, Murfett, et al., 2010; Cross & Whitcomb, 2017).

To date, only one study from an adversarial legal system has investigated changes in case processing following the introduction of a best-practice interviewing protocol. Pipe et al. (2013) compared case outcomes and dispositions of 760 CSA cases with interviews before (1994 to mid-September
1997; 350 cases) and after (mid-September 1997 to 2000; 410 cases) the introduction of the NICHD interview protocol. The NICHD protocol is a flexible but structured guide on how to use best-practice techniques in interviews with children (Lamb et al., 2008; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007). After controlling for several case-related variables, including number of victims, abuse severity, and age at interview, results showed that charges were significantly more likely to be filed, and trials significantly more likely to lead to guilty verdicts, in cases after the protocol was introduced.

The authors interpreted their findings in light of past research, which had shown that NICHD protocol interviews improve the quality of victim statements by decreasing inappropriate (i.e., specific, leading) questions, increasing appropriate (i.e., open) questions and techniques, decreasing the length of interviews, and increasing the number of forensically relevant details (Cyr & Lamb, 2009; Hershkowitz, 2001; Lamb, 2016; Lamb, Orbach, Hershkowitz, Esplin, et al., 2007; Sternberg et al., 1996; Sternberg et al., 2001). Given the child witness interview is central to most CSA prosecutions (Cross et al., 1994; Hoyano & Keenan, 2010), the authors concluded that the better outcomes for the group of protocol interview cases was a direct result of better quality interviews.

However, there are several significant limitations to Pipe et al.’s (2013) study. First and foremost, the actual quality of the interviews in the sample was not evaluated. Interview quality is usually assessed by examining the proportion of question types employed by interviewers (e.g., open, specific, and leading). More recently, the amount of important evidential information sought by interviewers has also been included in these evaluations (Benson & Powell, 2015b; Burrows et al., 2013). Given that there are significant difficulties with the
application of best-practice interviewing skills learned during training into the field (Lamb, 2016), it does not necessarily follow that interviews in the “post-protocol” condition were better in quality than interviews in the “pre-protocol” condition. Such a conclusion could only be drawn if at least a subset of interviews in each condition had been coded for interview quality.

Further, interviewers were trained to use the protocol in a five-day workshop, after which they were considered “trained” for the remainder of the study (3.5 years). This is problematic because, for one, such intensive training courses have been found to be largely ineffective in improving the quality of investigative interviews (see Lamb, 2016 for review); and for another, for improvements in interviewing to be maintained over time, very specific types of training programs are required (spaced out, with multiple modules, and with intensive feedback over time: Benson & Powell, 2015b; Lamb, 2016). An alternative explanation for Pipe et al.’s (2013) findings is that the increase in the number of people charged and prosecuted for CSA offences merely reflects a trend in increased charging and prosecution rates that has been present over the last few decades (Australian Law Reform Commission, 2010; Eastwood et al., 2006). Thus, it is possible that post-protocol cases resulted in more charges simply due to this overall trend, not due to improvements in interview quality.

Another limitation of Pipe et al.’s (2013) study is that available corroborative evidence (medical, material, witnesses) was not controlled for when evaluating the outcome of the cases. Corroborative evidence – including eyewitnesses, physical evidence, number of victims, and (in some studies) medical evidence – is one of the key variables associated with decisions to charge and prosecute in CSA cases (Bradshaw & Marks, 1990; K. D. Brewer et al., 1997;
Cross et al., 1994; Davis et al., 1999; Gallagher, 1999; Lewis & Klettke, 2012; Powell, 2008). There is also some evidence that corroborative evidence increases conviction rates (Bradshaw & Marks, 1990; Palusci et al., 1999), although too few studies to date have examined factors related to conviction in actual CSA trials to draw any firm conclusions. Most evidence of factors related to verdicts comes from vignette studies with mock jurors, which have been shown to have very poor external validity (D. R. Read, Connolly, & Welsh, 2006). The only recent study that has examined factors predictive of jury verdicts in real CSA trials supports the importance of corroborative evidence in these cases. Blackwell and Seymour (2014) found that three variables – similar fact evidence (evidence of the defendant’s prior behaviour considered to be similar to that charged at the trial), witnesses to the offending, and positive medical or DNA evidence – significantly predicted guilty verdicts in their sample of 137 CSA trials in New Zealand. Thus, corroborative evidence should be taken into account in any study examining factors related to CSA trial outcomes.

One final limitation of Pipe et al.’s (2013) study is that only 23 (6%) of the 364 cases in which charges were filed were resolved at trial. The majority (291; 80%) of cases were resolved by plea agreement, a process that differs significantly between jurisdictions. There are more systemic pressures, for example, for defendants to plead guilty in US jurisdictions than in Australian jurisdictions (for a more detailed discussion of the differences, see Brook, Fiannaca, Harvey, Marcus, & McEwan, 2016). As a result, only limited conclusions can be drawn from Pipe et al.’s (2013) study about the influence of the quality of the child witness interview on trial outcomes generally, and also specifically on trial outcomes in the Australian context.
Given these limitations, further research is necessary to investigate the extent to which adherence to best practice interviewing is related to trial outcomes. The current study aimed to address this issue. It was hypothesised that, once corroborative evidence was controlled for, 1) better-quality interviews (that is, interviews with a higher percentage of open-ended questions, fewer total numbers of questions, more desirable interviewer behaviours exhibited, and more evidential information) would be associated with more guilty verdicts, and 2) a greater number of cross-examination inconsistencies would be associated with fewer guilty verdicts. Although it is acknowledged that better evidence may also lead to a defendant being found “not guilty”, better quality interviews were predicted to be associated with more guilty verdicts for two reasons. First, the proportion of CSA cases that are successfully prosecuted is low compared to other offences (Australian Bureau of Statistics, 2004; NSW Bureau of Crime Statistics and Research, 2017), and second, research has suggested that this gap in successful prosecution rates is largely due to the insufficient, or poor quality, evidence that is available in these cases.

Method

**Trial and Child Witness Interview Transcripts**

Of the 85 complainants described in Study 1, eight were excluded because the outcome of their case was missing. This resulted in a final sample of 77 complainants, 16 male and 61 female, ranging in age from 6.87 to 18.78 years ($M = 13.83$, $SD = 3.30$).

**Procedure**

The trial transcripts of these 77 complainants were coded for child witness interview inconsistencies (see Study 3), outcome, and evidence characteristics
(see Study 1). The outcome for each complainant’s case was re-coded into a binary variable: “guilty on at least one count” and “other outcome” (not guilty on all counts, mistrial, or hung jury). The different types of corroborative evidence (medical or physical evidence and eyewitnesses) were collapsed into one category of “number of corroboration categories present”, given the small number of cases in which this type of evidence was present. “Number of victims” was retained as a separate category to quantify the precise number of victims present (rather than just the absence or presence of multiple victims). The term “victims” was used instead of “complainants” as the variable includes complainants, co-complainants, as well as non-complainant victims.

Child witness interviews were coded for four measures of interview quality, which were the same as those used in Study 1: interviewer questions (including total number of questions, and proportion of open-ended questions), interviewer behaviour, interview length, and investigative questions. From here on, the different interview quality variables and child witness interview inconsistencies will be referred to as the “variables of interest”. The “number of corroboration categories present” and “number of victims” variables will be referred to as the “control variables”. The control variables were chosen on the basis of past research; no other “evidence characteristics” variables were included in this study.

**Results**

Complete data was not available for all cases. The child witness interviews for 10 complainants were missing, and information about interview length was missing for a further 12 cases. Mean substitution and multiple imputation were considered as means to generate missing values, however, due to
the real-world implications of these findings both methods were deemed inappropriate for this dataset. Therefore, to maximise sample size, pairwise exclusion was utilised in analyses where necessary in place of deleting the entire case with missing data points.

A number of regression analyses were carried out to systematically investigate the influence of child witness interview quality and corroborative evidence on outcome. The results of these analyses are presented in two sections below. In the first section, logistic regression analyses were used to examine the relationship between each variable of interest and each control variable and trial outcome. Hierarchical logistic regressions were then conducted to investigate whether these relationships changed once the variance associated with the control variables was accounted for. In the second section, an overall model predicting the likelihood of at least one guilty verdict using both variables of interest and the control variables was developed. The variables included in the final model were chosen on theoretical grounds as well as on the relationships found in the first section.

**Individual Contributions of Variables**

To determine whether any of the variables were related to the outcome, binomial logistic regression analyses were conducted. As can be seen in Table 9.1, only the number of victims significantly predicted the outcome. The odds ratio suggests that with each additional victim, the chance of the defendant being found guilty of at least one charge doubles.
Table 9.1  
*Binomial Regression Analyses Between Each Variable and Outcome*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>N</th>
<th>Wald $\chi^2$</th>
<th>$p$</th>
<th>$B$ (SE)</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Victims</td>
<td>77</td>
<td>4.03</td>
<td>.05</td>
<td>0.78 (0.39)</td>
<td>2.17</td>
</tr>
<tr>
<td>Corroboration</td>
<td>77</td>
<td>2.22</td>
<td>.14</td>
<td>0.79 (0.53)</td>
<td>2.21</td>
</tr>
<tr>
<td><strong>Variables of interest</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Q</td>
<td>67</td>
<td>0.12</td>
<td>.73</td>
<td>0.00 (0.00)</td>
<td>1.00</td>
</tr>
<tr>
<td>Open Q</td>
<td>67</td>
<td>1.01</td>
<td>.32</td>
<td>-3.44 (3.42)</td>
<td>0.03</td>
</tr>
<tr>
<td>Specific Q</td>
<td>67</td>
<td>1.43</td>
<td>.23</td>
<td>2.95 (2.48)</td>
<td>19.16</td>
</tr>
<tr>
<td>Leading Q</td>
<td>67</td>
<td>0.33</td>
<td>.57</td>
<td>-1.97 (3.42)</td>
<td>0.14</td>
</tr>
<tr>
<td>Int Behav</td>
<td>67</td>
<td>0.00</td>
<td>.96</td>
<td>0.11 (2.32)</td>
<td>0.12</td>
</tr>
<tr>
<td>Evi</td>
<td>67</td>
<td>1.89</td>
<td>.17</td>
<td>1.85 (1.35)</td>
<td>6.34</td>
</tr>
<tr>
<td>Evi App</td>
<td>67</td>
<td>0.55</td>
<td>.46</td>
<td>-0.61 (0.82)</td>
<td>0.54</td>
</tr>
<tr>
<td>Length (mins)</td>
<td>55</td>
<td>0.71</td>
<td>.71</td>
<td>0.40 (0.01)</td>
<td>0.99</td>
</tr>
<tr>
<td>CWI inconsistencies</td>
<td>77</td>
<td>2.01</td>
<td>.16</td>
<td>-0.03 (0.02)</td>
<td>0.97</td>
</tr>
</tbody>
</table>

*Note:* Int Behav = desirable interviewer behaviours; Evi = proportion of evidence categories sought; Evi App = proportion of evidence categories sought appropriately; CWI = Child witness interview; Q = questions.

Next, to investigate whether any of the variables of interest were significant predictors of outcome *after* the influence of the control variables was taken into account, hierarchical binomial regressions were carried out. Both control variables were entered into the first block of the regression, and the variable of interest being investigated was entered into the second block. The first block of all models was significant, with model $\chi^2$ values ranging from 6.29 to
8.58, and significance values ranging from $p = .01$ to $p = .04$. The contributions of each variable of interest to the model (Block 2) are shown in Table 9.2.

### Table 9.2

Hierarchical Binomial Regression Analyses Between Each Variable of Interest and Outcome

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$N$</th>
<th>$\Delta$ Model $\chi^2$</th>
<th>$p$</th>
<th>Wald $\chi^2$</th>
<th>$p$</th>
<th>B (SE)</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Q</td>
<td>67</td>
<td>0.13</td>
<td>.72</td>
<td>0.13</td>
<td>.72</td>
<td>0.00 (0.00)</td>
<td>1.00</td>
</tr>
<tr>
<td>Open Q</td>
<td>67</td>
<td>0.54</td>
<td>.46</td>
<td>0.53</td>
<td>.47</td>
<td>-2.61 (3.59)</td>
<td>0.07</td>
</tr>
<tr>
<td>Specific Q</td>
<td>67</td>
<td>0.48</td>
<td>.49</td>
<td>0.47</td>
<td>.49</td>
<td>1.81 (2.64)</td>
<td>6.09</td>
</tr>
<tr>
<td>Leading Q</td>
<td>67</td>
<td>0.01</td>
<td>.91</td>
<td>0.01</td>
<td>.91</td>
<td>-0.42 (3.61)</td>
<td>0.66</td>
</tr>
<tr>
<td>Int Behav</td>
<td>67</td>
<td>0.11</td>
<td>.74</td>
<td>0.11</td>
<td>.74</td>
<td>-0.84 (2.52)</td>
<td>0.43</td>
</tr>
<tr>
<td>Evi</td>
<td>67</td>
<td>0.99</td>
<td>.32</td>
<td>0.95</td>
<td>.33</td>
<td>1.36 (1.39)</td>
<td>3.88</td>
</tr>
<tr>
<td>Evi App</td>
<td>67</td>
<td>0.13</td>
<td>.71</td>
<td>0.13</td>
<td>.71</td>
<td>-0.32 (0.87)</td>
<td>0.73</td>
</tr>
<tr>
<td>Length</td>
<td>55</td>
<td>1.50</td>
<td>.22</td>
<td>1.27</td>
<td>.26</td>
<td>-0.01 (0.01)</td>
<td>0.99</td>
</tr>
<tr>
<td>CWI inconsistencies</td>
<td>77</td>
<td>5.47</td>
<td>.02</td>
<td>3.40</td>
<td>.07</td>
<td>-0.05 (0.03)</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Note: Int Behav = desirable interviewer behaviours; Evi = proportion of evidence categories sought; Evi App = proportion of evidence categories sought appropriately; $Q =$ questions; CWI = Child witness interview; $\Delta$ Model $\chi^2 =$ change in the Model $\chi^2$ after adding the focus variable (Block 2).

As illustrated by the Wald statistics, none of the variables of interest made a significant contribution to the prediction of outcome in these hierarchical regressions. However, the change in Model $\chi^2$ showed that when “child witness interview inconsistencies” was added to the control variables, this addition significantly improved the predictive power of the model.

---

8 Model $\chi^2$ and significance values differed slightly in the first block due to the different $N$ values for two of the full hierarchical regressions.
Predicting Outcome Using Multiple Variables of Interest

Given the number of different interview quality variables, correlations were conducted to determine whether any of these variables were significantly correlated (and thus potentially redundant in further analyses). Correlation coefficients and associated significance values are shown in Table 9.3.

Table 9.3

<table>
<thead>
<tr>
<th></th>
<th>Total Q</th>
<th>Open Q</th>
<th>Specific Q</th>
<th>Leading Q</th>
<th>Int Behav</th>
<th>Evi</th>
<th>Evi App</th>
<th>Length (mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific Q</td>
<td>-.48 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading Q</td>
<td></td>
<td>-.06</td>
<td>-.79 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Int Behav</td>
<td></td>
<td>-.24 *</td>
<td></td>
<td>-.01</td>
<td>-.52 **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evi</td>
<td></td>
<td>.32 **</td>
<td>-.27 *</td>
<td>.37 **</td>
<td>-.30 *</td>
<td>.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evi App</td>
<td></td>
<td>-.22</td>
<td>.11</td>
<td>-.10</td>
<td>-.01</td>
<td>.20</td>
<td>-.02</td>
<td></td>
</tr>
<tr>
<td>Length (mins)</td>
<td>.67 **</td>
<td>-.50 **</td>
<td>.49 **</td>
<td>-.14</td>
<td>-.14</td>
<td>.32 *</td>
<td>-.27 *</td>
<td></td>
</tr>
</tbody>
</table>

Note: Int Behav = desirable interviewer behaviours; Evi = proportion of evidence categories sought; Evi App = proportion of evidence categories sought appropriately. ** p < .01, two-tailed. * p < .05, two-tailed. N = 74 for all variables except “length”; N = 60 for “length”.

As can be seen, a number of interview quality related variables were significantly correlated. Of note, higher proportions of open-ended questions (one key marker of a good interview) were significantly associated with shorter interviews (in terms of both the numbers of questions asked and the length of the interview in minutes). Higher proportions of open-ended questions were also associated with lower proportions of specific questions and fewer evidential

---

9 Only “interviewer behaviour” and “investigative questions asked appropriately” met the assumption of normality; thus, all correlation coefficients except the correlation between these two variables are Spearman correlations.
categories sought. Higher proportions of specific questions were associated with more evidential categories sought and interviews being longer (again in terms of both number of questions asked and minutes). As would be expected, the total number of questions in an interview was highly associated with interview length, and both these variables were positively associated with the number of evidential categories sought. The number of evidential categories sought appropriately was not correlated with any other variable.

Based on the correlations and the results of the previous logistic regressions, three variables of interest were included in the final model: “open-ended questions”, “number of evidential categories sought”, and “child witness interview inconsistencies”. “Total number of questions” and “length” were excluded as they were significantly correlated with the majority of interview quality variables. As was seen in Study 3, length is also associated with child witness interview inconsistencies. Of the remaining variables, “open-ended questions” was chosen because these questions are the most consistently defined question type (thus improving generalisability of any findings), and also the question type that differentiates best between effective and less effective interviewers (Powell et al., 2005). “Number of evidential categories sought” was chosen as this variable was developed with the aid of prosecutors, reflects evidential details necessary for a successful prosecution, and thus should theoretically be associated with successful prosecutions. This variable was chosen instead of “number of evidential categories sought appropriately” as the Wald statistics of both the straight binomial (Table 9.1) and binomial hierarchical (Table 9.2) regressions suggested that the former variable was the better predictor (though both predictors were non-significant). Finally, “child witness interview
“inconsistencies” was included as this variable significantly improved the predictive power of the model when entered after the control variables (see Table 9.2), and past literature has suggested inconsistencies have a significant influence on trials (see Study 3).

The full hierarchical regression model, including the three variables of interest and the control variables, was statistically significant $\chi^2(5) = 19.42, p < .01$, indicating that it was able to distinguish between complainants whose cases would, and would not, result in at least one guilty verdict. The model explained between 25.2% (Cox and Snell R square) and 33.6% (Nagelkerke R square) of the variance in outcome, and correctly classified 71.6% of cases. The change in model $\chi^2$ after adding the variables of interest was significant ($p = .01$), indicating that adding the additional predictors significantly improved the predictive power of the model. However, as shown in Table 9.4, only “number of victims” and “child witness interview inconsistencies” made unique statistically significant contributions to the model. The strongest predictor was number of victims, with an odds ratio of 4.98. This indicates that for every additional victim, the odds of at least one guilty verdict increased by about 5. Conversely, the odds ratio for child witness interview inconsistencies of 0.93 indicates that for each additional inconsistency, the odds of at least one guilty verdict decreased by 7%. In other words, for approximately every 14 inconsistencies, the odds of a complainant’s case resulting in at least one guilty verdict decreased by 1.
Table 9.4

*Hierarchical Logistic Regression Predicting Likelihood of At Least One Guilty Verdict (N = 67)*

<table>
<thead>
<tr>
<th></th>
<th>Wald</th>
<th>df</th>
<th>p</th>
<th>B (SE)</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Victims</td>
<td>7.01</td>
<td>1</td>
<td>.01</td>
<td>1.61 (.61)</td>
<td>4.98</td>
</tr>
<tr>
<td>Corroboration</td>
<td>1.48</td>
<td>1</td>
<td>.22</td>
<td>0.87(0.71)</td>
<td>2.38</td>
</tr>
<tr>
<td>Open Q</td>
<td>0.05</td>
<td>1</td>
<td>.83</td>
<td>-0.90 (4.22)</td>
<td>0.41</td>
</tr>
<tr>
<td>Evi</td>
<td>2.12</td>
<td>1</td>
<td>.15</td>
<td>2.36 (1.62)</td>
<td>10.61</td>
</tr>
<tr>
<td>CWI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inconsistencies</td>
<td>4.86</td>
<td>1</td>
<td>.03</td>
<td>-0.08 (0.04)</td>
<td>0.93</td>
</tr>
<tr>
<td>Constant</td>
<td>3.36</td>
<td>1.00</td>
<td>.07</td>
<td>-3.25 (1.77)</td>
<td>0.04</td>
</tr>
</tbody>
</table>

*Note: Evi = proportion of evidence categories sought; Q = Questions; CWI = Child witness interview.*

**Discussion**

The aim of this study was to investigate the extent to which adherence to best practice interviewing is related to trial outcomes in Australia. The hypothesis that better-quality interviews would be associated with more guilty verdicts (once corroborative evidence was controlled for), was not supported. However, the second hypothesis that a greater number of cross-examination inconsistencies would be associated with fewer guilty verdicts (once corroborative evidence was controlled for), was partially supported. Out of all variables examined, the number of victims was most predictive of outcome. Number of corroboration categories present was not significantly related to outcome, contrary to past literature (Blackwell & Seymour, 2014; Bradshaw & Marks, 1990; Lewis & Klettke, 2012). This may be due to the limited number of cases in the current study in which corroborative evidence was present.
Only one study has previously examined the relationship between best-practice interviewing and case outcomes. Pipe and colleagues (2013) found that cases using protocol-interviews (assumed to be of better quality) were associated with more charges and guilty verdicts at trial compared with pre-protocol interviews. The current study’s results do not directly support this conclusion. None of the interview quality variables examined (question types, interview length, interviewing behaviour, or evidential information sought) were significantly predictive of trial outcome – either on their own, when corroborative evidence was controlled for, or (in the case of open-ended questions and evidential information sought), when their individual contribution was examined in a model. Thus, one interpretation of the current results is that once a case proceeds to trial, interview quality is not a significant factor in determining trial outcome.

Although this finding is contrary to Pipe et al.’s (2013) conclusion regarding trial verdicts, that study’s results could be explained by their very small sample size of 23 trials, for which no information was available about the trials themselves. This resulted in the authors not being able to examine trial variables, such as how many incidents were charged, and what evidence was actually presented at trial (cf. what evidence was present in a police report; determinations of the admissibility of certain types of evidence are only made at trial). Additionally, the quality of interviews was not assessed, and some potentially confounding variables (e.g., corroborative evidence) were not taken into account when conducting significance tests.

An alternate interpretation of the current results is that – consistent with Pipe et al.’s (2013) conclusion – there was an association between interview
quality and trial outcome in the current study, but either a) the sample of interviews did not include enough variability in quality (i.e., they were all too poor) to detect it, or b) the relationship between interview quality and outcome is so small and/or indirect that the current study lacked the power to detect it. In terms of the first possible explanation, the range of the proportion of open-ended questions used in the interviews was indeed relatively small, which is a limitation of the current study. Nevertheless, other interview quality variables such as total number of questions should have had enough variability for a relationship to be detected – if one were present (see Study 1 for ranges, means, and standard deviations of all interview quality variables).

The second possible explanation is thus more likely. Given the range of factors that have the potential to influence the outcome of a trial (including case-related factors such as age of complainant, corroborative evidence, victims, type of abuse, and delay [see Study 1]; and legal factors, such as what evidence is ruled admissible, whether trials are held jointly or separately, how witnesses are cross-examined, and the type of judicial directions given: Australian Law Reform Commission et al., 2005; Victorian Law Reform Commission, 2004), the effect of child witness interview quality on outcome may be small or indirect. This interpretation is supported by the finding that a model incorporating a number of variables (corroborative evidence, number of victims, proportion of open-ended questions, proportion of evidential categories sought, and number of inconsistencies) significantly predicted verdicts, and correctly classified just under three-quarters of cases. Of note, although “number of victims” was a significant predictor in isolation, the change to the model was significant when
additional predictors were added, suggesting it is possible that interview quality variables had *some* influence on outcome, even if this influence was minor.

How interview quality may influence trial outcome in a more indirect way is demonstrated by the finding regarding inconsistencies originating from the child witness interview. Study 3 demonstrated that answers to open-ended questions were more likely to lead to inconsistencies during cross-examination than answers to either specific or leading questions. The current findings suggest that these inconsistencies, in turn, are related to trial outcome. In the full model, cross-examination inconsistencies were a significant predictor of at least one guilty verdict, with more inconsistencies resulting in decreased chances of a guilty verdict. This trend was also evident in the model that included just cross-examination inconsistencies and the control variables, although as a sole predictor in that model it did not quite reach significance. These findings support concerns expressed by prosecutors in previous studies, namely that lengthy and detailed interviews “inhibit the State’s prospects of conviction because the inconsistencies [resulting from these details] are used to cast doubt on the evidence” (Burrows & Powell, 2014b, p. 195; see also Burrows & Powell, 2014d; Burrows, Powell, & Benson, 2016). It also supports findings from studies with mock jurors, which have suggested that jurors consider inconsistencies in witness statements as strong markers of inaccurate evidence (Berman et al., 1995; N. Brewer et al., 1999; Connolly et al., 2008), and supports the notion that raising inconsistencies is an effective cross-examination technique.

The implications of these findings are twofold. First, they highlight the need for interviewers to be educated regarding the potential influence of child witness interviews in court. In particular, interviewers need to be made aware of
how complainants’ evidence is tested during cross-examination, and how inconsistencies raised here can influence the outcome of the case. Second, the findings suggest that educating legal professionals and jury members about the nature of inconsistencies in testimony may be beneficial in reducing the impact of inconsistencies on trial outcomes.

The results of the current study should be interpreted in light of the fact that only cases that reached trial were evaluated. Thus, in line with Pipe et al. (2013), interview quality may have a more significant role to play in determining which cases actually reached trial. For instance, cases in which interviewers failed to seek critical evidential details may have dropped out of the system prior to trial. Such attrition could explain why the number of evidential categories sought was not a significant predictor of trial outcome – all cases that reached trial had been assessed as having sufficient evidence for a successful prosecution. Another interview variable that could have been affected in this manner is interviewer behaviours; these behaviours are recommended on the basis that they encourage children to provide their best possible statements (American Professional Society on the Abuse of Children, 2012; Benson & Powell, 2015b; Lamb, 2016; Powell & Snow, 2007). Thus, if too few of them were utilised, it is possible that the quality of statements was affected and the case did therefore not proceed to trial.

In conclusion, the current study did not find evidence of a direct relationship between interview quality and trial outcomes. However, results suggest that some aspects of interview quality may indirectly influence outcomes, namely in the form of inconsistencies raised during cross-examination, which were linked both to interview questioning and trial outcomes. The findings suggest that when determining the relationship between the child witness
interview and the legal process, it is necessary to examine the trial as a whole, rather than just trial outcomes. Future studies should replicate the current findings, ideally with interviews ranging in quality on all key indicators of best-practice interviewing.
CHAPTER 10: GENERAL DISCUSSION

The aim of this thesis was to examine the relationship between investigative interview quality, CSA trial process, and trial outcome. The broader objective was to determine whether problems with the quality of child witness interviews could account for some of the difficulties experienced prosecuting these cases. CSA is one of the most difficult crimes to prosecute and it suffers from high attrition rates (Australian Law Reform Commission, 2010; Eastwood et al., 2006). The low rate of successful prosecution may partially be due to the type of evidence that is available in these cases; there are rarely any eyewitnesses (Australasian Institute of Judicial Administration, 2012; Cossins, 2000), and corroborating evidence, such as medical evidence of penetration, physical injury, or the presence of semen is also usually absent (Australasian Institute of Judicial Administration, 2012; Cossins, 2000; Gothard, 1987; Oates, 2007; W. A. Walsh et al., 2010). As a result, the complainant’s evidence – usually given as their child witness interview – is central to these cases (Burrows & Powell, 2015). It is thus conceivable that interview quality might influence prosecutions and trials. However, prior to this thesis, no studies had investigated how these interviews are utilised in court, and there was only limited research examining how – or whether – they influence CSA trials.

Four original studies were presented that examined different facets of this complex relationship. Study 1 provided a detailed and current description of the CSA cases currently being prosecuted in Australian courts, including an analysis of the quality of the child witness interviews in these cases, and the relationship between these interviews with other case characteristics. Study 2 examined how
legal professionals discuss child witness interviews during CSA trials, with a particular focus on the interviews’ usefulness as evidence-in-chief. Study 3 investigated whether questioning in investigative interviews is related to cross-examination inconsistencies, and explored the nature of the inconsistencies raised; and finally, Study 4 examined the extent to which adherence to best practice interviewing is related to trial outcomes. The current and final chapter in this thesis presents a global discussion of the key findings of these studies. Findings will be discussed in three sections: 1) the relationship between the child witness interview and the trial process; 2) the relationship between the child witness interview and cross-examination; and 3) the relationship between the child witness interview and case outcomes. The chapter will conclude with an overall discussion of limitations, implications for policy, and recommendations for future research.

The Relationship Between the Child Witness Interview and the Trial Process

An examination of the characteristics of cases that reached trial suggested that, overall, child witness interview quality (as defined by interviewing experts) did not play a significant role in determining which cases were prosecuted. An important caveat here, of course, is that no data was available about cases not accepted for prosecution, so a direct comparison was not possible. However, a number of findings support this conclusion. First, the overall quality of the interviews that reached trial was very poor (Study 1). The proportion of open-ended questions was low, only slightly higher than the proportion of leading questions, while the vast majority of questions were specific. This low proportion of open-ended questions is contrary to interviewing guidelines, which suggest that interviewers should maximise the use of open-ended questions, only utilise
specific questions once open-ended questions have been exhausted, and avoid leading questions altogether (Lamb & Fauchier, 2001; Powell, 2002; Powell & Snow, 2007). If interview quality was a determining factor in deciding which cases reached trial, it would be reasonable to expect the quality of the interviews that reached trial to be slightly better.

Further, the proportion of open-ended questions asked was not significantly related to any of the case characteristics examined (average age at child witness interview, type of case in terms of worst offence, repeated offence or not, corroborative evidence present or not; Study 1). The same was true for three other interview quality variables examined (proportion of interviewer behaviours exhibited, the length of the interview per occasion, and the proportion of investigative questions asked appropriately). If interview quality did influence decisions to prosecute, a relationship between some of these variables might be expected. For example, a relationship between the presence of corroborative evidence and interview quality could be expected, so that cases with no corroborative evidence were associated with better interviews. Similarly, cases with younger children and cases with less severe charges – all factors that have been found to be associated with a lessened likelihood of prosecution (Bunting, 2008; Christensen et al., 2014; Davis et al., 1999; Fitzgerald, 2006; Leach et al., 2016) – might logically be expected to require better quality interviews to improve the strength of the case. However, this expectation was not supported by the data. Thus, it appears that interview quality was not a factor in the decision of whether or not to accept a case for prosecution.

Although interview quality did not appear to influence decisions about whether to accept cases for prosecution, it was raised as an issue during the trials
themselves. When discussions between legal professionals at trial were examined in Study 2, they suggested that judges, prosecutors, and defence lawyers were aware of some issues with the quality of interviews in their trials. One key concern was that interviews were too long, leading to complainants forgetting their evidence, as well as complainant and juror fatigue. In one notable case, a judge halted the complainant’s evidence-in-chief and ordered a break after a juror fell asleep watching a complainant’s long interview. These concerns about interview length support findings from a previous focus group study with prosecutors, which had also suggested that interview length is a significant issue at trial (Burrows & Powell, 2014b). An analysis of interview length showed that interviews in the current sample were indeed quite long, lasting one hour, and with complainants being asked over 250 questions on average (Study 1). In comparison, one recent study found that interview times could be reduced to 40 minutes on average, without any loss of important evidential information (Benson & Powell, 2015b). Such a reduction in interview length would likely lead to significant improvements for complainants and jurors at trial.

A further critique of interview quality that emanated from trial discussions was that interviews were poorly structured and unclear (Study 2). Legal professionals were concerned that interviews did not follow any chronological order, did not clearly relate to the charges on the indictment, and jumped “from place to place”. The consequences of this lack of structure included confusion for both complainants and legal professionals, both during cross-examination and voir dires. One judge expressed specific frustration at interviewers “continually” interrupting the children’s narratives and moving to irrelevant topics at pivotal times – for example, just after a disclosure of abuse. These findings are consistent
with past studies, in which prosecutors also expressed concerns about interviewers interrupting children’s disclosures of abuse; for example, by asking irrelevant questions about contextual details (Burrows & Powell, 2014b, 2014d). The concerns about interview structure that arose during trial discussions were backed up by the analysis of interview quality, which showed that only one third of desirable interviewer behaviours – including sticking to one occurrence at a time, and allowing the child to talk without interruption – were exhibited in the current sample on average (Study 1). Thus, although some aspects of interview quality were not mentioned as being problematic (such as the low proportion of open-ended questions), there is evidence that other aspects – length and structure – did have undesirable downstream effects on trial proceedings.

One way in which issues with interviews may be resolved is through editing the recordings prior to use at trial and through the use of supplementary questioning by prosecutors (AIJA, 2012; Corns, 2001; Freckelton & Selby, 2013). Such editing was done in the vast majority of cases in the current sample (Study 1); however, trial discussions suggest that edits were not made for interview quality purposes but rather to remove inadmissible evidence (Study 2). Of those edits that were discussed at trial, the majority concerned the removal of uncharged acts evidence – that is, evidence of the accused’s motive, or tendency to act or think in a particular manner, or evidence concerning the accused’s relationship to the complainant. This finding directly contradicts the recommendations made in a recent study with Australian prosecutors, which argued for the routine inclusion of this type of evidence in child witness interviews (Darwinkel, Powell, & Tidmarsh, 2014). Given how frequently such evidence was edited out of interviews in the current sample, the suggestion to include it more frequently is
problematic. Based on the current findings, (and without any change in the law) it is likely that any such evidence collected in interviews will either be edited out prior to trial, or lead to more legal arguments at trial lengthening the proceedings.

Despite the ability to edit interviews before trial, it is imperative that interviewers remember that anything that is recorded during the interview could be used in court. This includes children’s behaviour during breaks, and children’s demeanour, as was demonstrated in Study 2. Here, the discussions between legal professionals showed that children’s actions and demeanour during an interview were often used to judge the veracity of their claims. At times, their behaviour was also used as basis for the defence to argue for a reliability warning to be given. Prosecutors had previously voiced concerns regarding the reduced formality of interviews, suggesting that this might undermine children’s perceived credibility in the eyes of a jury (Burrows & Powell, 2014c); the current findings support these concerns and suggest that interviewers should also be mindful of children’s behaviour both during the interview and during breaks. Indeed, children’s demeanour and behaviour forms an important part of the jury’s assessment of their credibility, and judges will instruct juries at the conclusion of the trial to take this into account in reaching its decision.

**Relationship Between the Child Witness Interview and Cross-Examination**

The part of a trial where the child witness interview seemed to have the most significant influence (apart from evidence-in-chief) appeared to be cross-examination. A number of findings suggest that the behaviour and questioning strategies employed by interviewers have a significant relationship with the complainant’s cross-examination experience. For example, Study 2 highlighted a number of cases in which mistakes made by the interviewers (such as putting the
wrong version of events to the complainant and mixing up details across incidents of abuse) resulted in confusion for the complainant during cross-examination. The examination of discussions between legal professionals suggested that the defence will use such mistakes to argue that the child’s account is unreliable, despite the fact that it was the interviewer who made the mistake in the first place.

The type of information sought by the interviewer also appeared to have an influence during cross-examination. Past studies have suggested that interviewers tend to seek too much detail and irrelevant information from children, which is then used to undermine complainants’ accounts during cross-examination (Burrows & Powell, 2013, 2014b, 2014d). Current findings support this supposition. Although interviewers sought most of the “important evidentiary information” (approximately 80%, see Study 1), half of this information was sought inappropriately (that is, it was sought using developmentally inappropriate language and concepts, and included questioning about irrelevant, minutiae details). Taken together with the findings that there were significantly more peripheral than central child witness interview inconsistencies during cross-examination (Study 3), and that over one-fifth of inconsistencies concerned topics not necessary from an evidential perspective (Study 3), these results suggest that the questioning around minute or irrelevant details in the interviews was an important contributing factor to many of the child witness interview inconsistencies raised during cross-examination. This is because peripheral or minute details are less likely to be remembered than central details (Brown et al., 2008; Toth & Valentino, 2008), are thus more likely to be contradicted by children in a retelling of their story, and can consequently be raised by the defence as a peripheral inconsistency during cross-examination. Peripheral details
in interviews also provide the perfect opportunity for the defence to deliberately create inconsistencies during cross-examination, in an effort to undermine the credibility of the complainant’s account (E. Davies & Seymour, 1998; Hanna et al., 2012; Kebbell et al., 2007).

Further, the amount of information collected in an interview played a significant role in the complainants’ cross-examination experience. As mentioned above, Study 1 showed that interviews were quite long on average. Sometimes this length resulted in complainants forgetting aspects of their evidence, meaning that they could not be effectively cross-examined on it, and might contradict something they had previously said (Study 2). Complainants also underwent long cross-examinations, during which they were asked almost 500 questions on average by defence lawyers (Study 1). One key tactic employed by the defence was raising inconsistencies; a total of 1087 inconsistencies were raised solely between the child witness interview and at least one other source for the sample of 73 complainants analysed in Study 3. Findings revealed that these three variables are connected; as the number of child witness interview inconsistencies increased, so did the length of cross-examination (Study 3). Similarly, when number of incidents was controlled for, the number of child witness interview inconsistencies were significantly associated with the overall length of the interview (Study 3). Notably, however, the number of child witness interview inconsistencies was not related to the total number of questions asked in an interview (Study 3). These findings suggest that the amount of information originally sought in an interview determines how long a complainant is likely to be cross-examined for, and how many inconsistencies the defence will raise.
Reducing the length of the interview is thus an avenue that should be targeted to improve the complainants’ cross-examination experience, or at least shorten it.

Consistent with this supposition that it is the amount of information contained in a child witness interview that is most important in determining the number of child witness interview inconsistencies raised during cross-examination, Study 3 showed that open-ended questions were more likely than either specific or leading questions to elicit information that would result in an inconsistency. There were no significant differences between leading and specific questions in resulting in inconsistencies. These findings were contrary to what was hypothesised based on the interviewing literature, where open-ended questions are recommended above specific questions. The most likely explanation for this finding is that open-ended questions led to more inconsistencies because (consistent with past research, see e.g., Sternberg et al., 1996; Hershkowitz, 2001) answers to these questions were longer and provided more information that could be used by the defence to create or highlight inconsistencies during cross-examination. This finding is significant, as it suggests that maximising the evidential quality of an interview cannot be achieved simply through increasing open-ended question usage, as has been suggested (Westera et al., 2017). More details in an interview do not equate to a better interview from an evidential perspective (cf. E. Davies & Seymour, 1998; Hershkowitz, 2001; Westera et al., 2013b). In line with findings from past studies with prosecutors (Burrows & Powell, 2014b, 2015), current results show that interviewers must be extremely mindful about what information is collected, and how it is collected. In particular, they support the assertion that broad open-ended questions may be just as damaging as too many specific questions (Burrows & Powell, 2014b), and
suggest a move towards recommending more focused open-ended questions. This suggestion will be further explored below.

The Relationship Between the Child Witness Interview and Trial Outcome

The final relationship examined in this thesis was that between the child witness interview and trial outcomes. Verdicts were available for 77 complainants; of these, guilty verdicts resulted in at least one charge in around 40% of complainants’ cases (Study 1). Given that past research has suggested corroborative evidence - including eyewitnesses, physical evidence, number of victims, and medical evidence – significantly improves the chances of a guilty verdict (Blackwell & Seymour, 2014; Bradshaw & Marks, 1990; Palusci et al., 1999), corroborative evidence was taken into account in determining what, if any, influence the quality of the child witness interview had on conviction rates in the current sample. Most cases did not have any corroborating physical evidence, and eyewitnesses and non-complainant victims were very rare (Study 1).

Nevertheless, the number of victims was the strongest predictor of outcome in all analyses (Study 4). The number of corroboration categories present was not significantly related to outcome in any of the analyses, which is not consistent with the findings from past research (Blackwell & Seymour, 2014; Bradshaw & Marks, 1990; Lewis & Klettke, 2012).

The contribution of the interview quality variables (introduced in Study 1) on outcome was examined in three scenarios: individually, while controlling for the influence of the number of corroboration categories present and number of victims, and collectively (in the case of open-ended questions and number of evidential categories sought). None of the variables was significantly predictive of outcome, nor made a significant individual contribution to the model, in any of
these scenarios (Study 4). This finding is not consistent with the only other study that has so far examined the influence of interview quality on trial outcomes (Pipe et al., 2013). Pipe et al. suggested that improved interview quality led to more guilty verdicts once a number of trial characteristics (including number of victims) were controlled for. However, these researchers did not code the actual quality of the interviews in their sample, did not take into account corroborative evidence, and based their conclusion about increased guilty verdicts on a very small sample of 23 trials. These are significant limitations and may explain the discrepancy with current results.

Next, the relationship between child witness interview inconsistencies (which had been coded for in Study 2) and trial outcome was examined. Findings revealed that when child witness interview inconsistencies were added after the control variables (corroborative evidence and number of victims), the predictive power of the model increased significantly (Study 4). Further, when the two selected interview quality variables and child witness interview inconsistencies were entered into a full hierarchical regression model (in which the control variables were entered first), child witness interview inconsistencies made a unique statistically significant contribution to the model. These findings suggest that child witness interview quality may influence trial outcomes indirectly; that is, interview quality influences how many inconsistencies are raised during cross-examination which, in turn, influence trial outcomes. Prosecutors had previously expressed concerns that lengthy and detailed interviews negatively influence prosecutions as they allow defence lawyers to undermine complainants’ evidence through the use of inconsistencies (Burrows & Powell, 2014b, 2014d; Burrows et al., 2016); current findings provide support for this claim.
Finally, the predictive power of three variables of interest (open-ended questions, number of evidential categories sought, and child witness interview inconsistencies) was examined collectively while controlling for corroborative evidence and number of victims. Given the range of factors – including case-related and legal factors – that have the potential to influence the outcome of a trial, it was of interest to determine whether a combination of variables related to the interview could predict outcomes. The full model significantly predicted verdicts, and correctly classified just under three-quarters of cases (Study 4). Further, adding the variables of interest significantly improved the predictive power of the model, indicating that collectively these variables had some influence on outcome. However, given that the two child witness interview quality variables did not have a significant individual influence on outcome, any contribution to trial outcome they might have had is likely to have been small.

**Limitations and Future Research**

There are a number of general limitations that are applicable to all studies presented in this thesis. These will now be discussed, alongside possible avenues for future research. First, as described in Study 1, the overall quality of interviews utilised in this thesis was generally poor. Although consistent with past research that has shown that field interviews generally do not adhere to best-practice guidelines (Lamb, 2016; Luther et al., 2015; Powell, Cavezza, et al., 2010; Powell, Murfett, et al., 2010; Wolfman et al., 2016), the generally poor quality of the current interviews could account for some of the non-significant findings in this thesis. For example, better quality interviews might be associated with more guilty verdicts, but given too few “good” interviews were present in the sample, the association was not detected. Further, prosecutors may have accepted cases
with poor interviews for prosecution simply because most cases passed on to the ODPP had poor quality interviews. Future research is thus needed examining the influence of interview quality on trials and outcomes using a sample of interviews with a broader range of quality. Future research should also include an evaluation of cases rejected for prosecution, and compare the interview quality in the “rejected” sample with interview quality in an “accepted” sample. Such research is necessary because, in line with Pipe et al. (2013), interview quality may be more influential prior to trial than at trial. Any such research will, however, first require interview quality to improve in the field which, in turn, will require adequate interviewer training programs to be adopted in all relevant organisations (Lamb, 2016). Such a feat has so far eluded researchers and policy-makers alike.

A second general limitation of this thesis is the low number of trials included in the studies, and the missing data (such as missing interviews and outcomes) within the sample. The trial and interview transcripts utilised for this thesis were obtained as part of a larger research study examining the effectiveness of alternative measures in CSA cases (see Powell, Westera, et al., 2016). Due to time constraints, resourcing issues, and different legal requirements across jurisdictions and courts, a larger or more complete dataset could not be obtained. This limitation affected Studies 3 and 4 in particular. For example, too few cases were available with more than one interview to compare question types for inconsistencies that arose between two interviews in Study 3, and only a limited number of predictors could be entered into the final model to predict case outcome in Study 4. These are areas for future research; in particular, re-examining the influence of different aspects of interview quality on case outcome in a larger dataset, while controlling for more confounding variables. Given the
current findings suggest that the influence of interview quality on outcome is not 
linear, but rather mediated by cross-examination inconsistencies and possibly 
other aspects of the trial, this indirect relationship could also be investigated 
statistically. Such an analysis would also require a larger dataset to achieve 
adequate power (Fritz & MacKinnon, 2007).

**Implications for Practice and Research**

**Recommendations for best-practice interviewing.** Despite these 
limitations, there are a number of important implications for researchers, policy 
makers, interviewers, and trainers arising out of the current findings. First, all 
studies show that reforms to current interviewing practices are needed to improve 
complainants’ experiences at trial, particularly complainants’ experience of cross-
examination. In line with recommendations in past research (Benson & Powell, 
2015b; Burrows & Powell, 2014b; Powell et al., 2007), current results show that 
interviews need to be structured in a logical order; each abusive occasion should 
be exhausted before moving on to the next, and interviewers should refrain from 
jumping between occurrences. Interviews also need to be kept short and focused 
around central details, and interviewers need to reduce the number of minute 
details that they request of children (Burrows & Powell, 2013, 2014c). If adopted, 
these recommendations should improve complainants’ memory of their evidence 
during cross-examination, and reduce the number of inconsistencies accidentally 
created. Reducing the length and amount of detail in interviews should also 
reduce the amount of information with which defence lawyers could attempt to 
“create” inconsistencies during cross-examination, and should also reduce the 
length of cross-examination overall. Finally, such changes should also aid jurors 
trying to follow the complainant’s story, and should help reduce complainant and
juror fatigue. Translating these recommendations into practice may necessitate some form of prosecutor instruction to clarify what information is required from an evidential perspective, and what sort of information can be left out (Burrows et al., 2013).

Second, there is strong evidence from the current studies showing that interviewing guidelines regarding the use of open-ended questions need to be modified. Clearly the wealth of evidence regarding the superiority of open-ended over closed questions in leading to reliable and accurate answers cannot be ignored (Brown et al., 2013; Dent & Stephenson, 1979; Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007; Powell & Thomson, 2001). However, it has also been consistently shown that open-ended questions result in much longer and detailed answers than other question types (Brown et al., 2013; G. Davies et al., 2000; Hershkowitz, 2001; Phillips et al., 2012; Sternberg et al., 1996), which is problematic as the current findings show that too much information (especially minute and peripheral details) in interviews is actually damaging to both complainants and the prosecution. This finding is in line with past research in which prosecutors described broad open-ended questions that elicited lengthy and irrelevant details as problematic (Burrows & Powell, 2014b). Thus, in order to reduce the number of inconsistencies raised during cross-examination, but still maintain the integrity of the evidence, best-practice interviewing guidelines should start recommending the use of more focused open-ended questions over the use of broad open-ended questions. Focused open-ended questions (e.g., cued, breadth or depth prompts) remind children of actions or details previously mentioned (Brown & Lamb, 2015; Gagnon & Cyr, 2017; Lamb et al., 2003), and involve questions such as, “tell me more about the part when he touched your
leg”, rather than questions like, “tell me everything that happened”. Focused open-ended questions are already recommended for certain groups of witnesses such as young children and children with intellectual disabilities (Brown, Lewis, Lamb, & Stephens, 2012; Gagnon & Cyr, 2017; Lamb et al., 2003). Thus, encouraging more focused open-ended questions to be used with all children should maintain the advantages of open-ended questions as demonstrated in the interviewing literature, while hopefully improving complainants’ experiences in court.

**Could expert evidence help?** Expert evidence is an exception to the rule against the admission of opinion evidence (see e.g., s. 79 Evidence Act 2008 (Vic); s. 79 Evidence Act 1995 (NSW)). This exception means that in certain circumstances (which are determined by the courts) people who are deemed to be experts in their field may give evidence to educate the court on matters that are not “common knowledge” (Australian Law Reform Commission et al., 2005; Freckelton & Selby, 2013). One way in which the impact of inconsistencies may thus potentially be mitigated is through expert evidence on human memory. However, such evidence is currently generally inadmissible in Australia¹⁰ (Freckelton & Selby, 2013), thus a change in the law would be required before experts could be called to assist prosecutions (or the defence) in these cases. Expert evidence is also controversial due to concerns that experts may usurp the role of the jury, and the possibility that trials may turn into “battlegrounds” for defence and prosecution experts (G. Davies, 2005; Heydon, 2017). Therefore, although a possible avenue to be considered in the future, expert evidence does

---

¹⁰ There are some exceptions in which expert evidence on memory is admissible, including where there are concerns about reliability about the memory of children who are under three years old, and eyewitness identification evidence; see Freckelton & Selby, 2013, for a more detailed discussion.
not currently appear to be a useful strategy to reduce the impact of inconsistencies in these cases.

**The importance of recognising the adversarial nature of our legal system.** In assessing the utility of child witness interviews, one aspect that has not received much attention in the interviewing literature to date is the influence of the legal system in which the interview is used. Current findings suggest that it is incredibly important for both researchers and interviewers to be aware of the key characteristics of the system in which they are working. In inquisitorial systems, there is a strong preference for documentary evidence (Cordon et al., 2003), and the case is decided by a judge. Complainants generally do not undergo cross-examination, and all relevant evidence is presented to the judge who assesses its importance to the case (Cordon et al., 2003; Finn, 2009). Defendants usually do not have the right to confront complainants in CSA trials (Cordon et al., 2003; Myklebust, 2017) and, in comparison to adversarial trials, complainants’ involvement in the process frequently ends after they are interviewed by the judge before the trial (Myklebust, 2017). This process is significantly different to the adversarial legal system, in which there is a strong focus on oral evidence (Cossins, 2010), and the case is decided by a jury. As a result, there are strict rules of evidence, and cross-examination is used to test the complainant’s evidence (Cordon et al., 2003). Because the case is decided on the evidence presented by the prosecution and defence, and there is no independent inquiry by the court, the manner in which the evidence is given in court – and whether the complainant’s case appears credible to the jury – is crucial in determining the outcome of a case (Cordon et al., 2003). Children’s involvement does not end with their
investigative interview; their interview is played in court and is subject to cross-examination in front of a jury.

Given these differences, what amounts to a “good” child witness interview for evidentiary purposes arguably differs between adversarial and inquisitorial criminal trials. For example, it is conceivable that in a system without cross-examination, longer responses to questions in investigative interviews have a favourable effect on court decisions (Myklebust & Bjørklund, 2009); in contrast, when interviews are to be subject to cross-examination, too much information in interviews may have a negative effect on prosecutions (Study 3 and 4; Burrows & Powell, 2014b, 2014d; Burrows, Powell, & Benson, 2016). Differences will also exist between what information should be included in interviews, given the fact that there are strict rules of evidence in adversarial criminal trials, and no such rules in inquisitorial trials. Future research investigating the effect of child witness interviews on trials and outcomes must therefore differentiate between the effect of these interviews in inquisitorial trials, and the effect of these interviews in adversarial trials. Practitioners must also be aware of the environment in which their interview will be used, particularly in adversarial countries where complainants undergo cross-examination questioning seeking to discredit them and their testimony (Bowden et al., 2014; Cossins, 2009; Zydervelt et al., 2017).

The need for increased interdisciplinary communication and research. The area of investigative interviewing bridges (at least) two disciplines; law and psychology. However, most research that established the principles behind best-practice interviewing of children was undertaken from child development, language, and memory perspectives; legal professionals have had little influence in the development of these guidelines (Burrows & Powell, 2014b,
2014c). This is despite the fact that legal professionals are the end-users of interviews, which frequently form the central piece of evidence in CSA prosecutions (Powell, Murfett, et al., 2010). It is likely that many of the problems highlighted in this thesis regarding the quality of interviews have arisen because investigative interviewers and trainers are unaware of what makes a “good” interview from a legal perspective, nor how interviews are used in court (Burrows et al., 2013; Guadagno et al., 2006; Hoyano & Keenan, 2010; McConachy, 2002). Moving forward, it will thus be crucial to improve interdisciplinary communication, feedback and collaboration between legal professionals and interviewers, to maximise the utility and quality of investigative interviews as evidence-in-chief (McConachy, 2002; Powell, 2008; Victorian Law Reform Commission, 2004). This may be achieved through more interdisciplinary conferences and workshops, where professionals from both disciplines are exposed to alternate view points, and current issues can be discussed and possibly resolved.

**Concluding Comment**

In conclusion, the relationship between child witness interviews and CSA trials is complex. Current findings suggest that once a case reaches trial, interviews have the most influence during cross-examination, where interview quality affects how many inconsistencies are raised which, in turn, influences case outcomes. Interview quality appears to have little direct effect on trial outcomes, likely because there are numerous factors that can influence trials. Perhaps, then, moving onward researchers and investigative interviewers should not focus solely on the ability of interviews to influence case outcomes. Instead, given current results, it may be more advantageous to focus more broadly on the consequences
of interview quality for complainants’ at trial, and ensure that their interview supports – rather than undermines – them during cross-examination.
References


Criminal Procedure Act 1986 (NSW) (Austl.).

Criminal Procedure Act 2009 (Vic) (Austl.).


*Evidence Act 1906 (WA) (Austl.).*

*Evidence Act 1995 (NSW) (Austl.).*
Evidence Act 2008 (Vic) (Austl.).


*Longman v The Queen* 1989 168 CLR 79 (Austl.).


Magistrates’ Court and the Children’s Court of Victoria (2010). Submission to the Australian Law Reform Commission (ALRC FV 220, 1 July 2010). Sydney, Australia: ALRC.


Powell, M. B. (2002). Specialist training in investigative and evidential interviewing: Is it having any effect on the behaviour of professionals in


*S v. The Queen 1989* 168 CLR 266 (Austl.).


Westera, N. J., Kebbell, M. R., & Milne, B. (2013a). It is better, but does it look better? Prosecutor perceptions of using rape complainant investigative


